Variability of competitive performance of distance runners

WILLIAM G. HOPKINS and DAVID J. HEWSON

Department of Physiology and School of Physical Education, University of Otago, Dunedin, NEW ZEALAND

ABSTRACT

HOPKINS, W. G., and D. J. HEWSON. Variability of competitive performance of distance runners. Med. Sci. Sports Exerc., Vol. 33, No. 9, 2001, pp. 1588-1592. Purpose: The typical variation in an athlete's performance from race to race sets a benchmark for assessing the utility of performance tests and the magnitude of factors affecting medal prospects. We report here the typical variation in competitive performance of endurance runners. Methods: Repeated-measures analysis of log-transformed official race times provided the typical within-athlete variation in performance as coefficients of variation (CV). The types of race were cross-country runs (4 races over 9 wk), summer road runs (5 races over 4 wk), winter road runs (4 races over 9 wk), half marathons (3 races over 13 wk and 2 races over 22 wk), and marathons (2 races over 22 wk). Results: Typical variation of times for the fastest quartile of male runners was 1.2-1.9% in the cross-country and road runs, 2.7% and 4.2% in half marathons, and 2.6% in marathons. Times for the slower half of runners in most events were more variable than those of the faster half (ratio of slower/faster CV, 1.0-2.3). Times of younger adult runners were more variable than times of older runners (ratio of younger/older CV, 1.1-1.8). Times of male runners were generally more variable than those of female runners (ratio of male/female CV, 0.9-1.7). Conclusion: Tests of endurance power suitable for assessing the smallest worthwhile changes in running performance for top runners need $CV \le 2.5\%$ and $\le 1.5\%$ for tests simulating half or full marathons and shorter running races, respectively. Most of the differences in variability of race times between types of race, ability groups, age groups, and sexes probably arise from differences in competitive experience and attitude toward competing. Key Words: COEFFICIENT OF VARIATION, ELITE ATHLETES, FEMALE RUNNER, RELIABILITY, REPRODUCIBILITY, TEST-ING AND EVALUATION

The variability in a top athlete's performance from competition to competition is one of the key factors when the athlete's prospect of a medal is under consideration. In a competition against a handful of closely matched opponents, the smallest change in performance that has a substantial effect on the athlete's chance of winning is ~0.3 of this variability, expressed as a standard deviation (3). Researching such changes in performance with a laboratory or field test is practically impossible if variability of performance in the test is greater than that of performance in races, because the sample sizes have to be too large (3). The test would also have little or no practical value for monitoring these small changes in performance in an individual athlete. Clearly, information about the variability of competitive performance would be useful to researchers and practitioners who are concerned with factors that affect performance. Such information is now available for swimmers competing at national level (6). The purpose of the present study was to provide a similar analysis of competitive performance of distance runners.

0195-9131/01/3309-1588/\$3.00/0

MEDICINE & SCIENCE IN SPORTS & EXERCISE_® Copyright © 2001 by the American College of Sports Medicine

Submitted for publication August 2000. Accepted for publication November 2000.

METHODS

Subjects and races. Analysis of variability in performance requires race times for a sufficient number of athletes who have entered two or more races in a series. We focused on consecutive races in a series within a competitive season, because this time frame appears to be of most interest to researchers, athletes, and support professionals interested in factors that modify performance.

Competition organizers provided official times for runners in races ranging from club through national level. These data are in the public domain, so we did not seek written consent for their use from individual athletes. Data suitable for analysis were available for cross-country runs, summer and winter road runs, half marathons, and marathons. For most races, athletes competed in age groups, sometimes over different distances. Table 1 summarizes the race series, numbers of athletes, and performance times for athletes who competed in two or more races in each series.

Statistical analysis. We used the mixed linear modeling procedure (Proc Mixed) in the Statistical Analysis System (Version 6.12, SAS Institute, Cary, NC) to derive estimates of variability of performance for each race series. The dependent variable was the natural log of the time in an event; analysis of this transformed variable yields coefficients of variation (CV), which are variations in performance expressed as a percent of average performance (1). For each analysis, we modeled a within-athlete (error) variance and a between-athlete or true-score variance (free of error). The between-athlete CV shown in Table 1 represent TABLE 1. Performance time over selected distances for female and male runners who competed in two or more races within a given age group and type of event.

	No. of Athletes ^a in:		Distance	Time ^b	CV ^b
Age (yr)	≥1 Race	≥2 Races	(m)	(min)	(%)
Interclub—national cross-country runs (4 races over 9 wk)					
Female runners (2500–6000 m)					
<16	39	13	3000	15	21
16–19	82	7	4000	18	11
≥20	99	24	6000	28	13
Male runners (2500–12,000 m)					
<16	48	10	4000	16	14
16–19	104	12	8000	28	7
≥20	277	66	12,000	45	12
Interclub summer road runs (5 races over 4 wk)					
Female runners (5000–8000 m)					
20–39	14	8	8000	35	14
≥40	8	7	8000	37	11
Male runners (5000–8000 m)					
20-39	70	47	8000	30	12
≥40	32	26	8000	31	10
Interclub—national winter road runs (4 races over 9 wk)					
Female runners (10 000–12 000 m)					
≥20	54	19	10 000	44	15
Male runners (10 000–15 000 m)	01	10	10,000		10
>20	234	94	10 000	38	14
Half marathons: region A (3 races over 13 wk)	204	04	10,000	00	17
Female runners					
	109	16	21 095	103	11
Male runners	105	10	21,000	100	
All	372	69	21 095	92	11
Half marathons: region B (2 races over 22 wk)	072	00	21,000	52	
Female runners					
20-39	652	17	21 095	113	13
20 33	100	8	21,000	116	15
+0-+3 ≥50	Q2	8	21,035	128	9
≥00 Male runners	52	0	21,000	120	5
20_20	736	26	21 005	05	18
20-39	387	10	21,035	08	18
40-49 ~50	250	10	21,000	107	16
≤ 30 Marathana: ragion P (2 ragon over 22 w/c)	235	19	21,095	107	10
Fomala ruppare					
	010	6	40 105	064	10
20-49	219	0	42,190	204	12
⊂00 Male ruppore	47	0	42,190	201	21
	514	40	42 105	216	01
20-43 ~ E0	014 150	42	42,190	210	21
200	199	24	42,195	250	22

CV, between-subject standard deviation expressed as a coefficient of variation.

^a Total for all distances in the given age group.

^b Estimates derived from statistical modeling of times for all runners who completed \geq 2 races of any distance in the age group.

the usual percent standard deviations that would be observed in any given race in the series between athletes who had competed in at least two races; these CV are the square root of the sum of the within- and between-athlete variances provided by the mixed modeling. The within-athlete CV shown in Table 2 are the square roots of the within-athlete variances. These CV are formally identical to the percent typical (standard) errors of measurement derived from reliability studies (1).

We performed separate analyses for runners of each sex in their age groups, where known. In some series, we merged age groups to get sufficient data for analysis. We examined plots of residual versus predicted values for each analysis, to check for extreme outliers and nonuniformity of the residuals. The following runners were deleted from the analyses, because the absolute values of one or more of their residuals was greater than 4 standard deviations: one male cross-country runner, four male runners from the summer road runs, and two male runners from half marathons of region B. These outliers were probably attributable to runners having the same name or to incorrect translation of a runner's number into a name by the race organizers. Names of one female and two male half marathon runners from region B appeared twice within the same race, presumably because of incorrect translation, so their data were also deleted.

Residuals generally showed a tendency to get larger for athletes with longer times in a given race. For this reason, we also analyzed runners of each sex in subgroups on the basis of their estimated mean time in all the races in the series. The estimate was the least-squares mean, which represents each runner's mean time if she or he had entered all races (2). The ranking of the least-squares means in each age group was used to split the runners into faster and slower halves; corresponding halves for different age groups ≥ 20 yr were then combined for the analysis of variability. (Runners in age groups <20 yr were not included, because the times for these runners were generally more variable than those of the adults, and the subgroup sample sizes were also much smaller than those of the adults.) For male runners, this analysis was repeated with the runners split into quartiles on the basis TABLE 2. Reliability of performance time expressed as the within-subject coefficients of variation (CV) for the faster and slower half of runners (within each age group \geq 20 yr) and for runners in a given age group.

	Female Runners		Male Runners	
	CV (%)	95%LL (%)	CV (%)	95%LL (%)
Interclub—national cross-country runs (4 races over 9 wk)				
Faster half	1.5	1.1-2.5	1.7	1.4-2.1
Slower half	2.5	1.8-4.1	1.9	1.6-2.4
Age <16	8.9	6.3-15	3.0	2.1-4.9
Age 16–19	1.5	1.0-3.1	3.4	2.5-5.6
Age ≥20	2.0	1.6-2.7	1.8	1.6-2.2
Interclub summer road runs (5 races over 4 wk)				
Faster half	1.6	1.2-2.5	1.6	1.4-1.9
Slower half	2.9	2.1-4.8	3.9	3.4-4.5
Age 20–39	2.8	1.9-4.9	3.4	3.0-3.9
Age ≥40	2.2	1.7-3.4	2.0	1.7-2.5
Interclub—national winter road runs (4 races over 9 wk)				
Faster half	1.1	0.7-1.9	1.8	1.5-2.1
Slower half	1.3	0.9-2.4	1.7	1.5-2.2
Age ≥ 20	1.1	0.9-1.6	1.7	1.6-2.0
Half marathons: region A (3 races over 13 wk)				
Faster half	2.0	1.3-4.0	2.9	2.3-3.7
Slower half	3.9	2.5-8.1	3.2	2.6-4.2
All ages	3.0	2.2-4.6	3.0	2.6-3.6
Half marathons: region B (2 races over 22 wk)				
Faster half	3.3	2.4-5.1	4.7	3.7-6.3
Slower half	3.7	2.8-5.7	6.3	5.0-8.4
Age 20–39	4.1	3.0-6.3	5.8	4.6-8.2
Age 40–49	3.2	2.1-6.5	5.7	4.3-8.6
Age ≥50	2.9	1.9-6.0	5.2	3.9-7.7
Marathons: region B (2 races over 22 wk)				
Faster half	3.8	2.4-9.7	3.8	3.0-5.0
Slower half	4.2	2.6-11	9.2	7.3-12
Age 20–49	4.0	2.5-10	7.1	5.8-9.0
Age ≥50	4.0	2.5-10	6.5	5.0-9.3

95%LL, 95% likely limits for the true value of the CV.

of speed; there were insufficient female runners to permit their analysis in quartiles.

Precision of the estimates of within-subject CV are shown as 95% likely limits (confidence limits), which represent the limits within which the true value is 95% likely to occur; the confidence limits were provided by Proc Mixed. We compared CV of two groups or subgroups by calculating the likely limits for the ratio of the CV, using the fact that the sampling distribution of the ratio of the sample to population variances in the two groups is an *F*-distribution. We regarded CV that differed by a factor of 1.15 or more as being substantially different, because the effect of such a difference on sample size in a controlled trial of competitive performance is a factor of 1.15^2 (1,3), or a change in sample size of 32%.

RESULTS

Readers familiar with competitive running times will see at once from Table 1 that the average runner competing in more than one race in each series was subelite, although the spread in ability included a small proportion of athletes of international caliber in each series. The spread in ability, represented in Table 1 as a coefficient of variation, was generally similar between series. The half marathon and marathon in region B had the widest spreads, probably because the mass-participation nature of these events resulted in a greater proportion of relatively slow runners than in the other events. Table 2 shows the within-athlete variability of performance times expressed as CV for the runners in each race series, running speed, age group, and sex.

Effect of race series. To compare the reliability of performance in the various race series, we combined the female and male CV of the faster runners, which showed a similar pattern of CV between series. Variability of times for cross-country, summer road, and winter road runs were virtually identical (e.g., ratio of winter road/cross-country CV, 1.0; likely range, 0.8-1.3). Times were more variable for the half marathon in region A than for the winter road races (ratio of CV, 1.7; likely range, 1.3-2.2) and for the half marathon in region B than in region A (ratio of CV, 1.5; likely range, 1.2–2.1). Times for the marathon tended to be a little less variable than for the half marathon in the same region (ratio of CV, 0.9; likely range, 0.7-1.2). The differences in CV between race series for the slower runners were generally similar to the differences for the faster runners, although the CV of the summer road runs and the marathon were relatively much larger than the CV of the other series for the slower runners.

Effect of running speed. We investigated this effect first by analyzing the CV for the faster and slower halves of adult runners (age, 20 yr). The effects were similar for both sexes, so we combined their CV. Times for the slower half of runners were more variable than those of the faster half in almost all race series (ratio of slower/faster CV, 1.0-2.3). The effects were clear-cut for summer road runs (ratio of CV, 2.3; likely range, 1.9-2.9) and for the marathon (ratio of CV, 2.2; likely range, 1.6-3.1).

The CV of the male runners in the fastest quartile (not shown in Table 2) were generally smaller than those of the next quartile. For the sake of brevity, we present only the CV for the fastest quartile: cross-country, 1.9% (likely range, 1.5-2.6%); summer road runs, 1.5% (likely range, 1.3-2.0%); winter road runs, 1.2% (likely range, 0.9-1.6%); half marathons in region A, 2.7% (likely range, 2.0-3.9%); half marathons in region B, 4.2% (likely range, 1.9-4.1%).

Effect of age group. Adequate data for age groups younger than 20 yr were available only for the cross-country series. On the basis of the likely limits for these CV, times for the male runners in both younger age groups are clearly more variable than those of the older male runners, whereas times for female runners are more variable only in the youngest age group. We excluded both these younger age groups from other analyses.

In the summer road runs, times for male runners of age 20-39 yr were more variable than those of age ≥ 40 yr (ratio of younger/older CV, 1.7; likely range, 1.3–2.1). Times for male runners in the young adult age groups (20-49 yr) in the half marathon and marathon also tended to be slightly more variable than those in the older age group (ratio of younger/older CV in half marathon, 1.1; likely range, 0.7–1.6; and ratio in marathon, 1.1; likely range, 0.7–1.6; and ratio in marathon, 1.1; likely range, 0.7–1.6; mates of the ratios of the CV were too imprecise to allow firm conclusions (because of smaller sample sizes for the female runners).

Effect of sex. CV for male and female runners differed by similar proportions in the faster and slower halves, so we combined the CV of these speed groups for the comparison. Times for adult male runners were more variable than those of female runners in almost all race series (ratio of male/ female CV, 0.9-1.7). The effect was clear-cut only for the winter road runs (ratio of CV, 1.5; likely range, 1.1-2.1).

DISCUSSION

The most useful results of the present study are the CV representing typical variation of performance for the faster female runners and fastest male runners in each type of race. As explained in the Introduction, these CV set benchmarks for the smallest worthwhile change in an athlete's performance and for the typical (standard) error of measurement of tests used to assess the smallest change.

If we take into account the precision of the estimates, it is reasonable to conclude that the CV for the fastest crosscountry and road adult males runners are all similar at ~1.5%, although the cross-country runners are probably a little more variable and the winter road runners a little less variable in their performance. The faster female runners in these series are probably a little less variable, with CV of ~1.3%.

Deciding on overall CV for the half marathon and marathon is more difficult, because of the discrepancy between the CV for the half marathon series in regions A and B. Obvious differences between these two series are the time between races and the number of races. Both these differences could account for the CV being smaller in the series in region A: with less time between races, individual differences in any real change in fitness between races are likely to be less, so the CV will be smaller; runners are also more likely to adopt a better pacing strategy for the second or third race on the basis of their memory of pace in the first or second race. Differences in environmental conditions in the two regions might also have produced differences in the CV. Region A (Otago, in the South Island of New Zealand) has a cooler climate than region B (Auckland, in the North Island), and it is known that athletes are more variable in their performance of the running stage of a triathlon in hot weather (Paton, C. D., and Hopkins, W. G., unpublished observations, 1999). On the other hand, we would expect performance in an environmentally challenging marathon to be even more variable than that of a half marathon, yet the times for the fastest runners in the marathon appear to be less variable than those in the half marathon in the same region on the same days. Our conclusion is that the best male runners in the half marathon and marathon under normal conditions probably have similar CV, ~2.5%. Female half marathon and marathon runners are less variable than male runners across speed groups and age groups, so it is reasonable to assume the female runners' observed CV of 2.0% in the half marathon is close to the true value, whereas their observed value in the marathon (3.8%) is inflated by some effect specific to the series in region B. Until more half marathon and marathon races are analyzed, we will assume that female runners have CV similar to those of male runners in these races: ~2.5%.

In the only other published study of the reliability of competitive performance, faster male swimmers of junior national level (mean age, 15 yr) had CV for race times of 1.0%; faster female swimmers may have been a little more variable (1.2%), but slower swimmers of both sexes were substantially more variable (1.5-1.6%). Comparison of the reliability of performance times between running and swimming is difficult, because the power developed by a swimmer is proportional to the cube of speed, whereas a runner's power is directly proportional to speed (4). To a first approximation, multiplying a coefficient of variation for swim time by a factor of 3 converts it to a coefficient of variation for power, but this factor does not take into account variability in swim times introduced by variability in the start and in the turns at the ends of the pool. The coefficient of variation for mean power of the swimmers is therefore probably ~2 times their coefficient of variation for time, that is, ~2-3%. The only athletes of comparable age in the present study are the cross-country runners in the <16-yr age group. The male runners in this group had similar reliability, but the female runners were much more variable, presumably because they had less experience of competitions.

The smallest enhancement of performance that has a substantial effect on a top athlete's chance of a medal is about one third of the typical variation of performance in competitions (3). It follows that top distance runners and their support professionals need to be concerned about

changes in performance of ~0.5% for cross-country and road races, and ~1.0% for half and full marathons. (These thresholds will need to be reduced, if performance in international endurance events is less variable than in the races we analyzed.) Tests suitable for delimiting such small changes in performance in a research setting need to have CV similar to, or preferably less than, the CV of the event they simulate-otherwise, the sample sizes are beyond the resources or pool of subjects available to the researchers (3). A practitioner using a test to monitor the performance of an individual athlete will also have little hope of noticing such small changes, unless the test has an even smaller coefficient of variation-otherwise, random variation in the athlete's change in performance between trials ($\sqrt{2}$ times the coefficient of variation of the test) will swamp the real change (one third the coefficient of variation of the event). Constant-speed tests to exhaustion have the highest reliability of all endurance tests (4) and probably have CV small enough to monitor for the smallest worthwhile changes in performance of individual athletes.

The fact that performance in half marathons and marathons is twice as variable as that in the shorter endurance events now needs explaining. It is possible that performance over longer durations is inherently more variable for physiological reasons, but if that were the case, we would expect to see some trend toward more variability as we move from the shorter to longer events in the other series (cross-country to summer road to winter road). If anything, we see less variability in this progression, so we suspect that duration of exercise per se does not affect variability of performance. Instead, we suggest that familiarity with competing over a given distance is the main factor: runners simply have less opportunity to compete or to practice competing over half marathon and marathon distances than over the shorter road and cross-country races. Familiarity with competing is also a likely explanation for the decrease in variability with increasing age; it would also help explain the greater vari-

REFERENCES

- HOPKINS, W. G. Measures of reliability in sports medicine and science. Sports Med. 30:1–15, 2000.
- HOPKINS, W. G., and J. R. GREEN. Combining event scores to estimate the ability of competitors. *Med. Sci. Sports Exerc.* 27:592– 598, 1995.
- HOPKINS, W. G., J. A. HAWLEY, and L. M. BURKE. Design and analysis of research on sport performance enhancement. *Med. Sci. Sports Exerc.* 31:472–485, 1999.

ability of the slower runners in each series if, as seems likely, the slower runners have less competitive experience.

Attitude toward competing may be another determinant of variability. Slower runners have little chance of winning, so they probably feel less motivated than the faster runners to attempt a current-best performance in every race. Attitude may also help explain the effects of age and sex on variability, if older runners and female runners are more likely to "run their own race." The lower variability of female runners is otherwise a puzzle. If we assume that most of the female runners in these races were not amenorrheic, any variation in performance across their menstrual cycles (5) was clearly insufficient to make them less consistent. Indeed, our results represent good evidence that the effect of the menstrual cycle on competitive endurance performance is negligible.

CONCLUSION

- The typical variation in competitive performance of faster adult male distance runners is ~2.5% in half and full marathons and ~1.5% in shorter endurance events.
- Performance tests suitable for tracking the smallest worthwhile changes in performance need typical errors of measurement similar to, or preferably less than, these typical variations.
- The smallest worthwhile changes are ~1% for half and full marathons, and ~0.5% for shorter endurance events.
- Female runners, older runners, and faster runners are less variable in their performance than male runners, younger runners, and slower runners.
- Most of the differences in variability probably arise from differences in competitive experience and attitude toward competing.

Address for correspondence: Will Hopkins, Department of Physiology, School of Medical Science, University of Otago, Box 913, Dunedin, New Zealand; E-mail: will.hopkins@otago.ac.nz.

- HOPKINS, W. G., E. J. SCHABORT, and J. A. HAWLEY. Reliability of power in physical performance tests. *Sports Med.* 31:211–234, 2001.
- LEBRUN, C. M., D. C. MCKENZIE, J. C. PRIOR, and J. E. TAUNTON. Effects of menstrual-cycle phase on athletic performance. *Med. Sci. Sports Exerc.* 27:437–444, 1995.
- STEWART, A. M., and W. G. HOPKINS. Consistency of swimming performance within and between competitions. *Med. Sci. Sports Exerc.* 32:997–1001, 2000.