

A Statistical Timetable for the Sub-2-Hour Marathon

SIMON D. ANGUS

Department of Economics, Monash University, Melbourne, AUSTRALIA

ABSTRACT

ANGUS, S. D. A Statistical Timetable for the Sub-2-Hour Marathon. *Med. Sci. Sports Exerc.*, Vol. 51, No. 7, pp. 1460–1466, 2019. **Introduction:** Breaking the sub-2-h marathon in an official event has attracted growing interest in recent times with commercial and international momentum building. Here it is shown that predicting *how likely* and *when* the sub-2-h barrier will be broken are statistically coupled considerations. **Methods:** Using a nonlinear limiting exponential model and calculating prediction intervals, a statistical timetable for the sub-2-h event is produced over a range of likelihoods. **Results:** At the benchmark odds level (1 in 10, or 10% likely), the expected sub-2-h arrival time is found to be May 2032. By estimating the model for male and female world record progressions, I find that limiting marathon times for males and females (at 1 in 10) are 1 h 58 min 5 s and 2 h 5 min 31 s, respectively. These times equate to a performance gap of 2.9% and 8.6%, respectively. The male estimate has remarkable similarity (~7 s) to Joyner's 1991 limiting human physiological estimate. Finally, I provide an estimate of the equivalent "sub-2-h" threshold for females and argue that a threshold of 130 min ("sub-130") could be an appropriate choice. **Conclusion:** The study is the first to address all three related aspects of world record marathon performance (sub-2 h, limits, gender equivalence) in a single, unified modeling framework and provides many avenues for further exploration and insight. **Key Words:** PREDICTION, PERFORMANCE LIMITS, PERFORMANCE GAP, NONLINEAR REGRESSION, GENDER EQUIVALENCE

The "sub-2-h" question—when will the first male break the 2-h marathon barrier in an International Association of Athletics Federations (IAAF) competition?—has attracted attention from the athletic and exercise physiology community for decades and can be traced back at least to A.V. Hill's *Lancet* study of 1925 (1). Prospects of a male athlete going sub-2 h in an IAAF even in the near future seem high given that the most recent world record reduced the mark by a full 1 min and 18 s, and the Nike "Breaking2" project produced a time just 25 s outside the sub-2-h barrier. But *how likely?* and *when?* Although physiological studies have identified the

essential attributes of the ideal sub-2-h runner (2–4), there has been scarce statistical treatment of the empirical progression of the male world record and much less the female world record. The aim of this study is to provide and analyze a statistical framework capable of jointly exploring the *likelihood* and *timing* of a runner going sub-2 h, and so complement the physiological perspective. The key assumption used throughout is that the performance improvement generating mechanisms that have been operating in the recent epoch of world record progression (specifically, since 1950) can reasonably be assumed to continue over the prediction period. By fitting a robust nonlinear model to official world record performances and then calculating a spectrum of prediction intervals, I intentionally couple *likelihood* and *timing* as joint attributes of any statistical forecast. In other words, the sub-2-h question is reconceptualized as one of odds: choose an odds of success level, for example, "1 in 10" (or "10% likely"), and an arrival time for the sub-2-h world record can be calculated. In addition, I demonstrate how the statistical framework can provide further insights into several related topics. After exploring the primary question, "When will the male sub-2-h threshold be broken?" I then pursue, "What are the limits of human performance for the male and female marathon?" "How far, in performance gap terms, is the current world record holder from the human performance limit?" and finally, "What is the equivalent of the 'sub-2-h' threshold for females, and when will this be achieved?"

The odds approach used in the present study differentiates it in a fundamental way from previous empirical attempts to provide an answer to the sub-2-h question. All time-series work

Address for correspondence: Simon D. Angus, Ph.D., Department of Economics, Monash University, Wellington Road, Clayton, 3800 Victoria, Australia; E-mail: Simon.angus@monash.edu.

Submitted for publication November 2018.

Accepted for publication February 2019.

Supplemental digital content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal's Web site (www.acsm-msse.org).

0195-9131/19/5107-1460/0

MEDICINE & SCIENCE IN SPORTS & EXERCISE®

Copyright © 2019 The Author(s). Published by Wolters Kluwer Health, Inc. on behalf of the American College of Sports Medicine. This is an open-access article distributed under the terms of the Creative Commons Attribution-Non Commercial-No Derivatives License 4.0 (CCBY-NC-ND), where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially without permission from the journal.

DOI: 10.1249/MSS.0000000000001928

over 25 yr that I am aware of (3,5–7) have modeled only the expected value (average trend) properties of the world record marathon progression. From an odds perspective, by strictly abiding by the expected value model line, one is effectively computing when the sub–2-h event will occur with 50/50 (“1 in 2,” or “50% likely”) odds. However, there is no reason why the sub–2-h runner should abide by these odds; they will, after all, be a product of the expected value performance improvement generating process *and* idiosyncratic factors unique to their own personal athleticism (and that of the event). Indeed, history shows many examples of “extraordinary” performances (i.e., those that depart from the expected value line). The question of prediction thus returns again to *how extraordinary* a particular performance is, relative to the expected value. In the related area of limits of human performance, this insight has been better absorbed, with at least one author (8) taking a serious approach to statistical variation in long-run predictions (I return to this study in Discussion).

METHODS

Data: Marathon world record progression since 1950. Official male and female marathon world record performance times since 1950, which are recognized by the IAAF, were obtained from the IAAF statistics handbook and augmented by the most recent IAAF recognized world record that occurred after publication (9). Note that the aim is to predict official road marathons under the IAAF rules, records that may exist beyond the IAAF listing (e.g., those recognized by the Association of Road Racing Statisticians only) are not considered. Respectively for males and females, 27 and 30 observations are obtained characterized by the date of the world record run, and the elapsed time to the nearest second, of the performance (see Appendices A and B, Supplemental Digital Content 1, for a complete listing of official male and female world record times considered in this study, <http://links.lww.com/MSS/B517>). To assist estimation, world record times are normalized by the 2-h mark (120 min) for both males and females before estimation (i.e., 120 min shall equal 1.00).

Statistical modeling: estimating the model. The modeling problem can be stated as follows: what is the *likely* value of the world record, WR, in a given year, t ? Stated in this way and given that there exists an observational set of world record data to work with, the problem can be treated as a prediction problem where $WR(t)$ is conceived of as a random variable having an underlying functional form with some noise associated with it. Similar to Refs. (10) and (8), I assume that the $WR(t)$ series is described by an exponentially declining functional form, defining an asymptotic limit (i.e., the world record achieved after an infinite period of performance gains or the human performance limit). To focus on the current performance improvement regime as closely as possible, only data after 1950 are included, effectively constraining the lens to the exponential limiting tail of the logistic functional form assumed by Nevill and Whyte (10) and Denny (8). Specifically, denote by WR_∞ the asymptotic world record limit and

by β and λ the initiation and rate of change parameters, respectively, and count time, in years, as an offset from the initial time in the data series, $\Delta t = t - t_0$, then the following time evolution of the world record can be written as,

$$WR(\Delta t) = \beta e^{-\lambda(\Delta t)} + WR_\infty. \quad [1]$$

Equation 1 has the properties that $WR(0) = \beta + WR_\infty$, and that $WR(\infty) = WR_\infty$. To estimate the three parameters that define equation 1, I consider each i -th observation ($WR_i, \Delta t_i$) is produced by the expected value component described by equation 1 and an independent error term, $\varepsilon_i \sim iid$,

$$WR_i = \beta e^{-\lambda(\Delta t_i)} + WR_\infty + \varepsilon_i. \quad [2]$$

Estimation of equation 2 was carried out using Matlab’s *fitnlm* routine (11), a robust nonlinear numerical optimization routine that iteratively down-weights observations that have the properties of outliers (12). The output from this step gives the expected value world record progression,

$$WR(\Delta t) = \hat{\beta} e^{-\hat{\lambda}(\Delta t)} + \hat{WR}_\infty \quad [3]$$

where $\hat{\beta}$, $\hat{\lambda}$, \hat{WR}_∞ and are the estimated parameters. The estimated models (for males and females) given by equation 3 form the work-horse of further analysis. All data and code used to produce the results of this study are available for download at: <https://github.com/specialistgeneralist/SUB2-AngusMSSE-2019>.

Statistical modeling: calculation of prediction intervals. whereas previous studies have obtained some form of equation 3 either through linear or nonlinear estimation means (either on world record only (3,5,7,10) or “best in season” data (6)), all have gone on to extrapolate their version of equation 3 to find a particular crossing point (such as, for males, the sub–2-h “barrier”) or asymptotic limit. Although informative to some extent, by only using equation 3 to generate out-of-sample predictions, one is ignoring two forms of uncertainty: 1) uncertainty associated with the fitted parameters on historical data and 2) uncertainty associated with individual realizations of the model that are presently unknown. Statistical methods typically handle the former with *confidence intervals*: “what is the likely range of the expected value (or average) realization of the fitted model in some future time?” In such a case, it is the variance associated with the estimated parameters that is used to provide a boundary for a given level of uncertainty. However, the pressing question for the sub–2-h barrier and any threshold concern in any athletic pursuit is not “when will the *expected value* (or *average*) performance improvement arise that will break the threshold (with some likelihood)?” but rather “when will a *particular* performance arise that will break the threshold (with some likelihood)?” The second question entails computing the second type of uncertainty above, or in the language of statistical methods, one needs to estimate the *prediction interval*. I note that one limits-of-human-performance marathon study, that of Denny (8), makes the identical point and approaches the long-run limit

with a second-step generalized extreme value (GEV) distribution model, discussed hereinafter.

Although computing prediction intervals for nonlinear models under robust (iteratively weighted) estimation is non-trivial, methods exist in statistical packages to compute these with ease. In the present study, Matlab's *predict* method (11) is applied to the nonlinear regression model object that arises from the first step (Data: Marathon world record progression since 1950), an implementation of Lane and DuMouchel's hybrid methodology (see Section 4, on prediction intervals, in Ref. (13) for details).

Uncertainty and odds. Prediction intervals (like confidence intervals) are calculated for a given level of uncertainty, α . For example, the region between the lower and upper prediction interval for an expected value prediction at $\alpha = 0.10$ implies that one would expect any observation to fall within the interval with $1 - \alpha = 0.90$, or 90% likelihood. Importantly, this calculation assumes that the uncertainty (10%) is equally distributed above and below the corresponding interval (two-tailed). So, in effect, the likelihood that an observation will fall *outside and below* the lower interval (one-tailed) is $\alpha/2$, or in the example, 5%. In the language of odds, it could equally be said that with a "1 in 20" chance, an event below the lower prediction bound might be observed. For the sub-2-h question, this is exactly the statistically posed question needed to ask of the estimated model. To support ease of interpretation, I shall thus report (two-tailed) α values and "1 in..." odds (one-tailed focus).

All results are provided at a range of α (and corresponding one-tailed odds) levels of uncertainty to provide the reader with a range of risk/uncertainty appetites to work with. However, "1 in 10" (one-tailed; $\alpha = 0.20$, two-tailed) is used as the benchmark level of uncertainty, as this level of odds squarely places a threshold outcome in the "reasonable chance" category, but not in the "extremal" case. In other words, I will concentrate on the phrasing, "with a 1 in 10 chance of success, when will [outcome x] occur?"

RESULTS

Model estimation. Model estimates are given in Table 1 for males and females, respectively, and graphically by fitted models presented in Figures 1A and 2. Model fits indicate root mean square error of just 0.0085 (males) and 0.032 (females) and R^2 for both estimated models are above 0.97. For males,

TABLE 1. Estimated model outcomes and male and female world record progression.

| | World Record | | | | |
|-----------------------------------------------------|----------------|--------|----------------------------------------------------|--------|--|
| | Male | | Female | | |
| | Estimate (SE) | t-Stat | Estimate (SE) | t-Stat | |
| β | 0.149 (0.0078) | 19.21 | 0.659 (0.021) | 31.61 | |
| λ | 0.034 (0.0053) | 6.46 | 0.097 (0.0070) | 13.90 | |
| WR_{∞} | 1.012 (0.0090) | 112.1 | 1.125 (0.0158) | 71.06 | |
| $N = 27, df = 24$ | | | $N = 30, df = 27$ | | |
| RMSE = 0.0085, $R^2 = 0.98$, $F\text{-stat} = 508$ | | | RMSE = 0.032, $R^2 = 0.98$, $F\text{-stat} = 557$ | | |
| $P < 0.001$ for all estimates | | | $P < 0.001$ for all estimates | | |

RMSE, root mean square error.

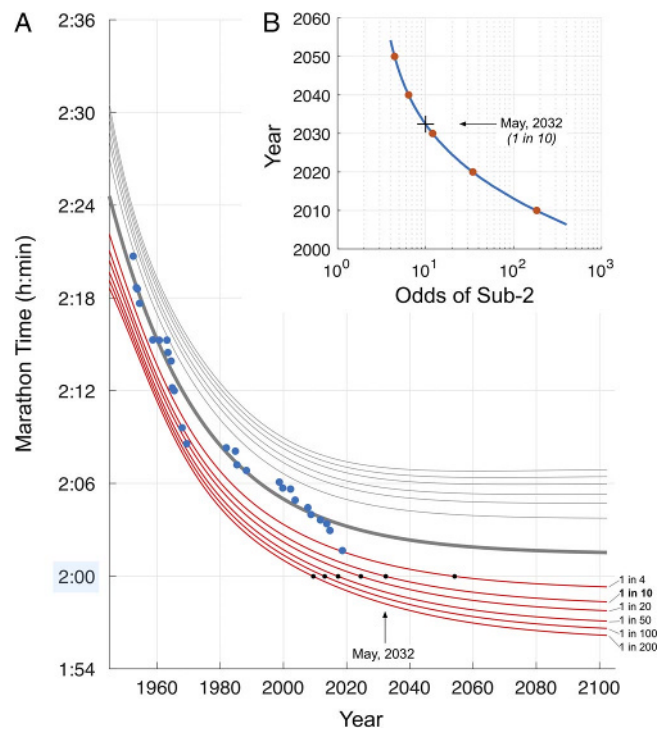


FIGURE 1—The progression of the male marathon world record, 1950 to present. **A**, Official IAAF marathon times, normalized such that 120 min = 1.0, vs year (blue markers), together with expected value model fit (thick gray line) and realizations of the upper (light gray) and lower (red) prediction intervals at odds levels as given. Black markers indicate crossing points with lower prediction bounds (see text for details), with the benchmark "1 in 10" odds crossing point indicated by an arrow. **B** (Inset), Continuous realizations of the sub-2-h crossing point for a variety of odds, crossings with decades 2010 ... 2050 indicated by red markers, with 1 in 10 odds crossing shown by black cross and arrow (see text for details).

this implies that the baseline model is, on average, accurate to the historical data to within less than 1%, or around 70s, over a 66-yr period, which saw a 19-min (1140 s) decline in the world record time. For females, accuracy is within around 3%, or around 200 s, over the same period where the world record declined by 1 h 22 min (>5000 s).

When will the male sub-2-h threshold be broken?

In Figure 1A (black markers), crossing points at 1 h 59 min 59 s for a variety of lower prediction interval boundaries are presented, having odds of success ranging from 1 in 4 to 1 in 200. At the benchmark odds of 1 in 10, the male sub-2-h observation should occur around May 2032. In other words, if an IAAF marathon were run in May 2032, then I would predict that there is a 1 in 10 chance (i.e., 10% likely) that a runner in the event will run "sub-2-h." Of interest, the analysis shows that the most recent world record time (Berlin, 2018) sits just inside the 1-in-4 odds prediction line: such a time at the event was expected to occur with around 25% likelihood. By inspecting the prediction lines, it can be seen that at the same event date, the odds of a sub-2-h run was a little higher than 1 in 50, or around 2% likely (or 98% unlikely). This analysis reminds us that a sub-2-h run could occur any time now, but the likelihood is still very rare. In Table 2, a statistical schedule

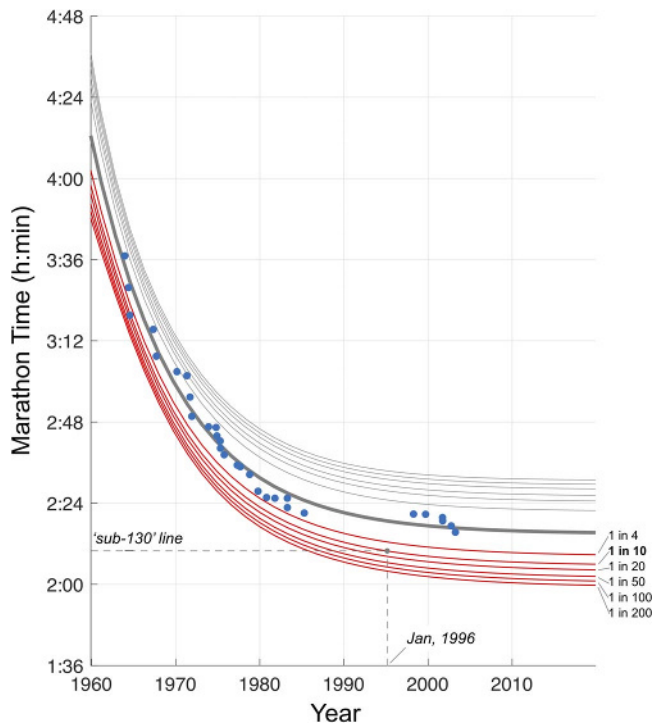


FIGURE 2—The progression of the female marathon world record, 1950 to present. See caption to Figure 1A for details. *Dashed lines* and annotations give the position of the suggested “sub-130” focus target for the female equivalent of the “sub-2-h” marathon. See text for details.

for the male sub-2-h world record marathon is presented to summarize these insights.

With the same machinery, I compute a sub-2-h schedule for a continuum of odds, as presented in Figure 1B, and show how, as one improves the odds of success, the arrival date moves further into the future at an increasing rate: the odds were just 1 in 183 in 2010; by 2020, these improve markedly to 1 in 34; then for 2030, 2040, and 2050, the odds improve more gradually from 1 in 12, to 1 in 6, and 1 in 4, respectively.

What are the limits of human performance for the male and female marathon? The estimated model (3) lends itself to simple calculation of the asymptotic limit of the performance improvement process. I note again that, whereas sub-2-h prediction requires that the performance improving epoch of recent history applies a few years, or decades into the future, applying the model to the asymptotic case requires that the current epoch is indeed the *final* epoch: no further material change will arise in the performance improving mechanisms at play today.

With this caveat, the asymptotic limit for male and female world record times is again a question of likelihood. In Figure 3A

TABLE 2. A statistical schedule for the “sub-2” male marathon world record.

| Odds | Chance, % | When? |
|----------|-----------|--------------|
| 1 in 4 | 25 | March 2054 |
| 1 in 10 | 10 | May 2032 |
| 1 in 20 | 5 | June 2024 |
| 1 in 50 | 2 | April 2017 |
| 1 in 100 | 1 | January 2013 |
| 1 in 200 | 0.5 | June 2009 |

The benchmark, 1 in 10, prediction is indicated by italics.

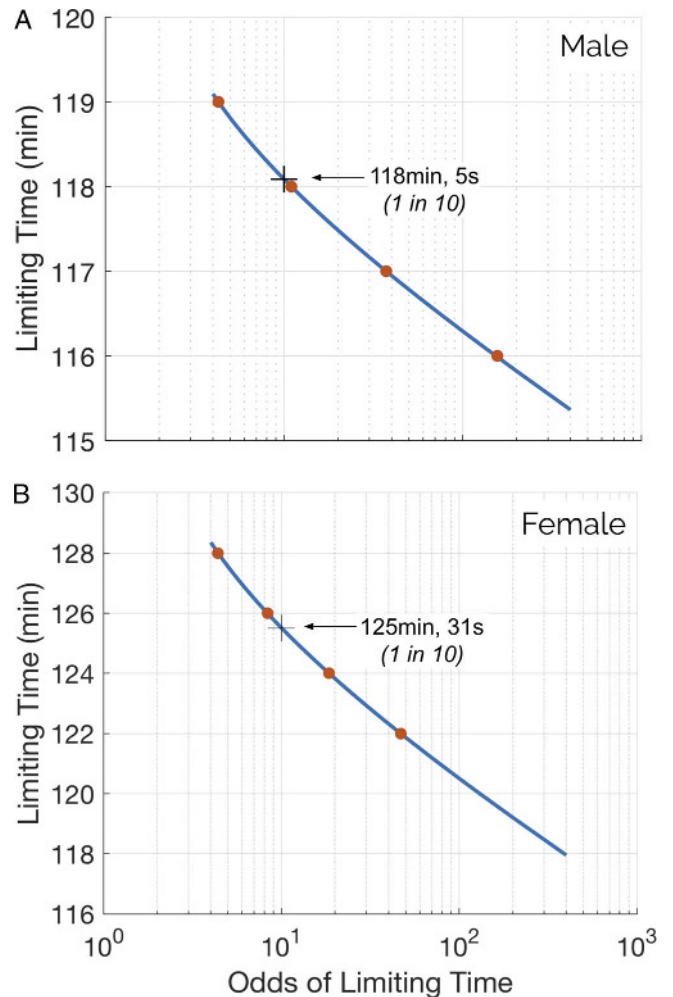


FIGURE 3—Asymptotic limiting times for the marathon world record. Male (A) and female (B), with significant limiting times indicated by red markers, and 1 in 10 odds limiting time given by black cross and annotation (see text for details).

and B, the limiting time, in minutes, for a range of odds is presented. At the benchmark likelihood of “1 in 10,” the male and female limiting times are given as 118 min 5 s (1 h 58 min 5 s) and 125 min 31 s (2 h 5 min 31 s), respectively.

One can see from Figure 3 that, in the limit, the sub-2-h threshold has strongly differing likelihoods. For males, to the question, “will a male ever break 2 h for the marathon?” I find that the likelihood approaches 1 in 2 (50% likely), whereas, for females, the likelihood of ever seeing a sub-2-h time approaches 1 in 100 (1% likely).

How far, in performance gap terms, is the current world record holder from the human performance limit? Setting aside the sub-2-h question for the moment, the current world record for males and females can be compared with their respective asymptotic limit at the benchmark 1 in 10 likelihood and the performance gap computed. For males, with a current world record of 121 min 39 s and an asymptotic limit of 118 min 5 s, the performance gap stands at 3 min 34 s, or around 2.9%. For females, with a current world record of 135 min 25 s and an asymptotic limit of 125 min 31 s, the gap stands at 11 min 35 s, or around 8.6%.

What is the equivalent of the “sub-2-h” threshold for females, and when will this be achieved?

Although the sub-2-h question for males has received academic and commercial interest, far less attention has been paid to the equivalent threshold for females. However, now that we have an estimate of the asymptotic limit for both genders, it is simple to obtain the performance equivalent of the male “sub-2-h” barrier for females. First, I calculate the performance gap from the 120-min barrier to the asymptotic limit for males just computed (118 min 5 s) and find a gap of 1.62%. Next, I use this gap to calculate the equivalent performance threshold for females by adding 1.62% to the female asymptotic limit (125 min 31 s) and obtain 127 min 33 s (2 h 7 min 33 s). With the current female world record set at 135 min 25 s, this time sits around 8 min off the threshold. With this result in hand, a near, entirely arbitrary, but “rounded” goal of 130 min—the “sub-130” goal—might be put forward as a reasonable choice for the female threshold. It is intriguing to note when the “1 in 10” prediction boundary crossed this “sub-130” line for females. Whereas for males, the sub-2-h crossing point (at 1 in 10) lies more than a decade away, for females, the “sub-130” crossing point occurred in January 1996 (refer to Fig. 2).

DISCUSSION

The present study is the first to my knowledge that provides an integrated analysis of three related questions in marathon performance analysis: the sub-2-h barrier, the limits of human performance, and a gender equivalence analysis.

First, on the question of sub-2-h prediction, the most influential academic work of Joyner et al. (3) came to a substantially earlier prediction for the sub-2-h threshold of either 2021–2022 (using data since 1960) or 2036 (using data since 1980). However, differences are simple to explain as Joyner et al. (3) use linear extrapolation without accounting for variability, compared with the nonlinear, stochastic modeling of the present work. Nevertheless, my preferred prediction (at 1 in 10 odds) of May 2032 sits in between the range used by Joyner et al. (3). Alternative approaches to the question have all come up with far more pessimistic views on the matter. Weiss et al. (6), using “season best performances” (rather than IAAF world records) and a mixture of nonlinear time-series analysis (but without variability) and experience curves, comment that the sub-2-h marathon is “unlikely to happen before the year 2100” (p. 400), whereas Tucker and Santos-Concejero (14), who compare the performance gender gap over a range of sports and, noting the above average gap in the marathon, conclude similarly, entitling their article, “The unlikeliness of the sub-2-h marathon.” Nevertheless, for reasons returned to below, the average gender gap is potentially a misleading basis on which to make such a pessimistic conclusion.

The limits of human performance are an area that has attracted a richer set of academic studies. Two related empirical studies that both use nonlinear fits to the world record progression are those of Nevill and Whyte (10) and Denny (8). Nevill and Whyte (10) use an expected value logistic modeling

approach for male world record progression over a longer series, and given that the logistic function and limiting exponential function (of the present study) share the same tail behavior, Nevill and Whyte (10) provide a reasonable test of the framework of my approach. If I apply the same expected value only approach of Nevill and Whyte (10) (i.e., the main curve, not the prediction interval boundary) and truncate my data to 2005 per the comparison, I obtain a limiting male world record time of 124 min 0 s, which indeed falls within 30 s to Nevill and Whyte’s (10) 123 min 38 s (Table 3 in the reference). Alternatively, Denny (8) comes at the question from a biological standpoint, comparing greyhounds, horses, and humans in the same treatment. Similarly to Nevill and Whyte (10), Denny (8) applies a logistic estimation approach, in addition to a population-based approach to the limiting estimation. However, in contrast to Nevill and Whyte (10), Denny (8) takes the variation around the curve into account, with a second-step GEV estimation. By this two-step approach, Denny (8) produces long-run estimates of 120 min 28 s for males and 124 min 58 s for females. On face value, these estimates seem quite high compared with the present study. However, results are not directly comparable as Denny (8) uses “annual best” times, with years that have runners found to be repeat “annual best” time holders being omitted (due to assumptions of the GEV step) rather than official IAAF world record times only. Although “annual best” reflects Denny’s (8) biological/population framework for the “fastest example of the species per year,” it necessarily oversamples from fast but non-world record performances, and so, estimates of both the curve and the variation from the curve will necessarily be blunted. Incidentally, for completeness, I apply GEV analysis to the residuals from the male and female fitting procedure in the present study and find that the first parameter (“a”) is not significantly of negative sign, and so, the maximum variation from the GEV is not defined. This matches my intuition around world record-only data having greater variance than “annual best” data.

A further, related consideration in the limits of human performance analysis is the evolution of the so-called gender gap, that is, the performance difference between female and male performance, expressed as a percentage of the faster performance. At present, the literature seems to have settled on an approximately 10% gender gap in endurance running performance, including the marathon (14–16). This number seems at odds with the gender gap one would calculate for the limiting times presented in the current study of 6.3%. Indeed, if one applies the fitted *expected* value models of this study to male and female performance over 1950 to 2100, one finds that the gender gap is steady and bounded by the range 9.6% to 11.6% from 1990 to 2100, fitting quite closely to the empirical findings of, for instance, Thibault et al. (15). However, because of the nature of the *prediction interval* curvature, the gender gap in the male and female *1 in 10* lines over the same period lies in a stable range of just 4.2% to 6.6%, encompassing the limiting gender gap of 6.3%. This exercise once again indicates the broader perspective one obtains by

taking into account the natural statistical variation of the problem at hand.

Of particular note is Joyner's (2) much earlier study that took a purely physiological approach to the question of the limiting male marathon time (i.e., $\dot{V}O_{2\max}$, running economy, etc.). Remarkably, Joyner's final estimate was 1 h 57 min 58 s, just 7 s outside my own preferred (1 in 10) limiting performance estimate for the male marathon of 1 h 58 min 5 s. Given the completely different basis for these estimates, the similarity in outcomes strongly triangulates the work of the present study.

The question of an equivalent "sub-2-h" threshold for females has been studied explicitly in Ref. (16). Here (and similar to Ref. (14)), the authors took a 10%–13% performance similarity between males and females as given, and, when applied to the 120 min "sub-2-h" marathon threshold, found that the current female world record was effectively already equivalent to the sub-2-h threshold, commenting that (current female world record holder) "Radcliffe's performances were exceptional." Without detracting from Radcliffe's remarkable sequences of world record runs, the results of my analysis shown in Figure 2 indicate that Radcliffe's most recent world record fell very near the expected value curve for the female world record progression, providing a different perspective to the analysis of Hunter et al. (16). Indeed, what is exceptional is why the recent period has not seen more energetic pursuit of the female world record progression by all stakeholders. On this point, I agree with Hunter et al. (16), who note that most likely asymmetric opportunities exist for females to compete at the highest level in athletic competition. Following this line of inquiry, I note that the demographics of elite marathoning, detailed in Ref. (17), show that for males, when comparing the average running speed of the top 10 athletes by continent of origin, male African runners demonstrate significantly faster running speeds than European runners to the tune of around 2.5%. However, this result is not borne out for females, with female African and European elite marathoners having identical mean running speeds ($5.00 \text{ m}\cdot\text{s}^{-1}$), suggesting that, if one assumes that European talent search mechanisms are somewhat exhaustive, there are potentially missing African elite females in the marathon. For example, if I apply the male African/European running speed ratio to Radcliffe's world record marathon speed (treating Radcliffe as a representative top-ranked European runner), I obtain a hypothesized running speed of $5.1932 \text{ m}\cdot\text{s}^{-1}$, or a marathon time of 2 h 12 min 4.5 s. If such a prototypical African athlete had turned up at the recent 2018 Berlin marathon where the male world record was set, a time of 2 h 12 min 4.5 s would have fallen, just like the male world record time, just inside the 1 in 4 (25% likely) prediction band

for the female world record marathon progression model (Fig. 2), indicating that this hypothetical computation is realistic. In any case, this exercise serves to bolster the view that African female marathoners are highly likely to be a genetic pool of future world record marathon progression and correspondingly point to institutional, economic, and social questions as to why this prototypical African female was not, indeed, running to victory on the streets of Berlin in 2018.

Finally, I conclude by noting various areas for development and extension of the present work. First, Eichner (18) has recently voiced concern about the prevalence of performance-enhancing drug abuse in second- and third-tier Kenyan runners. The modeling results of the present work could be used as a flagging device; "exceptional" performances can be benchmarked against the odds of occurrence triggering further inquiries. Second, the present analysis could equally be applied to a "pack" of top 10 or more performances in a given year (19) and used as the basis for point prediction. Third, the present analysis could easily be applied to any other world record progression across the athletic spectrum. Here, ensemble estimation methods (estimating performance development across sports in a single model) would be a fruitful line of inquiry. Fourth, I note that the present study leans heavily on the assumption that the current performance enhancement epoch/regime for males and females continues not only to the near or decadal future, but also, for limiting analysis, into the very far future. This assumption would be interesting to explore empirically by conducting formal structural break analysis of the performance improvement trajectory. In a related direction, one could focus on the functional stability of individual aspects of the performance enhancement epoch. For instance, one could study either the recent trajectory of economic incentives for elite male and female marathon runners, or the same for clothing and material technology enhancements. Here, negative evidence for the current/final epoch assumption would require at least one clear break in the functional form of one or other contributing performance dynamic, with *systematic* change (of whatever functional kind) over the period constituting positive evidence.

I thank colleagues and runners who have assisted with the formation of this study and have commented on prior approaches to the problem. In particular, I am indebted to three anonymous referees who each provided many thoughtful suggestions on an earlier manuscript that have served to improve the clarity and quality of the work. All residual errors or omissions are, of course, my own.

The author declares no conflict of interest regarding the study. The results of the present study do not constitute endorsement by the American College of Sports Medicine. The results of the study are presented clearly, honestly, and without fabrication, falsification, or inappropriate data manipulation.

REFERENCES

- Hill AV. The physiological basis of athletic records. *Lancet*. 1925; 206(5323):481–6.
- Joyner MJ. Modeling: optimal marathon performance on the basis of physiological factors. *J Appl Physiol*. 1991;70(2):683–7.
- Joyner MJ, Ruiz JR, Lucia A. The two-hour marathon: who and when? *J Appl Physiol*. 2011;110(1):275–7.
- Lundby C, Robach P. Performance enhancement: what are the physiological limits? *Phys Ther*. 2015;30(4):282–92.

5. Whipp BJ, Ward SA. Will women soon outrun men? *Nature*. 1992; 355(6355):25–5.
6. Weiss M, Newman A, Whitmore C, Weiss S. One hundred and fifty years of sprint and distance running—past trends and future prospects. *Eur J Sport Sci*. 2015;16(4):393–401.
7. Sousa CV, Sales MM, Nikolaidis PT, Rosemann T, Knechtle B. How much further for the sub-2-hour marathon? *Open Access J Sports Med*. 2018;9:139–45.
8. Denny MW. Limits to running speed in dogs, horses and humans. *J Exp Biol*. 2008;211(24):3836–49.
9. IAAF (2017), IAAF World Championships London 2017 Statistics Handbook. Available at: <http://www.iaaf.org/ebooks/2017/WCH/index.html>. Accessed November 2018.
10. Nevill AM, Whyte G. Are there limits to running world records? *Med Sci Sports Exerc*. 2005;37(10):1785–8.
11. *MATLAB Release 2017*. Natick (MA): The MathWorks, Inc.
12. Holland PW, Welsch RE. Robust regression using iteratively reweighted least-squares. *Communications Stat Theory Methods*. 1977;A6:813–27.
13. Lane TP, DuMouchel WH. Simultaneous confidence intervals in multiple regression. *Am Stat*. 1994;48(4):315–21.
14. Tucker R, Santos-Concejero J. The unlikeliness of an imminent sub-2-hour marathon: historical trends of the gender gap in running events. *Int J Sports Physiol Perform*. 2017;12(8):1017–22.
15. Thibault V, Guillaume M, Berthelot G, et al. Women and men in sport performance: the gender gap has not evolved since 1983. *J Sports Sci Med*. 2010;9(2):214–23.
16. Hunter SK, Joyner MJ, Jones AM. The two-hour marathon: what's the equivalent for women? *J Appl Physiol*. 2015;118(10):1321–3.
17. Marc A, Sedeaud A, Guillaume M, et al. Marathon progress: demography, morphology and environment. *J Sports Sci*. 2013; 32(6):524–32.
18. Eichner ER. Top marathon performance: interesting debate and troubling trends. *Curr Sports Med Rep*. 2015;14(1):2–3.
19. Kruse TN, Carter RE, Rosedahl JK, et al. Speed trends in male distance running. *PLoS One*. 2014;9(11):e112978.