



# The Texas Sharpshooter in the Three Grand Tours (1933-2013): No Evidence for Superior Time Trial Performances in the “Epo Era”

Hein F. M. Lodewijkx, Peter Verboon

Faculty of Psychology and Educational Sciences, Open University of The Netherlands, Heerlen, The Netherlands  
Email: [Hein.Lodewijkx@ou.nl](mailto:Hein.Lodewijkx@ou.nl), [Peter.Verboon@ou.nl](mailto:Peter.Verboon@ou.nl)

Received 2 September 2014; revised 17 October 2014; accepted 18 November 2014

Copyright © 2014 by authors and OALib.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## Abstract

Studies examining effects of doping in professional road racing building on archival records of the three major European stage races—the Tour de France, Giro d’Italia, and Vuelta a España—concluded that riders’ final performances in the “epo era” (>1990) strongly improved, yet declined since 2004. These observations can be criticized. First, we argue that time trial performances are more valid than final performances to indirectly evaluate doping effects. Second, we will pay attention to an informal logical flaw—the *Texas sharpshooter fallacy*—which may have biased findings and conclusions presented in the studies. To empirically substantiate our critique, we analyzed mean kilometers per hour (km/h) performances realized by winning riders in all time trials on flat and rolling terrain in the three tours (1933-2013,  $N = 325$ ). Regression analyses revealed no evidence for nonlinear in- or decreases in riders’ speed beyond the 1990s, but a straightforward linear progress over time of  $b = 0.16$  km/h per year ( $R^2 = 0.50$ ,  $p \leq 0.001$ ). Findings corroborate our comments on previous archival studies and qualify opinions about effects of the “epo epidemic” on cyclists’ achievements, since the time trial performances delivered in these years are no exemption to the observed linear progress in speed.

## Keywords

Epo Doping, Multi-Stage Cycling Races, Professional Road Racing, Time Trials

Subject Areas: Education, Sports Science

## 1. Introduction

An increasing number of doping scandals plagued the cycling world the last twenty years [1]. The most recent

**How to cite this paper:** Lodewijkx, H.F.M. and Verboon, P. (2014) The Texas Sharpshooter in the Three Grand Tours (1933-2013): No Evidence for Superior Time Trial Performances in the “Epo Era”. *Open Access Library Journal*, 1: e1045.  
<http://dx.doi.org/10.4236/oalib.1101045>

incident, evoking impassioned societal debates and worldwide media coverage, concerns American ex-professional cyclist Lance Armstrong. According to the United States Anti-Doping Agency [2] [3] (p. 5) he was involved in “a massive team doping scheme, more extensive than any previously revealed in professional sports history”. To indirectly examine effects of doping in professional cycling, El Helou *et al.* [4] and Perneger [5] assessed riders’ speed progress over time based on archival records of the three European Grand Tours (Tour de France, Giro d’Italia, and Vuelta a España). They reported significance in- and decreases in riders’ final kilometers per hour (km/h) performances beyond the 1990s. However, as will be delineated below, these studies may suffer from methodological errors. First, we will pay attention to the dependent variable used in these studies to indirectly appraise effects of doping agents on riders’ speed, arguing that time trial performances are more suitable than final performances to critically examine these effects. Second, we will focus on an informal logical flaw—the *Texas sharpshooter fallacy*—which may have led to a misrepresentation of developments in cyclists’ performance progress in the three tours in the epo era.

### 1.1. The Epo Era in Professional Cycling

The use of the banned, artificial glycoprotein hormone erythropoietin, or epo, became widespread in endurance sports such as cycling in the beginning of the 1990s [6]. Verbruggen [7], former president of the International Cycling Union (UCI), the sports’ governing body, agrees with this view. So does Vandeweghe [1], chairman of the Flanders Cycling Federation (WBV, Wielerbond Vlaanderen). Both maintain that epo doping enhances endurance performance by as much as 20 percent. Other studies, which examined annals of the cycling sport, appear to support these statements. Perneger [5] investigated mean km/h performances of riders who ranked fifth in the overall standings of the three major tours in the period 1990-2009. He reported that between 1990 and 2004, riders’ speed increased by 0.16 km/h per year and further observed a decrease in speed of 0.22 km/h per year since 2004. In 2001, the World Anti-Doping Agency (WADA [8]) implemented a direct test for epo, while a test for homologous blood transfusions (blood harvested from a compatible donor) was put into operation at the 2004 Summer Olympic Games in Athens [8]. Perneger [5] attributed the incline in riders’ speed to illicit doping practices and interpreted the decline since 2004 as evidence for the successes of WADA’s policy. El Helou *et al.* [4] also investigated archival data of the cycling sport using a time-series model. They analyzed mean km/h performances of riders who reached the first ten places in the final rankings of eleven European races from 1892 to 2008, which included all famous one-day classic races as well as the three major European stage races. Their analyses yielded four distinct periods: Before WW I (P1), between the two world wars (P2), after World War II (P3), while the fourth period (P4) starts after 1993. They reported that, relative to the third period (1946-1992), riders’ performances in the fourth period (1993-2008) showed an improvement in speed of 6.38%. Notice that their fourth period collides with the years of the epo era. They further put forward that the improvement they observed agreed with the 6.3% - 6.9% increase in aerobic exercise capacity found in laboratory studies assessing the performance-enhancing (or ergogenic) effects of epo administration on this capacity. However, in their review of epo studies, Lundby and Olsen [9] estimated improvements of 8% - 12%, while Vandeweghe [1] and Verbruggen [7] mentioned improvements up to 20%. These divergent estimates indicate disagreement between sources about these improvements.

### 1.2. Time Trial vs. Final Performances

Lodewijkx and Brouwer [10] [11] criticized the conclusions of the two historic studies alluded to above, but we will refrain from re-iterating these comments in the current paper. Importantly for the present thesis, in another study, Lodewijkx and Verboon [12] argued that these studies inappropriately used individual riders’ ultimate km/h performances as the dependent variable to evaluate effects of doping use on their achievements over the years. The study compared Armstrong’s time trial wins, realized in the Tour de France (1999-2005), to performances of other riders who, from 1934 to 2010, won similar races in the three major tours facing trial distances comparable to Armstrong’s (50 - 61 km). Findings indicated that, initially, he did realize significantly faster performances relative to the other riders ( $M = 4.70$  km/h,  $p \leq 0.05$ ,  $R^2 = 0.064$ ). However, after statistically controlling for the significant influence of competition year, he ultimately raced somewhat slower ( $M = -0.43$  km/h,  $p = 0.80$ ,  $R^2 = 0.001$ ). Only one of his wins exceeded the bounds of the 68%-CI (*i.e.*,  $\pm 1SD$  from the sample mean), while all his remaining six achievements fell within the bandwidth of this very stringent criterion. Thus, when considering the historic performance variation in these trials of limited distance, all individual accom-

plishments of the controversial American were far from superior. Moreover, in disagreement with results of the two historic studies presented above, the findings yielded a clear-cut linear increase in riders' mean km/h performances from 1934 to 2010 of  $b = 0.20$  km/h per year, which explained  $R^2 = 0.57$  ( $p \leq 0.001$ ) of the variation in riders' performances.

The foregoing findings entail that the choice of dependent variable may misrepresent conclusions concerning the influence of doping agents on riders' performance progress over the years. As already proposed by [5], selecting riders' final performances as the measure to evaluate this influence may be invalid, because the very same performances are also the result of joint and coordinated labors of the total group of riders participating in the three-week, multi-stage cycling races. Hence, these group labors can be considered confounding variables which may strongly hinder a sound assessment of individual riders' final km/h performances. Unpaced time trialing constitutes the "moment of truth" in cycling. In these races, riders in person compete for the fastest time and they cannot profit from the efforts of other riders in the race through drafting. In time trials of 5 - 70 km riders may deliver exceptional power outputs, ranging between 320 - 450 W [13]. These races therefore demand the utmost of individual riders' aerobic exercise capacity. Consequently, it is reasonable to argue that an assessment of these individual performances will raise the likelihood to indirectly identify the influence of banned, ergogenic doping aids on riders' sportive feats over the years. Consistent with Vandeweghe [1], we therefore maintain that individual riders' time trial performances are more valid than final performances to evaluate as to whether or not El Helou and colleagues and Perneger are correct in their observations and conclusions.

### 1.3. The Texas Sharpshooter Fallacy

We further argue that conclusions drawn in all three studies (admittedly, including our own) are perhaps unsound, because they all could have been subject to the same false logic: The Texas sharpshooter fallacy. This informal fallacy derives from a story about a Texan who practiced shooting his gun at the side of a barn located in his backyard. He randomly shot at the wall of the barn. After practicing several weeks, he drew a target onto the biggest cluster of hits and then claimed to be a sharpshooter. Applied to the observations discussed above, the logic illustrates that ("causal?") conclusions or inferences based on selective or limited data may not be valid. In combination with the incorrect choice of dependent variable, this fallacy might explain why findings of the three studies strongly contradict each other.

As noted, our time trial study [12] yielded a linear increase in riders' time trial performances from 1934 to 2010. El Helou *et al.* [4] reported an improvement in riders' final km/h performances from 1993 to 2008, while Perneger [5] described a significant downturn in the km/h achievements of the fifth-ranking riders in the years between 2004 and 2009. However, when considering results of the latter two studies, El Helou *et al.* state that the improvement they observed also held for the ten best classified riders in the three Grand Tours. Their sample thus includes the achievements of the fifth-ranking riders on which Perneger based his results. Yet, El Helou and co-workers did not report a significant downturn in speed after 2004. Finally, Perneger himself states that the trends he reported were far from consistent. It was the strongest for the Vuelta, modest for the Tour, and absent in the Giro. So, we conclude that findings between all three studies are strongly equivocal.

This inconsistency can be explicated by different samples, differences in dependent measures, selective statistical analyses, and the use of restricted ranges of observations. Indeed, our own time trial study may have built on selective data, since it evaluated km/h performances of riders who faced a limited range of trial distances (50 - 61 km). It is conceivable that statistical findings could be different, when reckoning the total variation in trial distances riders faced in the three major races over time. The 6.38% progress in the fourth relative to the third period obtained by [4] is also subject to criticism, owing to the specific way this increase was statistically tested. The researchers aggregated all riders' performances demonstrated in the third period and compared the resulting mean performances to the aggregated mean performances of riders in the fourth period. Because they also found a strong linear increase in performance from the third to the fourth period, the difference between the two periods might thus have been influenced by the relatively slower speeds demonstrated by riders in the years following WW II in the third period. Thus, the improvement they reported might be attributed to the selection of specific comparison groups. As a final point, Perneger's findings may be subject to the same fallacy, not only because his sample scrutinized a very restricted range of years in which riders competed, but also because he might have utilized an invalid dependent variable.

## 2. Research Questions

To re-assess developments in riders' progress in speed over the years, we decided to investigate all time trial wins demonstrated by riders on flat and rolling terrain in the three European tours from 1933 to 2013. In 1933, organizers of the Giro scheduled the first time trial ever in cycling. Our first research question examines whether El Helou *et al.*'s [4] and Perneger's [5] conclusions hold true. Based on El Helou and co-workers, we expect superior performances of riders in the 1990s and thereafter. Building on Perneger, we expect slower performances in the period since 2004 compared to the preceding epo era. Our second research question explores whether riders' expected faster or slower performances in the two distinct "critical" periods can be considered statistical outliers.

## 3. Method

### 3.1. Design and Sample

The annals of the French website "*Association Mémoire du Cyclisme*" [14] provided all information concerning our variables relating to the three tours. As noted, the first time trial ever was scheduled in the 1933 Giro and this initiative was immediately followed by the Tour organizers in 1934 and in 1941 by the Vuelta organizers. Team time trials, mountain time trials (racing uphill) and prologues were not included in the study, because they are considered different sub-disciplines in cycling [12]. The total number of time trials is  $N = 325$ . However, Van Dyck's victory in the first time trial of the 1947 Vuelta was not measured in time units, but was awarded points. Therefore, we discarded his performance, resulting in  $N = 324$  to conduct analyses.

### 3.2. Descriptive Statistics and Correlations

Descriptive statistics are presented in **Table 1**, Panels A to C. Regarding the total sample, Panel A shows that the trial distances ranged between 2.7 and 139 km ( $M = 43.83$ ;  $SD = 19.51$ ). Riders' mean km/h performances varied between 25.81 and 56.22 ( $M = 45.65$ ;  $SD = 4.65$ ). We expressed riders' mean time performances in hours, minutes and seconds and they varied between 00:03:02 and 03:49:36 with  $M = 00:59:00$  ( $SD = 00:30:25$ ). The correlation between the two measures,  $r(324) = -0.50$ ,  $p \leq 0.001$ , indicates a common variance of 25% between the two variables. Panel B presents the descriptive statistics for the three tours. Mean trial distances in the three races vary,  $F(2, 322) = 10.69$ ,  $p \leq 0.001$ . Collapsed across trials, distances were significantly larger in the Tour compared to the Italian and the Spanish race, which did not differ significantly from each other. Regarding Panel C, on the basis of Lodewijkx and Brouwer [11], we partitioned the years of competition into eight different periods of approximately ten years, accounting for El Helou *et al.*'s critical year 1993 (period 7). This year served as the standard to classify the ten-year periods after WW II. The epo era (period 7) comprises the years between 1993 and 2002, while Perneger's years (period 8) constitute the years between 2003 and 2013. Comparisons of trial distances between the pre-epo years (across periods 1 to 6) vs. the epo years vs. Perneger's years produced no significant differences,  $F(2, 322) = 0.68$ ,  $p = 0.50$ .

The correlation between competition year and trial distances,  $r(325) = -0.19$ ,  $p \leq 0.001$ , indicates a small decline in distances over the years. However, the scatter plot in Panel A of **Figure 1** shows that the relationship is not linear, but can best be described as an M-curve. Riders faced larger distances in the post-WW II years, shorter distances in the 1970s, growing distances in the 1980s, and shorter distances again after 2003. This relationship entails that we should control for the variation in trial distances when examining our research questions. Also, Panel B in **Figure 1** shows a nearly perfect relationship between trial distances and riders' mean time performances,  $r(324) = 0.97$ ,  $p \leq 0.001$ . As the distances grow larger, riders perform slower. Note that this substantive relationship holds true, regardless of the year in which riders competed and, hence, applies to riders' wins in the epo era ( $r = 0.97$ ) and Perneger's years ( $r = 0.97$ ) as well. We will attend to these robust correlations in more detail in the concluding sections of this paper.

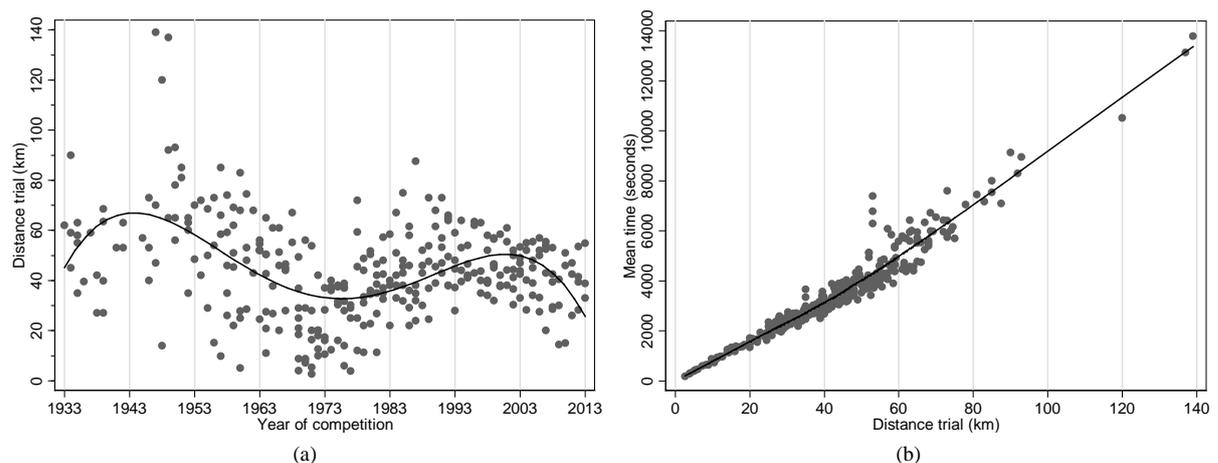
### 3.3. Analyses

We initially checked whether the relationship between competition year and riders' mean km/h performances indeed showed any signs of nonlinearity, using the Lowess scatter plot procedure [15], a nonparametric method for exploring and estimating nonlinear regression functions. We next statistically tested the same question by

**Table 1.** Descriptive statistics of time trials broken down by grand tours and time periods.

Samples	<i>N</i> (Trials)	Years	Distance (km)	$M_{\text{distance}}^1$	$M_{\text{km/h}}(SD)^2$	$M_{\text{time}}(SD)^2$
A: Total Sample	325	1933-2013	2.7 - 139	43.83	45.65 (4.56)	00:59:00 (00:30:25)
<b>B: Grand Tours</b>						
Tour	129	1934-2013	6 - 139	49.75 <sup>a</sup>	45.32 (4.07)	01:07:20 (00:35:52)
Giro	99	1933-2013	5 - 81	40.90 <sup>b</sup>	46.51 (3.54)	00:53:19 (00:19:31)
Vuelta	97	1941-2013	2.7 - 93	38.95 <sup>b</sup>	45.20 (5.85)	00:53:40 (00:29:24)
<b>C: Time Periods</b>						
1. –	19	1933-1942	27 - 90	52.75	38.14 (4.68)	01:25:33 (00:30:25)
2. –	22	1945-1952	14 - 139	72.18	39.01 (3.61)	01:52:23 (00:50:40)
3. –	35	1953-1962	5 - 85	50.14	43.50 (3.43)	01:09:50 (00:31:19)
4. –	52	1963-1972	2.7 - 67	33.05	44.30 (3.72)	00:46:07 (00:27:58)
5. –	54	1973-1982	3.8 - 72	32.51	45.91 (2.42)	00:42:43 (00:18:55)
6. –	50	1983-1992	22 - 87.5	47.08	47.05 (2.55)	01:00:23 (00:19:59)
Pre-Epo Years (Total)	232	1933-1992	2.7 - 139	43.85 <sup>a</sup>	44.16 (4.29)	01:01:16 (00:34:50)
7. Epo Years	47	1993-2002	28 - 64	46.12 <sup>a</sup>	49.34 (2.30)	00:56:10 (00:11:41)
8. Pernerger's Years	46	2003-2013	14.4 - 57	41.40 <sup>a</sup>	49.34 (3.17)	00:50:29 (00:14:01)

Notes: <sup>1</sup>Within panels, means without a common superscript differ significantly,  $p \leq 0.05$ . For time periods, the pre-epo years (collapsed across time periods 1 - 6), were contrasted with the epo and Pernerger's years. <sup>2</sup>Km/h = kilometers per hour performance; time = time performance in hours, minutes and seconds. Both performance measures are based on  $N = 324$ .

**Figure 1.** Lowess scatter plots of the relationships between competition year and trial distances (Panel a) and between trial distances and mean time performances (Panel b).

linear regression analyses (OLS). Reckoning the variation in trial distances over time, we decided to control for the influence of this variable in the analyses, as well as for differences between races and the interaction effects between Competition Year  $\times$  Distance, Competition Year  $\times$  Races, and Distance  $\times$  Races. To these ends, we mean centered the interval variables ( $M_{\text{year}} = 1978$ ;  $M_{\text{distance}} = 43.83$  km) and effect-coded the races variable: First contrast Tour = 1, Giro = -1, Vuelta = 0; second contrast Tour = 0, Giro = 1, Vuelta = -1.

Regarding our second research question, as follow-up of the regression analyses we examined whether riders' estimated residual performances can be considered outliers, concentrating on performances realized by riders in the 1990s and thereafter. To settle on outliers, we used the rather stringent criterion of  $z \geq \pm 1.96$  ( $\pm 2SD$ ) instead of the generally accepted  $\pm 3SD$ -criterion (or  $z \geq \pm 3.30$ ; [16]). The same analysis will also produce information concerning influential cases—such as extremely fast or slow performances, and large or small trial distances—that may have put the stability and validity of the regression findings at risk. Given the current data, the critical leverage value to decide on influential cases is  $h \geq 0.09$  [16]. We analyzed data by IBM-SPSS<sup>®</sup> (v. 20).

## 4. Results

### 4.1. Progress over Time

**Figure 2** presents the Lowess scatter plot of the relationship between competition year and riders' mean km/h performances. The correlation between the two variables,  $r(324) = 0.71$  ( $R^2 = 0.50$ ,  $p \leq 0.001$ ), designates that riders race faster over time. The plot reveals a straightforward linear progress in performance with no signs of any nonlinear deviations in the epo era and in Perneger's years. The same plot additionally shows that riders' achievements in the two critical periods show a strong variation, characterized by fast as well as slow performances.

Results of the linear regression analyses are summarized in **Table 2**. The nine variables included in the equation explained  $R_{\text{adj}}^2 = 0.604$  of the variation in riders' performances. The linear influence of competition year proved to be the strongest of all variables, indicating that riders race  $b = 0.16$  km/h faster per year. Furthermore, larger trial distances are associated with a decrease of  $b = -0.03$  km/h in speed to a kilometer increase in distance. Yet, the influence of this variable is less substantive relative to the influence of competition year, presumably owing to the fact that the mean km/h performance measure (as the dependent variable) already includes the distance variable. **Table 2** further reveals no significant difference in average speed between the Tour and the Giro, whereas the mean speed in the Giro is  $b = 1.47$  km/h faster than in the Vuelta. **Table 2** additionally shows that the five interaction effects together explained  $R_{\text{adj}}^2 = 0.060$  of the variation in km/h performances. The interaction effect between Competition Year  $\times$  Distance (A  $\times$  B) did not produce a significant effect. The analyses also yielded no differences between the French and the Italian race in interaction with competition year and trial distances (A  $\times$  C and B  $\times$  C), while the interaction effects of the latter variables with the Giro and the Vuelta were significant (A  $\times$  D and B  $\times$  D), explaining  $\Delta R^2 = 0.01 - 0.036$  of the differences in riders' speed. Auxiliary simple slopes analyses of the A  $\times$  D-interaction effect revealed that in the Vuelta the speed progress over time is faster ( $b = 0.21$  km/h per year) relative to the Giro ( $b = 0.11$  km/h per year). For the Tour the progress is intermediate ( $b = 0.16$  km/h per year). **Figure 3** presents the A  $\times$  D-interaction effect. It shows that the Italian race is characterized by a strikingly regular yearly progress in performance compared to the other two races, while performances are comparatively slower in the early years of the Spanish race compared to the Giro. Closer scrutiny of the B  $\times$  D-interaction effect indicated that trial distances in the Spanish race are associated with somewhat slower performances to a kilometer increase in distance ( $b = -0.10$  km/h) than in the Giro ( $b = -0.03$  km/h). Again, the relationship is intermediate in the French race ( $b = -0.06$  km/h). Last, **Table 2** indicates no sign of multicollinearity among the predictor variables, since none of them approached the critical benchmark for tolerance (0.1 - 0.2) which is considered a cause for concern [16]. This adds validity to our findings.

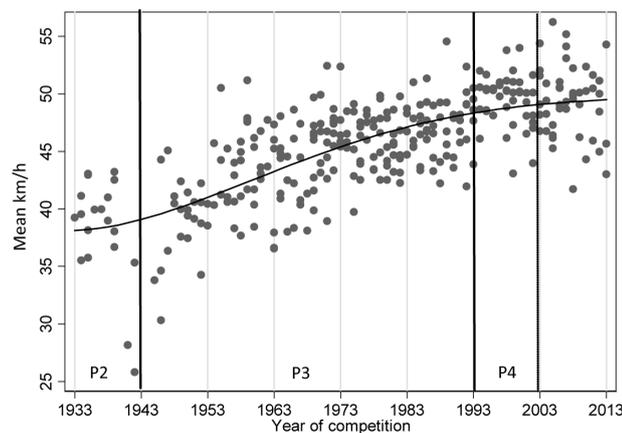
Most relevant for our first research question is the strong linear relationship we found between competition year and riders' winning performances. **Figure 4** presents the partial regression plot of this relationship with 95%-CI. In disagreement with conclusions presented by El Helou and colleagues, the 1990s are not characterized by nonlinear, superior performances compared to performances demonstrated by riders in preceding years. Furthermore, inconsistent with Perneger's conclusion, **Figure 4** also illustrates that riders' performances do not level off going from 2003 to 2013. Rather, they show a clear-cut linear progress in speed in this era compared to the foregoing epo years. This linear progress is in line with findings we discussed earlier relating to the Lowess scatter plot (see **Figure 2**).

To validate these conclusions, we conducted two supplementary analyses. Inspired by El Helou *et al.*, in a first analysis we compared riders' performances in the pre-epo years (across periods 1 to 6) to performances of riders in the epo years (period four). **Table 1** shows that performances in the epo years ( $M_{\text{km/h}} = 49.34$ ) at first were faster than performances in the pre-epo years ( $M_{\text{km/h}} = 44.16$ ;  $R^2 = 0.189$ ,  $p \leq 0.001$ ). After controlling for the influence of year of competition ( $b = 4.60$  km/h) and trial distances ( $b = -0.096$  km/h) on the rider-km/h performance relationship, the main effect was not significant anymore, revealing that riders in the epo years ( $M_{\text{km/h}} = 45.59$ ) raced 0.67 km/h faster than riders in the pre-epo years ( $M_{\text{km/h}} = 44.92$ ;  $R_{\text{adj}}^2 = 0.004$ ,  $p = 0.30$ ). The relative progress in speed of riders in the epo years equals 1.5%. In the second analysis we evaluated performances in the epo years against performances in Perneger's years. **Table 1** shows no initial performance differences between the two groups ( $M_{\text{km/h}} = 49.34$ ,  $p = 1.00$ ). However, after adjusting this main effect for the influence of competition year ( $b = -0.83$  km/h) and trial distances ( $b = -0.06$  km/h), findings revealed that riders in Perneger's years ( $M_{\text{km/h}} = 49.79$ ) outperformed riders in the epo years ( $M_{\text{km/h}} = 48.90$ ) with 0.89 km/h ( $R_{\text{adj}}^2 = 0.007$ ,  $p = 0.42$ ). This progress in speed amounts to 1.8%. Notice that the slight performance differences we

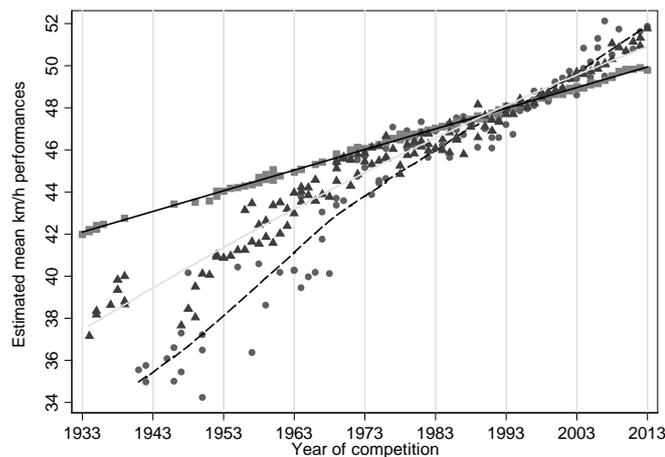
**Table 2.** Mean km/h performances regressed on competition year, trial distances, races, and interaction effects ( $N = 324$ ).

Effects	Variables	$b$	$SE$	$\beta$	$\Delta R^2$	$R^2_{adj}$	Tolerance
	Intercept	45.39***	0.19	-	-	-	-
Main effects	Comp. year (A)	0.16***	0.01	0.71	0.502	0.549***	0.93
	Distance (B)	-0.03**	0.01	-0.13	0.024		0.51
	Tour vs. Giro (C)	0.28	0.24	0.05	0.002		0.66
	Giro vs. Vuelta (D)	1.47***	0.27	0.25	0.025		0.58
	$A \times B$	0.01	0.01	0.08	0.005		0.52
Interaction effects	$A \times C$	-0.01	0.01	-0.01	0.010	0.060***	0.63
	$A \times D$	-0.06***	0.01	-0.20	0.036		0.71
	$B \times C$	0.01	0.01	0.05	0.001		0.61
	$B \times D$	0.04**	0.01	0.12	0.010		0.65
Total	-	-	-	-	-	0.604***	-

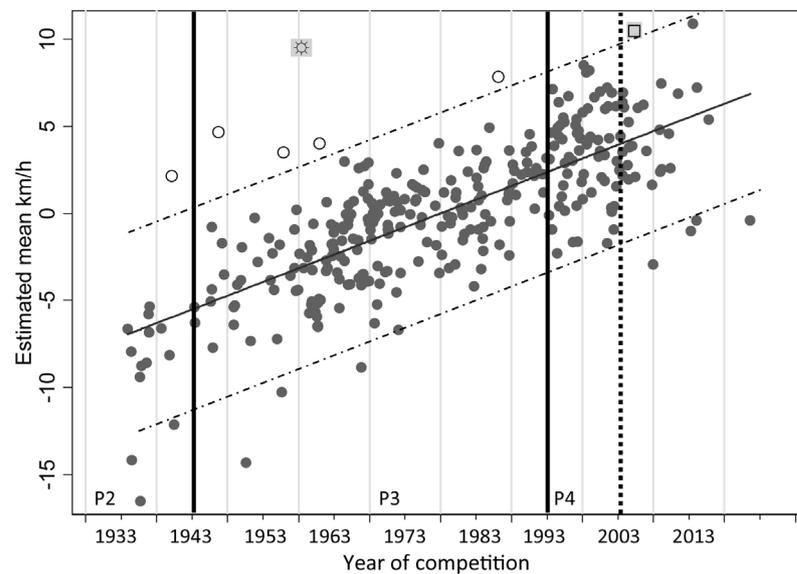
Notes: Degrees of freedom vary between  $df = 1, 314$  and  $df = 9, 314$ .  $b$  and  $SE$  are in kilometers per hour per year, or per kilometer increase in trial distance. \*\*  $p \leq 0.01$ ; \*\*\*  $p \leq 0.001$



**Figure 2.** Lowess scatter plot of the relationship between year of competition and riders' mean km/h performances in time trials ( $N = 324$ ). P2 to P4 present time periods distinguished by El Helou *et al.* The dotted line indicates the start of Perneger's years (2003).



**Figure 3.** Regression plot of the relationship between year of competition and riders' estimated mean km/h performances in time trials in the Tour (▲), Giro (■), and Vuelta (●). The figure illustrates the interaction effect between Giro/Vuelta and competition year ( $A \times D$ ).



**Figure 4.** Partial regression plot of the relationship between year of competition and riders' estimated mean km/h performances in time trials ( $N = 324$ ) with 95%-CI. Variables are mean centered. P2 to P4 present time periods distinguished by El Helou *et al.* The dotted line indicates the start of Perneger's years (2003). Fast performances beyond the 95%-CI are indicated by ○. ☆ indicates the performance of Roger Rivière in the 1959 Vuelta which genuinely can be regarded an outlier (see Table 3). □ presents Plaza's performance in the 2005 Vuelta.

obtained between the various comparison groups did not yield significant results. Still, they are indicative of the linear progress in speed we discussed previously. The results of these supplementary analyses thus agree with conclusions we drew earlier concerning research question one.

## 4.2. Outliers and Influential Cases

In answer to our second research question, Table 3 provides a summary of outliers and influential cases. It presents a total of 26 riders of whom eighteen demonstrated rather fast or slow residual performances beyond the 95%-CI and seven had large leverage values ( $h \geq 0.09$ ). The last column in the table shows that the latter values all concern riders who either faced extremely long or short trials. Returning to riders' slow and fast performances, Table 3 indicates that seven riders demonstrated comparatively fast residual performances. Yet, inconsistent with expectations based on El Helou *et al.*, only one (1.1%) of the 93 riders who performed in their critical years realized such an achievement: Spanish rider Ruben Plaza who achieved a formidable 56.22 km/h in his 38.9 km-long trial in the 2005 Vuelta (see Figure 4).

Eleven of the eighteen cyclists raced rather slowly. Only four (8.7%) of the 46 riders in Perneger's years delivered such performances: Bruseghin (Giro, 2008); Dowsett (Giro, 2013); Kessiakoff (Vuelta, 2012); and Cancellara (Vuelta, 2013). In conclusion, since a total of five (5.4%) of the 93 riders in the two critical periods we distinguished delivered comparatively slow or fast performances that surpassed the 95%-bandwidth, it is safe to state that our findings disprove our second research question.

## 5. Discussion

In disagreement with expectations based on [4] and [5], our findings yielded no support for our first research question. When taking time trial performances rather than final performances as the criterion, there is no evidence for nonlinear in- or declines in riders' speed beyond the 1990s. Across the three main European stage races, riders' mean km/h performances are characterized by a steady linear progress of 160 m per year and their achievements in the 1990s and 2000s are no exemption to this development. Inconsistent with observations of [4], the progress in speed we observed from the pre-epo to the epo era amounted to 1.5%, and not 6.38%. In a similar vein, not in line with [5], riders did not show a downturn in speed since 2003. Rather, they realized a progress in speed of 1.8% compared to riders in the epo era. We additionally found that the linear progress in

**Table 3.** Casewise diagnostics of riders with slow or fast residual km/h performances and influential cases.

Rider	Race (Year)	Distance	Observed km/h	Predicted km/h	Residual km/h	$z$	Leverage ( $h$ )
Impanis	Tour (1947)	139	36.32	37.58	-0.26	-0.09	0.18
Lambrecht	Tour (1948)	120	41.08	37.69	3.38	1.18	0.11
Coppi	Tour (1949)	137	37.56	37.15	0.41	0.14	0.15
<b>LeMond</b>	<b>Tour (1989)</b>	<b>24.5</b>	<b>54.55</b>	<b>47.53</b>	<b>7.02</b>	<b>2.45</b>	<b>0.03</b>
Coppi	Giro (1951)	81	39.12	42.87	-3.76	-1.31	0.12
Coppi	Giro (1952)	35	34.25	44.27	-10.02	-3.50	0.03
Poblet	Giro (1960)	5	46.15	45.59	0.57	0.20	0.11
Ventarelli	Giro (1960)	25	38.43	45.19	-6.76	-2.36	0.04
Merckx	Giro (1969)	49.3	39.84	45.62	-5.78	-2.02	0.01
Boifava	Giro (1971)	28	38.92	46.04	-7.12	-2.48	0.02
Bruseghin	Giro (2008)	39.4	41.71	49.38	-7.68	-2.68	0.03
Dowsett	Giro (2013)	54.8	43.01	50.30	-7.30	-2.55	0.07
Rodriguez	Vuelta (1941)	53	28.14	35.47	-7.33	-2.56	0.05
Rodriguez	Vuelta (1942)	53	25.80	35.69	-9.88	-3.45	0.05
Langarica	Vuelta (1946)	53	30.31	36.56	-6.25	-2.18	0.04
<b>Van Dyck</b>	<b>Vuelta (1947)</b>	<b>47</b>	<b>45.05</b>	<b>37.38</b>	<b>7.67</b>	<b>2.68</b>	<b>0.04</b>
Ruiz	Vuelta (1948)	14	40.45	40.85	-0.40	-0.14	0.09
Capo	Vuelta (1950)	93	37.43	33.56	3.87	1.35	0.14
<b>Magni</b>	<b>Vuelta (1955)</b>	<b>29</b>	<b>50.48</b>	<b>40.74</b>	<b>9.74</b>	<b>3.40</b>	<b>0.04</b>
Lorono	Vuelta (1957)	85	38.28	36.08	2.20	0.77	0.09
<b>Rivière</b>	<b>Vuelta (1959)</b>	<b>62</b>	<b>51.15</b>	<b>38.61</b>	<b>12.54</b>	<b>4.37</b>	<b>0.04</b>
<b>Gonz.-Linares</b>	<b>Vuelta (1971)</b>	<b>2.7</b>	<b>52.42</b>	<b>45.92</b>	<b>6.50</b>	<b>2.27</b>	<b>0.06</b>
<b>Merckx</b>	<b>Vuelta (1973)</b>	<b>10.5</b>	<b>52.35</b>	<b>45.67</b>	<b>6.68</b>	<b>2.33</b>	<b>0.04</b>
<b>Plaza</b>	<b>Vuelta (2005)</b>	<b>38.9</b>	<b>56.22</b>	<b>50.11</b>	<b>6.11</b>	<b>2.13</b>	<b>0.02</b>
Kessiakoff	Vuelta (2012)	39.4	44.94	51.53	-6.59	-2.30	0.03
Cancellara	Vuelta (2013)	38.8	45.65	51.76	-6.11	-2.13	0.04

Notes: In bold type face are riders with relatively fast residual performances with  $z \geq +1.96$ . Influential cases: leverage  $h \geq 0.09$ .

performance varied between races. It is significantly faster in the Spanish than in the Italian race, while no significant differences emerged between the French and the Italian race.

We also did not find evidence for our second research question. In disagreement with El Helou and colleagues, only one of the 93 riders in their critical era delivered a rather fast performance: Ruben Plaza in the 2005 Vuelta. However, it can be doubted whether his achievement can be regarded extremely fast. Not only because we took the rather rigorous standard of  $+2SD$  to evaluate conspicuous performances, but also because the circumstances during the trial must have been very favorable. To illustrate, closer inspection of the classification of the trial revealed that Alberto Ongarato, who reached the 100<sup>th</sup> position, already raced at a speed of nearly 50 km/h [14]. Likewise, only four of the 46 riders in Perneger's era demonstrated relatively slow wins beyond the 95%-confidence interval. It remains to be seen whether these performances can be attributed to the successes of WADA's anti-doping measures, as suggested by Perneger. Alternatively, they can also be reasonably explicated by unfavorable weather, road and terrain conditions, which are of the utmost importance in unpaced time trialing [17]. In our case, all four riders faced extremely hilly courses, a circumstance which reduced their overall speed considerably.

## Implications and Comments

How can we settle these findings with generally shared opinions concerning, for instance, the proposed powerful ergogenic effects of epo doping on cyclists' performances? Our reply would be that this relationship is overvalued [18] [19] and might even lack scientific evidence [20]. To critically examine the strength of this relationship, Lodewijckx *et al.* [19] conducted a meta-analysis on the findings of seventeen laboratory studies that hitherto investigated this relationship. In these studies, participants were treated with epo or not. The studies subsequently assessed the degree of performance improvement attributable to epo administration, measured by maximal oxygen uptake ( $\text{VO}_{2\text{max}}$ ) and maximal aerobic power output ( $W_{\text{map}}$ ). The analysis yielded modest effect sizes,  $d = 0.41 - 0.49$ ,  $r = 0.19 - 0.44$  and  $r^2 = 0.04 - 0.19$ , indicating that a considerable 81% - 96% of the differences in performance improvement observed in the studies cannot be attributed to epo administration. The largest epo-induced improvement in  $\text{VO}_{2\text{max}}$  yielded by the analysis equaled an increase in speed of about 1 km/h. Perneger [5] arrives at the same conclusion concerning this improvement in speed. However, we emphasize that this slight increase is solely restricted to laboratory situations. It is a well-known fact that such improvements cannot be directly extrapolated to multi-stage cycling races that last three weeks [19]-[21]. All these observations imply that judgments pertaining to the strong ergogenic effects of epo doping on aerobic exercise capacity are far from conclusive. In turn, this means that the relationship between epo doping and performances of endurance athletes such as cyclists in real contests is overestimated too. In our view, this constitutes the most parsimonious explanation for the null results we obtained regarding riders' expected superior time trial performances in the epo era.

Moreover, the differences in progress in speed between races we found also put great pressure on conclusions drawn in previous historic studies. Notice that Perneger [5] reported similar trends in his study. This variation is inconsistent with generalizing statements pertaining to the hypothesized strong effects of epo doping used by riders in the three tours in the last twenty years. For instance, how can such statements be reconciled with the prominent regular performance progress observed in the Italian race? Additionally, arguing from the same perspective, it is also inconsistent to explain the very same differences by referring to riders' use of different doping agents in the three races as an explanatory variable. Another explanation for these differences, suggested by [5], concerns the assumed, rather lenient attitudes to adhere to and to put into practice WADA's anti-doping regulations in the different countries. However, the real reasons to clarify these differences might be much more mundane. For instance, the faster yearly progress in the Vuelta vs. the Giro we found, can mainly be attributed to the rather slow performances realized by riders in the early years of the Spanish race (see [Figure 3](#)). During these years, the Vuelta struggled with civil war and the Franco regime, a deep economic recession, a competition which for many years restricted itself to local Spanish riders and did not include more famous and very competitive foreign Italian and French riders, and, last, sometimes extremely bad weather, race and road conditions [22].

The same problems arise when trying to elucidate the nearly perfect correlation we obtained between trial distances and riders' mean time performances,  $r = 0.97$ . As a single variable, trial distances are responsible for a substantive 94% of the performance differences between riders over the years, leaving 6% of these differences unexplained. We already described that this strong relationship is not influenced by the years in which riders competed and thus applies to riders' wins in the epo era and the years thereafter. In descriptive terms, some studies suggest that the ergogenic effects of epo doping on aerobic exercise capacity are "dramatic" [23], whereas effects of blood transfusions are labeled "gigantic" in a recent study by Lundby, Robach and Saltin [24]. However, if these observations were valid, one would expect the trial distance-time performance relationship to be strongly moderated by the years in which riders won their trial, most notably the 1990s. Yet, this is not what our findings tell us.

Last, it is a rather biased and circular way of reasoning to mainly concentrate on riders' performances in the epo era and to try to find explanations for these achievements, thereby completely ignoring comparable advances in foregoing years. In our view, the simplest explanations for the performance improvement over time we found, concern winning riders' outstanding athletic capacities combined with other facilitative, performance-enhancing factors such as more favorable road, terrain and race conditions, less demanding racing programs, growing insights from exercise physiology, more sophisticated and effective training regimes, improved technology of bikes and racing gear, increased specialization of riders, growing commercial interests with concomitant increases in financial incentives and pressures to perform, and improvements in nutrition and hydration, leading to an enhanced maintenance of riders' energy balance during stage events.

## 6. Conclusion

To conclude, findings of the current study corroborate our comments [10]-[12] on previous archival studies, which examined effects of doping use and anti-doping efforts on riders' final km/h performances in the three major European races in the years of the "epo epidemic" and thereafter. As suggested in the Introduction of this paper, the choice of an incorrect dependent variable combined with the Texas sharpshooter fallacy may indeed have led to biased conclusions relating to these developments. When reckoning individual riders' time trial performances, our findings compellingly revealed there is no solid evidence whatsoever that justifies making any such statements.

## References

- [1] Vandeweghe, H. (2013) Who Has Faith in Cyclists Anymore? [In Dutch: Wie gelooft de coureurs nog?]. Borgerhoff & Lamberigts, Ghent.
- [2] USADA (2012) Statement from USADA CEO Travis T. Tygart Regarding the U.S. Postal Service Pro Cycling Team Doping Conspiracy. <http://cyclinginvestigation.usada.org>
- [3] USADA (2012) Reasoned Decision of the United States Anti-Doping Agency on Disqualification and Ineligibility. <http://cyclinginvestigation.usada.org>
- [4] El Helou, N., Berthelot, G., Thibaut, V., Tafflet, M., Massif, H., Campion, F., *et al.* (2010) Tour de France, Giro, Vuelta, and Classic European Races Show a Unique Progression of Road Cycling Speed in the Last 20 Years. *Journal of Sports Sciences*, **28**, 789-796. <http://dx.doi.org/10.1080/02640411003739654>
- [5] Perneger, T.V. (2010) Speed Trends of Major Cycling Races: Does Slower Mean Cleaner? *International Journal of Sports Medicine*, **31**, 261-264. <http://dx.doi.org/10.1055/s-0030-1247593>
- [6] Catlin, D.H., Hatton, C.K. and Lasne, F. (2006) Abuse of Recombinant Erythropoietins by Athletes. In: Molineux, G., Foote, M.A. and Elliott, S.G., Eds., *Erythropoietins and Erythropoiesis*, Birkhäuser Verlag, Basel, 205-228. [http://dx.doi.org/10.1007/3-7643-7543-4\\_13](http://dx.doi.org/10.1007/3-7643-7543-4_13)
- [7] Verbruggen, H. (2001) The EPO Epidemic in Sport. *Bloodline Reviews*, **1**, 3-4.
- [8] WADA (2013) World Anti-Doping Agency Prohibited List. <http://www.wada-ama.org>.
- [9] Lundby, C. and Olsen, N.V. (2011) Effects of Recombinant Human Erythropoietin in Normal Humans. *Journal of Physiology*, **589**, 1265-1271. <http://dx.doi.org/10.1113/jphysiol.2010.195917>
- [10] Lodewijckx, H.F.M. and Brouwer, B. (2011) Some Empirical Notes on the "Epo Epidemic" in Professional Cycling. *Research Quarterly for Exercise and Sport*, **82**, 593-608. <http://dx.doi.org/10.1080/02701367.2011.10599811>
- [11] Lodewijckx, H.F.M. and Brouwer, B. (2012) Tour, Giro Vuelta: Rapid Progress in Cycling Performance Starts in the 1980s. *International Journal of Sports Science*, **2**, 24-31.
- [12] Lodewijckx, H.F.M. and Verboon, P. (2013) Lance Armstrong's Era of Performance—Part I: Are His Time Trial Performances Much Different from Other Winners? *Journal of Athletic Enhancement*, **2**, 1. <http://dx.doi.org/10.4172/2324-9080.1000106>
- [13] Jeukendrup, A.E., Craig, N.P. and Hawley, J.A. (2000) The Bioenergetics of World Class Cycling. *Journal of Science and Medicine in Sport*, **3**, 414-433. [http://dx.doi.org/10.1016/S1440-2440\(00\)80008-0](http://dx.doi.org/10.1016/S1440-2440(00)80008-0)
- [14] Magnier, D., Picq, P., Debreilly, M., Zingoni, P., Haffreingue, H. and Bey, J.-L. (2014) Cycling Remembered. [In French: Mémoire du cyclisme]. <http://www.memoire-du-cyclisme.eu>
- [15] Cleveland, W.S. and Devlin, S.J. (1988) Locally-Weighted Regression: An Approach to Regression Analysis by Local Fitting. *Journal of the American Statistical Association*, **83**, 596-610. <http://dx.doi.org/10.2307/2289282>
- [16] Tabachnick, B.G. and Fidell, L.S. (2001) Using Multivariate Statistics. 4th Edition, Allyn and Bacon/Pearson Education Company, Boston.
- [17] Lucia, A., Earnest, C. and Arribas, C. (2003) The Tour de France: A Physiological Review. *Scandinavian Journal of Medicine and Science in Sports*, **13**, 275-283. <http://dx.doi.org/10.1034/j.1600-0838.2003.00345.x>
- [18] Kuipers, H. (2006) Putative Effects of Doping in Cycling. [In Dutch: Vermeende effecten van doping in de wielersport]. *Nederlands Tijdschrift voor Geneeskunde*, **150**, 2643-2645.
- [19] Lodewijckx, H.F.M., Brouwer, B., Kuipers, H. and Van Hezewijk, R. (2013) Overestimated Effect of Epo Administration on Aerobic Exercise Capacity: A Meta-Analysis. *American Journal of Sports Science and Medicine*, **1**, 17-27.
- [20] Heuberger, J.A.A.C., Cohen-Tervaert, J.M., Schepers, F.M.L., Vliegthart, A.D.B., Rotmans, J.I., Daniels, J.M.A., Burggraaf, J. and Cohen, A.F. (2012) Erythropoietin Doping in Cycling: Lack of Evidence for Efficacy and a Negative Risk-Benefit. *British Journal of Clinical Pharmacology*, **75**, 1406-1421. <http://dx.doi.org/10.1111/bcp.12034>

- [21] Hopkins, W.G., Hawley, J.A. and Burke, L.M. (1999) Design and Analysis of Research on Sport Performance Enhancement. *Medicine and Science in Sports and Exercise*, **31**, 472-485. <http://dx.doi.org/10.1097/00005768-199903000-00018>
- [22] Fallon, L. and Bell, A. (2005) *Viva la Vuelta! The Story of Spain's Great Bike Race*. Mousehold Press, Norwich.
- [23] Ninot, G., Connes, P. and Caillaud, C. (2006) Effects of Recombinant Human Erythropoietin Injections on Physical Self in Endurance Athletes. *Journal of Sports Sciences*, **24**, 383-391. <http://dx.doi.org/10.1080/02640410500131340>
- [24] Lundby, C., Robach, P. and Saltin, B. (2012) The Evolving Science of Detection of "Blood Doping". *British Journal of Pharmacology*, **165**, 1306-1315. <http://dx.doi.org/10.1111/j.1476-5381.2011.01822.x>