



**EFFECTS OF FALSE-START DISQUALIFICATION RULES ON  
RESPONSE-TIMES OF ELITE-STANDARD SPRINTERS**

Journal:	<i>Journal of Sports Sciences</i>
Manuscript ID	RJSP-2016-0387.R2
Manuscript Type:	Original Manuscript
Keywords:	Athletics, sprint start, exponentially modified Gaussian distribution, auditory performance

SCHOLARONE™  
Manuscripts

EFFECTS OF FALSE-START DISQUALIFICATION RULES ON RESPONSE-TIMES OF ELITE-STANDARD SPRINTERS

Abstract

The 100 ms ruling for false start disqualification at athletic competitions governed by the International Association of Athletics Federations has been in force since the early 1990's. Throughout this period, there have been marked changes to the rules that govern the disqualification of athletes from sprint events incorporating starts from blocks. This study analysed all available World and European Championship response-time data from 1999 to 2014 to examine effects of rule changes on competition response-times at major championships. The exponentially modified Gaussian distribution was used to model response-times and make comparisons relative to athletes' sex, ruling periods and competition rounds. Revised response-time thresholds of 115 ms and 119 ms were identified for men and women respectively, indicating that the current 100 ms rule could result in some false starts not being detected in competitive athletics. The study proposes that when using existing International Association of Athletics Federations approved systems, the false start detection threshold should be increased and that men and women athletes should have different thresholds because of substantial evidence of a sex-based difference in response-times in elite-standard athletes.

**Keywords:** sprint start; athletics; exponentially modified Gaussian distribution; auditory performance.

Introduction

Response-times (RT) of elite-standard 100 m sprinters has been identified as a benchmark against which to gauge the absolute limits of human auditory performance (Lipps, Galecki et al., 2011). A RT in the context of this paper is similar to that used in the founding paper by Mero and Komi (1990) defined as the complete response-time of an athlete which includes both the pre-motor time and motor time of the sprint start. Accurate identification of the RT period after

the start signal in sprint starts especially at major championships is vital for the fair and impartial refereeing of competitive athletics.

The International Association of Athletics Federations (IAAF) rule 161.2 stipulates that a false start occurs when a sprinter registers a RT less than 100 ms (International Association of Athletics Federations, 2015). Before January 2004, a RT of less than 100 ms (i.e. a false start) resulted in a warning on a competitor and the competitor was disqualified for two false starts in a race. This rule could result in multiple false starts in a single heat, causing disruption to timetables and having adverse effects on spectators. From January 2004, a false start by any competitor placed all athletes in a race on a first warning. Subsequently, any sprinter registering a false start would be automatically disqualified regardless of whether he or she was the original offender. In 2010, to eradicate gamesmanship (Athletics Australia, 2009), the IAAF made further important changes to the false start disqualification. Under the revised “No False Start” rule, no warnings are issued and a competitor who false starts is automatically disqualified and removed from the race.

The 100 ms period used to detect a false start is measured as the time taken by an athlete to produce a predefined force or acceleration threshold on the starting-blocks. Lipps, Galecki et al. (2011) noted that the IAAF justification for this 100 ms threshold was based on a considerably dated study that had only Finnish male sprinters, none of whom was of elite-standard (Mero and Komi, 1990). Pain and Hibbs (2007) suggested that genuine RTs of 85 ms are possible and Brown, Kenwell et al. (2008) presented similar results reporting that 21% of recorded

RTs were shorter than the 100 ms temporal threshold. While such RTs are possible, the recording systems used in both studies measured the time to the first change in force on the blocks produced by the athlete rather than the time taken to produce a predefined force. A revision of the false start detection threshold was proposed by Komi, Ishikawa et al. (2009), based on a study of sprint starts on four men and three women, Finnish national-level sprinters. While this study questioned the validity of the current 100 ms rule, it was weakened by having few participants with limited major championship status.

Previous studies have provided empirical evidence of a sex-based difference in the strength of elite-standard athletes (Lipps, Galecki et al., 2011) and of the general population (Der and Deary, 2006), but IAAF rule 161.2 continues to use a single RT threshold of 100 ms. Given the physiological differences in strength in men and women, it is possible that a sex-based difference in RTs require that rules for men’s and women’s events should be determined independently.

Clearly, a re-evaluation of the 100 ms threshold is required using analysis of data recorded using IAAF approved starting technologies and relevant data from international competitions. Collet (1999) studied the RTs of elite competitors between 1987 and 1997, while Tønnessen, Haugen et al. (2013) carried out a similar analysis for events between 2003 and 2009. While both of these studies added to the sprint start RT literature, the data used span only a single ruling period. Haugen, Shalfawi et al. (2013) studied elite-standard athletes over the period 1997 – 2011, but focused solely on the 100 m sprint across the three ruling periods of interest. They suggested a 20% increase in RTs over a 15 year

period, but did not provide a full analysis of the differences across ruling periods because of a lack of data for the third ruling period. All of the previous studies have been based on restricted sample sizes and many have not included data from athletes of international major championship status. Appropriate statistical analysis and modelling of a large sample of RTs for men and women at major international championships could be used to derive an appropriate temporal threshold for sprint start RTs in competitive athletics. Consequently, the primary aim of this study was to determine whether IAAF rule changes have influenced the competition RTs of elite-standard sprinters in major international championships by statistically modelling the historical RT data to see if there are differences across ruling periods, sex and round of competition. A secondary aim was to establish whether data collected across multiple major championship competitions could be combined and used to revise the IAAF 100 ms RT threshold for male and female sprinters accordingly by using a robust statistical modelling approach.

## Methods

### *Data Acquisition*

The collection and analysis of the data in this study was approved by the University of Limerick Education and Health Sciences Research Ethics committee (2013\_06\_07\_EHS). RT data were collated from the IAAF results (<http://www.iaaf.org/results>) and the European Athletics websites (<http://www.european-athletics.org/results/>) and included complete Championship event data for the 60 m and 60 m hurdles indoor championships, the 100 m, 110 m hurdles and the 200 m outdoor championships over the period

1999 to 2014 inclusive. Events over distances greater than 200 m were excluded since the RT has less influence on the performance and therefore would not provide valid information on the **shortest** possible RT of an elite athlete. The data contained race RTs for 1303 and 1007 men and women, respectively, with a total of 4560 and 3999 RT records for performances of men and women respectively. While available data contained additional information on the finish time, wind speed and name of the sprinter, the only variables of interest in this study were the RT of the sprinter, the year of the event and the round of competition. Based on an extensive literature search, this study constitutes the largest exploration of RTs of **elite-standard** sprinters to date and far exceeds the work of Mero and Komi (1990) which is proposed as the basis for the current 100 ms threshold (Lipps, Galecki et al., 2011).

***Exploratory Analysis***

Descriptive statistics for the centre and variability of RTs relative to **athletes' sex** and ruling periods were calculated. Box-and-whiskers plots provide visualisation of the RTs of men and women across ruling periods. The exploratory analysis also examined the number of false starts recorded in each ruling period. **Times from** OMEGA and SEIKO starting technologies at European and World Championships **were compared to see if they influenced** RTs of athletes.

***Statistical Modelling***

The exponentially modified Gaussian distribution (EMGD) has been widely used to model RTs (Dawson, 1988). The EMGD is a convolution of an exponential

distribution, controlling the tail of the distribution, and a Gaussian distribution to represent the group of quickest responses. Its probability density function (PDF) is:

$$f(RT|\mu, \sigma, \tau) = \frac{1}{\tau} \exp \left\{ \frac{\mu}{\tau} + \frac{\sigma^2}{2\tau^2} - \frac{RT}{\tau} \right\} \Phi \left( \frac{RT - \mu - \frac{\sigma^2}{\tau}}{\sigma} \right),$$

where,  $\mu$  and  $\sigma$  represent the mean and standard deviation of the Gaussian component and  $\tau$  is the rate parameter of the exponential distribution, while  $\Phi(t)$  is the cumulative standard normal distribution function evaluated at  $t$ . The parameters are constrained such that  $\min(RT) \leq \mu \leq \max(RT)$ ,  $0 \leq \sigma \leq \text{range}(RT)$  and  $0 \leq \tau \leq \text{range}(RT)$ . If RTs follow an EMGD with the above parameters its measure of location and scale can be represented as  $\mathbb{E}(X) = \mu + \tau$  and  $\text{Var}(X) = \sigma^2 + \tau^2$ , respectively.

### **Estimation**

The parameters (location, scale and rate) for the EMGD were estimated using the maximum likelihood approach provided in the retimes package (Massidda, 2013) based in the R statistical package (R Core Team, 2015). Maximum likelihood estimates of the parameters for the EMGD are described in detail in Silver, Ritchie et al. (2009).

### **Goodness-of-fit Statistics**

Pearson's Chi-Squared goodness-of-fit evaluated if the observed RT data came from an underlying EMGD. The null hypothesis assumed that the data fitted the distribution well, while the alternative assumed that the distribution was not a true reflection of the observed data. For details of this goodness-of-fit test, see Greenwood and Nikulin (1996). Statistical significance was  $\alpha = 0.05$ , with  $r - p - 1$  degrees of freedom, where  $r$  is the number of quantiles at which the fit

is compared, 10 is used in all calculations in this study, and  $p$  is the number of estimated parameters that is 3 for all cases when using the EMGD. The critical value is taken from the Chi-squared tables with 6 degrees of freedom at a significance level of 0.05 which corresponds to a critical value of 12.592.

**Comparison across ruling periods and rounds**

To compare differences across ruling periods and rounds, a simple probabilistic approach was taken. The probability of a RT above a given threshold,  $T$ , being observed can be expressed as  $\Pr(RT \geq T) = 1 - \Pr(RT < T)$ . In turn, the  $\Pr(RT < T) = F(T)$  where  $F(T)$  is the cumulative distribution function (CDF) of the EMGD evaluated at  $T$ . The CDF is formulated by integrating the PDF shown above, resulting in the EMGDs CDF being:

$$F(T) = -\exp\left(\frac{-T}{\tau} + \frac{\mu}{\tau} + \frac{\sigma^2}{2\tau^2}\right)\Phi\left(\frac{T - \mu - \frac{\sigma^2}{\tau}}{\sigma}\right) + \Phi\left(\frac{T - \mu}{\sigma}\right),$$

where,  $T$  is the threshold of interest,  $\mu$ ,  $\sigma$  and  $\tau$  are the parameters of the EMGD defined earlier and  $\Phi$  is again the CDF of the standard normal distribution. In addition the probability of superiority is used to compare populations of interest. The statistic returns a probability that a randomly chosen member of the “upper” population lies above the median of the “lower” population; a value of greater than 50% indicates a difference in the two populations.

**Revision of the False Start Detection Threshold**

For men and women, an EMGD was estimated to fit the sex-specific RTs across all ruling periods. The CDF of the EMGD presented earlier identified the threshold ( $T$ ) above which 99% of the observed RTs lay. The identified



threshold is proposed as an effective method to determine the true minimum RT by an elite athlete using the current technologies approved by the IAAF.

## Results

Men's events reported a median value of 156 ms and 159 ms for World and European Championships respectively, while women's events reported medians of 161 ms and 164 ms respectively. Table 1 provides descriptive statistics for the RTs of elite-standard sprinters relative to rule change periods set down by the IAAF for men and women. The median and a 95% coverage interval are reported because there was a positive skew in the RTs. The rows of the table represent the combination of ruling period and sex. The columns provide number of observations, median, a 95% coverage interval and number of false starts recorded in the period.

\*\*\*\*Table 1 near here\*\*\*\*

Box-and-whisker plots for the RTs by ruling period are presented in Figure 1. The left panel is for men, the right for women. The x-axis for both panels corresponds to the three ruling periods with the commonly scaled y-axis representing RTs in milliseconds. The horizontal dashed line at 100 ms denotes the current disqualification limit –RTs recorded below 100 ms correspond to false starts that are not a true reflection of the shortest valid RTs in elite-standard athletics. The circled numbers below the threshold line provide the number of disqualifications because of false starts according to athletes' sex and ruling period.

\*\*\*\*Figure 1 near here\*\*\*\*

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

The results of fitting the EMGD for men and women in each ruling period are provided in Table 2 and Figure 2. These also compare the distributional properties of RTs from first round events and finals, since races of greater importance **could** encourage athletes to take risks and anticipate the start signal. The parameters column presented in Table 2 provides the best estimates of the EMGD parameters given the data from the ruling period and **athletes' sex**, the properties column shows the location and scale of the distribution, while the final two columns provide a measure of goodness-of-fit of the estimated EMGD to the observed data with the corresponding p-value for the result of the test.

\*\*\*\*Table 2 near here\*\*\*

The estimated distributions for ruling periods, first round and finals by sex are shown in Figure 2. Panels A and B show the fit for each of the three ruling periods, the dotted line is the fit for the 1999-2003 ruling period, the dashed line is the estimated distribution for the 2004-2009 period, while the immediate disqualification ruling distribution is the solid line. Panel C and D provide a visualisation of the estimated EMGD fits for the first round heats, (dashed line), and finals, (solid line), for both men and women.

\*\*\*\*Figure 2 near here\*\*\*\*

The probabilistic approach for identifying a difference across ruling periods and rounds is determined by calculating the probability of observing a RT greater than 120 ms. **The respective probabilities for ruling periods** are 0.95, 0.98 and 1.00 for men and 0.97, 0.99 and 1.00 for women. The probabilities of observing RTs above 120 ms in first rounds are 0.98 and 0.99 and in finals are 0.98 and 0.97 for men and women respectively. Table 3 provides the probability of

superiority for pairs of populations and in all cases the probability was above 50% indicating that the compared distributions differed.

\*\*\*\*Table 3 near here\*\*\*\*

Revision of the false start disqualification limit calculated from the historical data relevant to athletes' sex resulted in both thresholds exceeding the current 100 ms limit. Men's event data proposed a revised limit of 115 ms while the women's event data suggested a limit of 119 ms.

## Discussion

The median values for RTs at World and European Championships filtered by athletes' sex, showed absolute differences of 3 ms for men's and women's events. This small difference indicates that the different measurement technologies used at World and European Championships was not an important factor in this study and supports the decision to analyse pooled data from European and World Championships. The 95% coverage interval of RTs of the entire data shows a lower bound of 122 ms, 22 ms above the threshold set down by the IAAF. For a 100 m sprinter 22 ms can be a meaningful advantage. For example, in the 2012 Olympic Games in London the winning margin between the first and second athletes in the women's 100 m event was only 30 ms, and at a typical speed of 10 m/s, this difference represents a 30 cm advantage. If the 100 ms threshold fails to detect a false start, it is possible that an athlete could anticipate the start without disqualification and this could alter the results of an event.

***Ruling Periods***

The median and 95% coverage intervals for each ruling period shown in Table 1 indicate shifts in the distributions of the RTs of athletes in periods after a rule change by approximately 5 ms for men and 7.5 ms for women. While the shift in RTs of sprinters after the rule change in 2004 is not particularly noticeable in Figure 1, a shift does appear after the rule change in January 2010. This is consistent with the findings of Haugen, Shalfawi et al. (2013) and Tønnessen, Haugen et al. (2013) that were restricted to 100 m sprinters. This is also in line with the probability of superiority shown in Table 3 with men and women showing a greater probability of superiority after the second rule change. The comparison between the earliest and latest ruling periods is noteworthy with 72.1% and 74.5% probability of superiority for men and women respectively, indicating that shifts in distributions occurred between these ruling periods. Figure 2 panels A and B, suggest that for ruling periods, the RT of men and women has increased. It is clear that in both sexes, a shift has occurred in the distribution density after the automatic disqualification ruling in January 2010. The sex-based difference can also be identified as panels A and B of Figure 2 are on the same scale, with the densities for women shifted further to the right. The probability of observing a RT greater than 120 ms increases for men and women relative to the ruling periods. The probability of observations greater than 120 ms after the introduction of the automatic disqualification rule in 2010 is 1, irrespective of athletes' sex. This suggests that the minimum observed RT of athletes recorded using the current IAAF approved technology and rules is greater than 120 ms, some 20 ms above the enforced threshold.

### *Competition Rounds*

A comparison of the RTs in first round heats and finals is shown in panels C and D of Figure 2. A reduction of the RTs in finals suggests that athletes are more likely to anticipate the start signal when the race is more important, this is consistent with the observed shift in RTs by competition round shown in Tønnessen, Haugen et al. (2013) that examined only the 2003 – 2009 ruling period. A shift to the left indicating a decrease in RT is evident in both cases, suggesting that athletes are more likely to take risks in finals than in first round heats. This difference is evident in Table 3 where the probabilities of superiority for men and women are 0.656 and 0.638 respectively, therefore regardless of athletes' sex, RTs in finals are typically shorter than in first rounds. A sex-based difference is evident by the comparison of the plots with the RT distribution for women shifted further to the right. The probabilistic approach for the difference between RTs greater than 120 ms in first rounds and finals shows a relative difference for women, this could be because men take a risk at all levels whereas women require a marked incentive to take that risk (Whitley, Nelson et al., 1999).

### *Sex-based Difference*

The assertion by Lipps, Galecki et al. (2011) of a sex-based difference in RTs is evident throughout this study and can be further explained by the strength and rate of force development differences between men and women, allowing men to produce the required force to reach the false start threshold in a shorter time. There was a sex-based difference in RT both across ruling periods and competition rounds. This prominent sex-based difference supports the case for

independent false start detection limits for men and women as suggested by Lipps, Galecki et al. (2011). In addition this evident sex-based difference is consistent with the suggestion of Lipps, Galecki et al. (2011) that the force threshold for women should be decreased, while this paper does not address the percentage decrease of this threshold, we consider the estimated 22% on the RTs at the Beijing Olympics is relevant to all competitions. While this study has consistently reported a sex-based difference in RTs of athletes, this difference might be a consequence of the current starting-block sensor technology and event detection algorithms. The replacement of a force based threshold with an appropriate algorithm to detect initial rise in block force could establish whether a true sex-based difference exists in the RTs of elite-standard athletes.

**Rule Revision**

Revised RT thresholds were established for men and women independently by estimating an EMGD fit for each of the sexes. The revised threshold was calculated as the RT above which 99% of the observed RTs lay. Based on this and under the constraints of the current starting-block technology, we recommend that the RT threshold for men should be increased to 115 ms, and for women to 119 ms, because of the sex-based difference reported throughout this study and by Lipps, Galecki et al. (2011).

**Technology**

This paper has presented evidence that the 100 ms ruling as a disqualification limit for false starts in elite-standard athletics is liberal. This needs to be considered in the context of current technology that measures athletes' RT. The

IAAF approved technology does not record the **shortest** RT, or the first movement of the athlete, but rather the addition of the time to produce a predefined force along with the RT. Pain and Hibbs (2007) suggested that RTs of 85 ms are possible; this paper does not contradict these findings but rather suggests that these shorter RTs are similar to the time until the initial **response** of an elite athlete after a start signal.

### Conclusions

From this study it is clear that the IAAF 100 ms false start threshold is inadequate for the fair and impartial refereeing of **elite-standard** athletics, when using the IAAF approved **starting-block** systems. Similar to the work of Lipps, Galecki et al. (2011) a **sex-based** difference has also been identified from the historic data. It is suggested that the ruling for men and women be revised to account for the strength differences between the sexes. Revised RT thresholds, estimated from the historical data, both **for** men and women are provided in the results section.

In conclusion, this study has examined a large dataset of RTs of **international-standard** athletes, combined with an extensive literature review, **and** shown that the current governance of false start disqualification is inadequate under the current measurement system utilised by the IAAF.

### Disclosure Statement

No potential conflict of interest was reported by the authors.

Supplementary Material

Data and analysis scripts can be provided on request from the corresponding author.

References

Athletics Australia. (2009). "Implementation of IAAF "No False Start" Rule." from [http://www.easternsuburbs.org.au/assets/console/customitem/attachments/New\\_Start\\_Rule\\_Guidelines\\_141109.pdf](http://www.easternsuburbs.org.au/assets/console/customitem/attachments/New_Start_Rule_Guidelines_141109.pdf).

Brown, A. M., Z. R. Kenwell, B. K. V. Maraj and D. F. Collins (2008). ""Go" signal intensity influences the sprint start." *Medicine and Science in Sports and Exercise* **40**(6): 1142-1148.

Collet, C. (1999). "Strategic aspects of reaction time in world-class sprinters." *Perceptual and Motor Skills* **88**(1): 65-75.

Dawson, M. R. W. (1988). "Fitting the ex-Gaussian equation to reaction time distributions." *Behavior Research Methods, Instruments, & Computers* **20**(1): 54-57.

Der, G. and I. J. Deary (2006). "Age and sex differences in reaction time in adulthood: Results from the United Kingdom health and lifestyle survey." *Psychology and Aging* **21**(1): 62-73.

Greenwood, P. E. and M. S. Nikulin (1996). *A guide to chi-squared testing*, John Wiley & Sons.

Haugen, T. A., S. Shalfawi and E. Tønnessen (2013). "The effect of different starting procedures on sprinters' reaction time." *Journal of Sports Sciences* **31**(7): 699-705.

International Association of Athletics Federations. (2015). "Competition Rules 2016-2017." from <http://www.iaaf.org/about-iaaf/documents/reules-regulations>.

Komi, V., M. Ishikawa and S. Jukka (2009). "IAAF sprint start research project: Is the 100ms limit still valid." *New studies in athletics* **24**(1): 37-47.

Lipps, D. B., A. T. Galecki and J. A. Ashton-Miller (2011). "On the implications of a sex difference in the reaction times of sprinters at the Beijing Olympics." *PLoS ONE* **6**(10).

Massidda, D. (2013). "Retimes: Reaction Time Analysis." *R package version 0.1-2*.

Mero, A. and P. V. Komi (1990). "Reaction time and electromyographic activity during a sprint start." *European Journal of Applied Physiology and Occupational Physiology* **61**(1-2): 73-80.

Pain, M. T. G. and A. Hibbs (2007). "Sprint starts and the minimum auditory reaction time." *Journal of Sports Sciences* **25**(1): 79-86.

R Core Team (2015). "R: A Language and Environment for Statistical Computing (R Foundation for Statistical Computing, Vienna, 2012)." *URL*: <http://www.R-project.org>

Silver, J. D., M. E. Ritchie and G. K. Smyth (2009). "Microarray background correction: maximum likelihood estimation for the normal-exponential convolution." *Biostatistics*: kxn042.



Tønnessen, E., T. Haugen and S. A. I. Shalfawi (2013). "Reaction time aspects of elite sprinters in athletic world championships." Journal of Strength and Conditioning Research **27**(4): 885-892.

Whitley, B. E., A. B. Nelson and C. J. Jones (1999). "Gender differences in cheating attitudes and classroom cheating behavior: A meta-analysis." Sex Roles **41**(9-10): 657-680.

## Tables

- Table 1: Descriptive statistics for RTs of elite sprinters (ms).
- Table 2: Estimated parameters, distributional properties and goodness-of-fit statistics of the EMGD.
- Table 3: Probability of Superiority between distributions.  
(2nd group named is the "upper" distribution)

## Figures

- Figure 1: RTs across ruling periods.  
Box-and-whisker plots of valid RTs of sprinters across World and European Championships 1999-2014. Numbers circled are the number of disqualifications recorded as false starts.
- Figure 2: EMGD fits to RT data.  
Panels A and B present the estimated distributions for men and women respectively for the three ruling periods. The dotted line in each plot relates to the ruling period from 1999-2003, the dashed line represents the introduction of the group warning in January 2004, while the solid line represents the current ruling of automatic disqualification introduced in January 2010. Panels C and D present the estimated distributions for men and women respectively for first round heats and finals. The dashed line represents the first round RTs of athletes between 1999 and 2014, while the solid line represents the RTs in the finals of these same events.

Sex	Ruling Period	N	Median	95% Coverage Interval	No. False Starts
Men	1999-2003	1407	151.0	(116.0, 230.0)	1
	2004-2009	1545	155.0	(121.0, 250.0)	0
	2010-2014	1608	161.0	(128.0, 282.7)	8
Women	1999-2003	1160	153.0	(119.0, 244.0)	0
	2004-2009	1237	161.0	(126.2, 256.8)	0
	2010-2014	1602	168.0	(130.0, 278.0)	1
Total	1999-2014	8559	159.0	(122.0, 259.0)	10

Table 1: Descriptive Statistics for RT data (ms).

Sex	Ruling Period	Parameters			Properties		Goodness-of-fit	
		$\mu$	$\sigma$	$\tau$	$E(X)$	$s(X)$	$\chi^2$	p-value
Men	1999-2003	132.80	14.66	21.97	154.77	26.41	16.03	0.0136
	2004-2009	134.76	11.57	27.76	162.52	30.07	42.87	<0.0001
	2010-2014	139.95	10.92	32.33	172.28	34.12	60.25	<0.0001
	First Rounds	134.96	12.61	33.60	168.56	35.89	68.53	<0.0001
	Final Rounds	137.30	12.53	15.21	152.51	19.71	11.57	0.0723
Women	1999-2003	131.32	12.14	27.62	158.94	30.17	8.95	0.1765
	2004-2009	138.32	11.91	31.24	169.56	33.43	3.80	0.7039
	2010-2014	141.82	11.81	36.51	178.33	38.37	13.73	0.0328
	First Rounds	137.89	12.11	38.24	176.13	40.11	7.60	0.2688
	Final Rounds	136.90	14.04	24.15	161.05	27.93	7.97	0.2403

Table 2: Estimated parameters, distributional properties and goodness-of-fit statistics of the EMGD.

Sex	Comparison	Probability of Superiority
Men	1999-2003 v 2004-2009	0.591
	2004-2009 v 2010-2014	0.631
	1993-2003 v 2010-2014	0.721
	First Rounds v Finals	0.656
Women	1999-2003 v 2004-2009	0.684
	2004-2009 v 2010-2014	0.692
	1993-2003 v 2010-2014	0.745
	First Rounds v Finals	0.638

Table 3: Probability of Superiority between distributions.  
(the 2<sup>nd</sup> group named is the “upper” distribution)

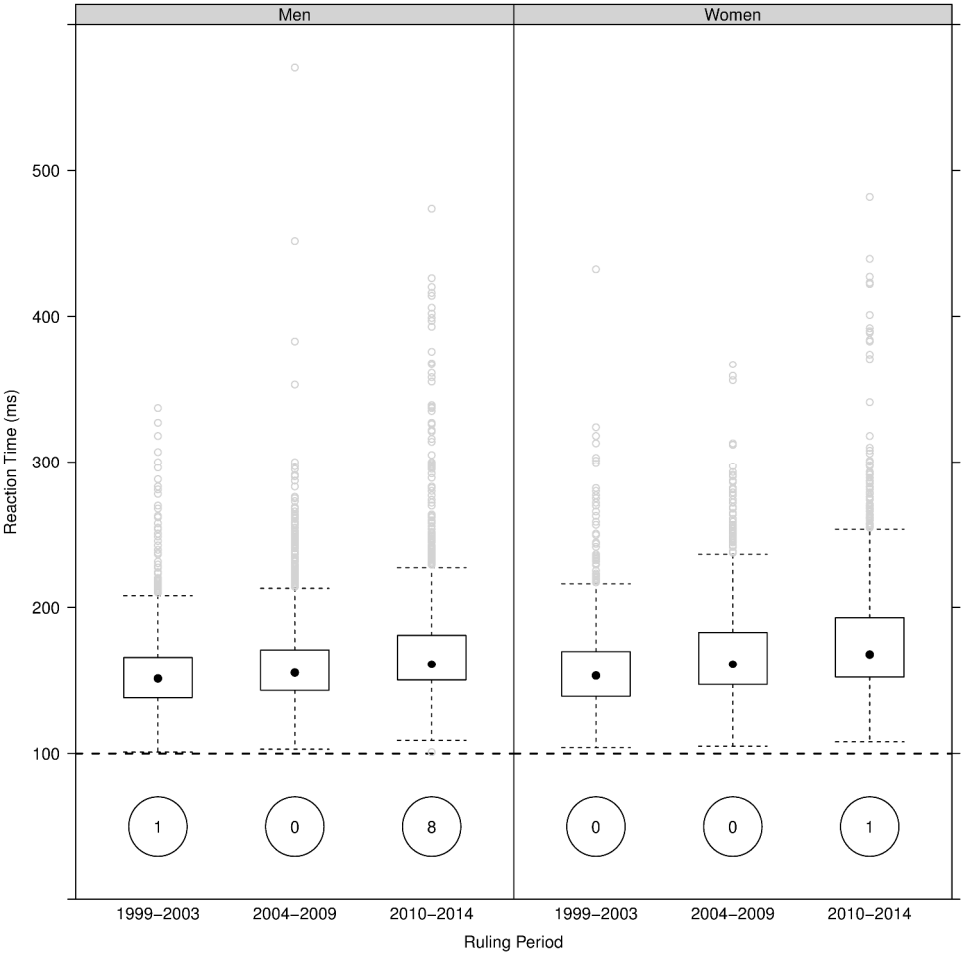


Figure 1: RTs across ruling periods. Box-and-whisker plots of valid RTs of sprinters across World and European Championships 1999-2014. Numbers circled are the number of disqualifications recorded as false starts.

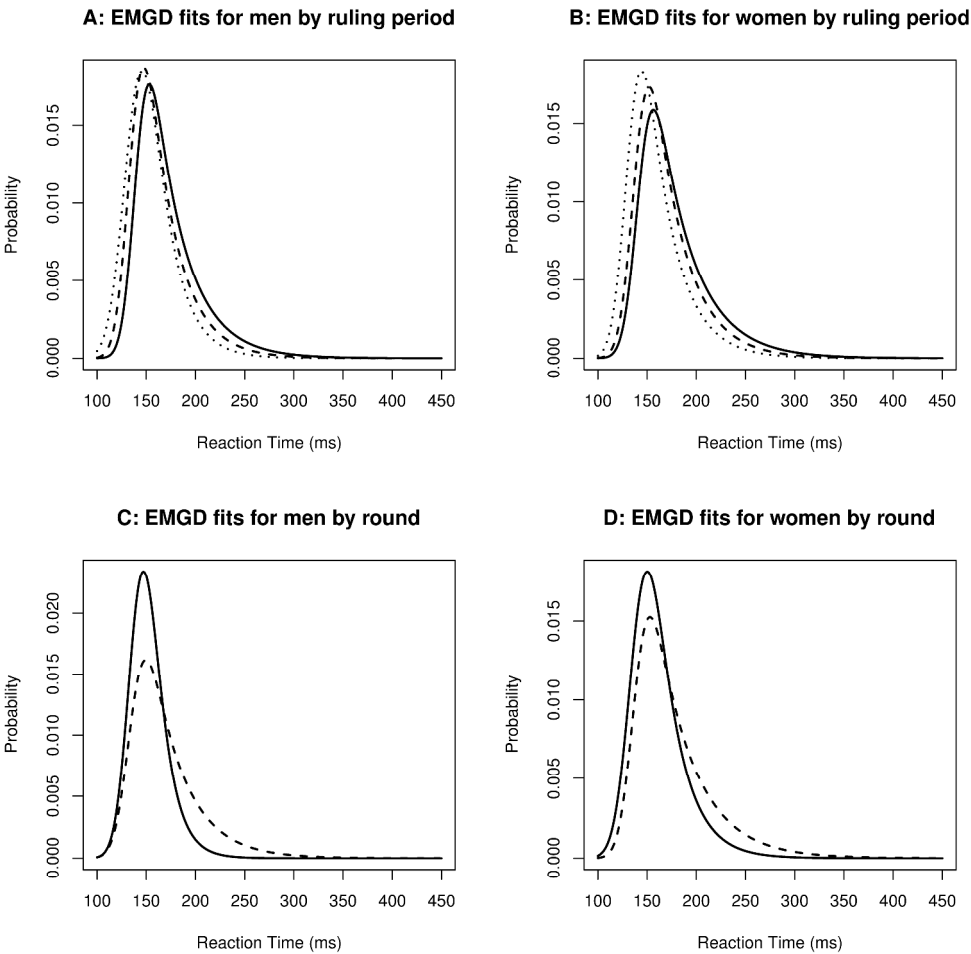


Figure 2: EMGD fits to RT data.

Panels A and B present the estimated distributions for men and women respectively for the three ruling periods. The dotted line in each plot relates to the ruling period from 1999-2003, the dashed line represents the introduction of the group warning in January 2004, while the solid line represents the current ruling of automatic disqualification introduced in January 2010. Panels C and D present the estimated distributions for men and women respectively for first round heats and finals. The dashed line represents the first round RTs of athletes between 1999 and 2014, while the solid line represents the RTs in the finals of these same events.