Effects of false-start disqualification rules on response-times of elite-standard sprinters

Kevin C. Brosnan, Kevin Hayes & Andrew J. Harrison

To cite this article: Kevin C. Brosnan, Kevin Hayes & Andrew J. Harrison (2016): Effects of false-start disqualification rules on response-times of elite-standard sprinters, Journal of Sports Sciences, DOI: 10.1080/02640414.2016.1201213

To link to this article: http://dx.doi.org/10.1080/02640414.2016.1201213

Published online: 28 Jun 2016.
Effects of false-start disqualification rules on response-times of elite-standard sprinters

Kevin C. Brosnan a, Kevin Hayes a and Andrew J. Harrison b

aDepartment of Mathematics and Statistics, University of Limerick, Limerick, Ireland; bDepartment of Physical Education and Sports Science, University of Limerick, Limerick, Ireland

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 1990s. 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 (RT) data from 1999 to 2014 to examine effects of rule changes on competition RT at major championships. The exponentially modified Gaussian distribution was used to model RT and make comparisons relative to athletes’ sex, ruling periods and competition rounds. Revised RT 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 RT in elite-standard athletes.

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, & Ashton-Miller, 2011). A RT in the context of this article 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. Much of the literature in the area of false-start detection in elite athletics refers to reaction times; however, the correct term is response-times (for more detail see Mero and Komi, 1990). 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 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 & Komi, 1990). Pain and Hibbs (2007) suggested that genuine RTs of 85 ms are possible and Brown, Kenwell, Maraj, and Collins (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, and Jukka (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 et al., 2011) and of the general population (Der & 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, and Shalfawi (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, and Tønnessen (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 1999–2014 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 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

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

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 \(\min(\text{RT}) \leq \mu \leq \max(\text{RT}), 0 \leq \sigma \leq \text{range}(\text{RT})\) and \(0 \leq \tau \leq \text{range}(\text{RT})\). If RTs follow an EMGD with the above parameters, its 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 by Silver, Ritchie, and Smyth (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

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

where $T$ is the threshold of interest, $\mu$, $\sigma$ and $\tau$ are the pa-
rameters 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 pre-
presented earlier identified the threshold ($T$) above which 99% of the
observed RTs lie. 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 posi-
tive 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>
<thead>
<tr>
<th>Sex</th>
<th>Ruling period</th>
<th>N</th>
<th>Median</th>
<th>95% coverage interval</th>
<th>No. false starts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Men</td>
<td>1999–2003</td>
<td>1407</td>
<td>151.0</td>
<td>(116.0, 230.0)</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>2004–2009</td>
<td>1545</td>
<td>155.0</td>
<td>(121.0, 250.0)</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>2010–2014</td>
<td>1608</td>
<td>161.0</td>
<td>(128.0, 282.7)</td>
<td>8</td>
</tr>
<tr>
<td>Women</td>
<td>1999–2003</td>
<td>1160</td>
<td>153.0</td>
<td>(119.0, 244.0)</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>2004–2009</td>
<td>1237</td>
<td>161.0</td>
<td>(126.2, 256.8)</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>2010–2014</td>
<td>1602</td>
<td>168.0</td>
<td>(130.0, 278.0)</td>
<td>1</td>
</tr>
<tr>
<td>Total</td>
<td>1999–2014</td>
<td>8559</td>
<td>159.0</td>
<td>(122.0, 259.0)</td>
<td>10</td>
</tr>
</tbody>
</table>

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.

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 an-
anticipate the start signal. The parameters column pre-
presented in Table 2 provides the best estimates of the EMGD
parameters given the data from the ruling period and ath-
letes’ sex, the properties column shows the location and
scale of the distribution, while the final two columns pro-
vide 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.

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.
Panels 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.

The probabilistic approach for identifying a difference
across ruling periods and rounds is determined by calculat-
ing 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 pro-
ability was above 50% indicating that the compared distribu-
tions differed.

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⁻¹, 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 et al. (2013) and Tønnessen 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. Panels A and B in Figure 2 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

<table>
<thead>
<tr>
<th>Sex</th>
<th>Ruling period</th>
<th>Parameters</th>
<th>Properties</th>
<th>Goodness-of-fit</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>µ</td>
<td>σ</td>
<td>τ</td>
</tr>
<tr>
<td>Men</td>
<td>1999–2003</td>
<td>132.80</td>
<td>14.66</td>
<td>21.97</td>
</tr>
<tr>
<td></td>
<td>2004–2009</td>
<td>134.76</td>
<td>11.57</td>
<td>27.76</td>
</tr>
<tr>
<td></td>
<td>2010–2014</td>
<td>139.95</td>
<td>10.92</td>
<td>32.33</td>
</tr>
<tr>
<td></td>
<td>First rounds</td>
<td>134.96</td>
<td>12.61</td>
<td>33.60</td>
</tr>
<tr>
<td></td>
<td>Final rounds</td>
<td>137.30</td>
<td>12.53</td>
<td>15.21</td>
</tr>
<tr>
<td>Women</td>
<td>1999–2003</td>
<td>131.32</td>
<td>12.14</td>
<td>27.62</td>
</tr>
<tr>
<td></td>
<td>2004–2009</td>
<td>138.32</td>
<td>11.91</td>
<td>31.24</td>
</tr>
<tr>
<td></td>
<td>2010–2014</td>
<td>141.82</td>
<td>11.81</td>
<td>36.51</td>
</tr>
<tr>
<td></td>
<td>First rounds</td>
<td>137.89</td>
<td>12.11</td>
<td>38.24</td>
</tr>
<tr>
<td></td>
<td>Final rounds</td>
<td>136.90</td>
<td>14.04</td>
<td>24.15</td>
</tr>
</tbody>
</table>

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

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.
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 by Tønnessen 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 risk at all levels whereas women require a marked incentive to take that risk (Whitley, Nelson, & Jones, 1999).

**Sex-based difference**

The assertion by Lipps 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 et al., (2011). In addition, this evident sex-based difference is consistent with the suggestion of Lipps et al., (2011) that the force threshold for women should be decreased; while this article 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 et al., (2011).

Technology

This article 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 article 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 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 data set 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.

Acknowledgement

Kevin C. Brosnan was supported by an Irish Research Council Government of Ireland Scholarship [grant number GOIPG/2014/19]. Kevin Hayes was supported by Science Foundation Ireland Research Investigator’s Award [grant number SFI-12/IA/1683].

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

Kevin C. Brosnan was supported by an Irish Research Council Government of Ireland Scholarship [grant number GOIPG/2014/19]. Kevin Hayes was supported by the Science Foundation Ireland Research Investigator’s Award [grant number SFI-12/IA/1683].

ORCID

Kevin C. Brosnan http://orcid.org/0000-0002-0952-2134
Andrew J. Harrison http://orcid.org/0000-0002-5569-4885

References


