Talent selection criteria for olympic distance triathlon

ALESSANDRO BOTTONI1, ANTONIO GIANFELICI2, ROBERTO TAMBURRI1, MARCELLO FAINA2

1Italian Triathlon Federation, Rome, Italy
2Sport Medicine and Science Institute, Italian National Olympic Committee, Rome, Italy

ABSTRACT

Bottoni A, Gianfelici A, Tamburri R, Faina M. Talent selection criteria for olympic distance triathlon. J. Hum. Sport Exerc. Vol. 6, No. 2, pp. 293-304, 2011. Talent Selection allows to optimize the resources available for sporting talent in order to design the best strategy to achieve top level sporting results. Because of the unknown aspects of the performance model in Olympic triathlon the TS variables and their relationship with a future performance are far-off from being identified in order to make a talent prospective study possible. Currently most triathlon federations evaluate only the juvenile performance expressed in time trials test on swimming and running. The aim of the present study was to find the most appropriate variables for the Talent Selection in Olympic Triathlon, verifying those widely used by means of a retrospective research about particular juvenile features recognized in top world triathlon athletes. The variables are considered as input variables of a Talent selection model based on Fuzzy Logic that overcome the limits of traditional models based on cut-off selection. The present findings indicate that the exclusive evaluation of juvenile running and swimming performance in order to select triathlon talent is not appropriate. Diagnosis criteria should include several other variables that should also take into account mental ability, speed of abilities development, utilization of endogenous and exogenous resources, load and stress tolerance as several leading countries have done recently. Key words: FUZZY LOGIC, TALENT DIAGNOSIS, TALENT PROGNOSIS, TALENT IDENTIFICATION, EXPERT SYSTEM.
INTRODUCTION

The development of talent is a subject which creates great interest in today’s sporting associations. The identification and the capacity to promote talent assume an important and fundamental role in order to achieve significant results in an international field. Even more so, bearing in mind that Triathlon Federations often have to deal with limited numbers of young athletes available both in quantity and quality together with scarce availability of funds. This aspect is made worse by the present socio-sport background whereby juvenile sport activity is declining in many countries (Dolmann et al., 2005) and where even the base motor activity as well as the aerobic fitness have a falling trend (Tomkinson et al., 2003). From this point of view Talent Selection (TS) in Triathlon assumes great relevance as it allows to optimize the resources available for talent recruitment and talent promotion in order to design the best strategy to achieve top level sporting results. In this study the attention is focalized on talent selection. Other aspects such as recruitment and talent identification require a considerable amount of resources often not available for federations.

Regarding TS, the Science has given only a partial contribute to this complex issue over the year. The concept of TS has evolved from the evaluation of youth performance and abilities to a larger and dynamically interpreted evolving process (Abott, 2002). In the beginning, talent was just considered as a person who exhibits above average sporting performances whereby the genetic pattern plays a leading role. Recently it was shown that it is extremely improbable the existence of a perfect genetic profile as is extremely improbable to find a person with totally disadvantageous genetic profile (Williams & Folland, 2008). It was even stated that despite the enormous amount of attention that the performance's genes has received it is not possible to conclude whether the polymorphism is involved in performance phenotype (Bray et al., 2009). The study of genetic profile is not actually useful for talent detection (McArthur et al., 2005). In endurance sport a lot of attention was reserved to the aerobic variables like VO$_2$max, even in triathlon where this parameter is difficult to measure in race condition and could be specific to the exercise modality (Hue et al., 2000; Millet et al., 2009). Some studies did not find a significant difference between well trained elite athletes and top elite national ranked athletes (Gianfelici et al., 2009) and it was not possible to predict if young cyclist could develop a professional career from the valuation of the aerobic variables stating that these variables cannot be used for talent identification (Menaspa et al., 2010). There is a scientific community full consent about the importance of the environment and it was pointed out that gene and environment interact, not just over the short term, but also over the lifetime of an individual with permanent effects on the adult phenotype (Brutsaer & Parra, 2006).

Currently not only is it admitted that the influence of the deliberate practise (Ericsson et al., 1993) but also scientific community is led to believe that trainability is the most important part of the natural-born factor and there are a lot of evidences that peak performance is connected with the total amount of training. Taking in account all the above mentioned elements in order to define selection criteria it is a complex issue. Recently (Hohmann et al., 2003), four set of criteria has been considered: juvenile performance, speed of development of complex and partial performance, rate of utilization of resources and stress and load tolerance. Currently many Triathlon Federations follow a selection policy based on time trial in swimming and running for athletes from 14 to 16 years old. The tests mostly used are 200m swimming, 400m swimming, 2000m running and 3000m running. Some countries use also a trial test in cycling. Recently some sporting leader countries have introduced training load, stress tolerance and mental performance as determinant factor in TS.
Several studies tried to investigate the performance and the physiological characteristic in triathlon (Vleck et al., 2008; Millet et al., 2000, 2003, 2005 and 2009). Because of the unknown aspects of the performance model in Olympic triathlon the Talent Selection variables and their relationship with future performance are far-off from being identified in a prospective study. Some previous studies proposed a different retrospective approach (Hohmann, 2001) pointing out that Talent in Sport is a person that has retrospectively already shown peak performances in the past rather than a young person that is prospectively expected by means of scientific prognosis because of the exhibition of actual sports performances and performance prerequisites above average when compared to athletes of a similar age and living conditions. As Hohmann stated, the traditional selection model based on cut-off present several limits. The cut-off value define the expected groups composed by unexpected-non-achievers and expected high achievers but even other two groups composed by expected- non achievers and unexpected high achievers. A selection policy based on too much restrictive cut-off produce an unacceptable high number of not expected high achievers, while a low cut-off limit produces an unacceptable waste of resources gived to the expected - non achievers. In addition, once identified the variables, the traditional model doesn’t allow the variables integration and composition and each variables need to be processed independently. Recently a different approach, based on the development of an expert system for talent detection, was proposed (Rogulj et al., 2006, 2007) applying methodology of knowledge acquisition that in the final evaluation of the tested person has shown good results, founding a good correlation between the output results of the test measurements and the expert evaluations. Later was proposed (Papić et al., 2009) an approach based on Fuzzy Logic, a multi-valued logic based on fuzzy set and linguistic variables (Feng, 1984). Despite the use of fuzzy set which seem to be less accurate, Fuzzy Logic Model is used to control a very complicated process and as a expectation and diagnosis tool. Papić has compared two different approaches which achieved similar results in the recognition of a sport compatible for tested individuals: approach based on the kinesiology experts’ opinion on the compatibility of proposed hypothetical morphological model of a particular sport and the approach based on detection of differences between morphological characteristics of tested individuals and morphological characteristics of top athletes in various sports. This approach requires, in order to achieve good results, an onerous planning effort. According to Papić the improvements in performance of Fuzzy Model, based on expert opinion approach, requires a number of experts involved in research and a process of modeling and the adjusting of the membership functions as well as updating the normative values used for the specification of fuzzy sets.

MATERIAL AND METHODOLOGY

A group of 21 variables including juvenile performance, performance development, utilization of resources and training load, were identified and proposed in a questionnaire for a pilot study. The aim of the pilot study was to identify the best variables to study and question to propose in the retrospective research questionnaire. Regarding the juvenile performance both race performance and time trials test were considered. The juvenile time trial distances considered was 200m (t200) and 400m (t400) for swimming and 2000m (t2000) and 3000m (t3000) for running. The pilote questionnaire was gived to coaches and other people who work for national federations as youth coaches, involved in juvenile activity or in a federal project about talent, whose formation and competence allows to form an expert opinion about talent and talent selection. A synthesis of this study is reported in Figure 1 where is shown how the juvenile time trial performance is considered just one of the elements that should be taken in account. From the pilot study a questionnaire has been planned for the retrospective research about all the variables mentioned above. The personal best time on running (t5000, t10000) and on swimming (t1500) were also considered. The questionnaire also investigated the development of the performances during the past years and the volume (hours) of training per week at the different ages of development. As index of speed of performance
development (ISD) is considered the relative difference between personal best time at 14 y.o. and absolute personal best. A retrospective research until the age of 14 was extended to the top world level triathletes who have already shown peak performance. In other words those athletes have already shown their talent. Due to the rapid evolution of triathlon performance model over the past year (Vleck et al., 2008; Bernard et al., 2009) the retrospective study was backed up by athletes who became champions after the year 2000. The questionnaire was given to 66 male top world athletes (TW) in Olympic triathlon. Those considered were the top 5 in the World Championships, European Championships and Olympic Games from 2000 to 2008 (33% responses). The top 15 Italian athletes (TI) during the same period (93% responses) were also considered. At youth ages TW and TI were homogeneous for level. Statistical analysis was performed by the Mann-Whitney U test.

Figure 1. Rate of importance (from 0 up to 10) for coaches and other expert people of several possible diagnostic variables.
RESULTS

There were no significant differences in youth performance between TW and TI especially in running (Figure 2, Table 1), and TI swimming performances were even better at 16 y.o. (percent difference t400=3.6%, p<0.1). The same trend in swimming is maintained until 18 y.o. (percent difference t400=3%, p<0.1). In elite ages (Table 2, Figure 3) TI showed a higher value of t5000 (+3.9%, p<0.005) and t10000 (+3.3%, p<0.005) while the minimum insignificant difference of t1500 does not seem to be reflected in the race. Low correlation between juvenile test trials and the absolute performance in the race was observed, even combined the juvenile performance in some combined time trial event as some countries recently done (Figure 4). Regarding the speed of performance development TW showed (Table 3, Figure 5) higher values of ISD for t400 (TW:ISD=134±24·10^-3; TI:ISD=75±32·10^-3;p<0.001) and t3000 (TW: ISD=86±28·10^-3; TI: ISD=55±14·10^-3; p<0.005). Part of the questionnaire dealt also with volume training per week at different ages. In swimming and running TW showed more hours of training per week than TI for all ages of development (Figure 6). Training volume was tested only for the elite age and the difference in the amount of hour of training per week resulted particularly high (TW training volume was 28% greater than TI, p<0.01). The difference in results were also very high in running until 14 y.o (TW training volume was 36.3% greater than TI, p<0.01), from 14 y.o up to 16 y.o. (TW training volume was 33.2% greater than TI, p<0.01) and up to 18 y.o. (TW training volume was 20.1% greater than TI, p<0.05).

![Figure 2. Juvenile performance in swimming (t200, t400) and running (t2000, t3000) expressed in trial time.](image)

**Table 1.** Differences in time trial test between TW and TI over the years of development.

<table>
<thead>
<tr>
<th>Trial Test</th>
<th>age</th>
<th>TW (s)</th>
<th>TI (s)</th>
<th>difference (s)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Swim 200m</td>
<td>16</td>
<td>129.6±8.7</td>
<td>128.5±8.9</td>
<td>-1.1, -1.50</td>
<td>&gt;0.1</td>
</tr>
<tr>
<td>Run 2000m</td>
<td></td>
<td>358.4±4.9</td>
<td>361.7±5.6</td>
<td>3.3, 0.90</td>
<td></td>
</tr>
<tr>
<td>Swim 400m</td>
<td>14</td>
<td>288.8±9.2</td>
<td>283±8.8</td>
<td>-6.2, -2.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>16</td>
<td>278.6±8.3</td>
<td>268.8±7.2</td>
<td>-10.2, -3.6</td>
<td></td>
</tr>
<tr>
<td></td>
<td>18</td>
<td>269.5±6.3</td>
<td>261±5.4</td>
<td>-8.5, -3</td>
<td>&lt;0.1</td>
</tr>
<tr>
<td>P.B.</td>
<td></td>
<td>256.4±10.9</td>
<td>260.4±13.6</td>
<td>4, 1.2</td>
<td></td>
</tr>
<tr>
<td>Run 3000m</td>
<td>14</td>
<td>575.8±12.1</td>
<td>581.0±9.1</td>
<td>5.2, 0.9</td>
<td>&gt;0.1</td>
</tr>
<tr>
<td></td>
<td>18</td>
<td>552±13.2</td>
<td>553.5±12.9</td>
<td>1.5, 0.3</td>
<td></td>
</tr>
<tr>
<td>P.B.</td>
<td></td>
<td>539.4±12.7</td>
<td>537.6±14.5</td>
<td>-1.2, -0.2</td>
<td>&lt;0.1</td>
</tr>
</tbody>
</table>
Figure 3. Differences in running and swimming personal best time between TW and TI

Table 2. Differences in personal best time between TW and TI.

<table>
<thead>
<tr>
<th>Test Trial</th>
<th>Personal Best</th>
<th>difference (s)</th>
<th>(%)</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Swim 400m</td>
<td>TW 256.4±10.9</td>
<td>4s 1.2</td>
<td>&gt; 0.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>TI 260.4±13.6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Swim 1500m</td>
<td>TW 1026.7±29.8</td>
<td>5s 0.5</td>
<td>&gt; 0.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>TI 1031.7±52.9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Run 5000m</td>
<td>TW 871.4±17.1</td>
<td>33.6s 3.9</td>
<td>&lt; 0.005</td>
<td></td>
</tr>
<tr>
<td></td>
<td>TI 905.0±18.7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Run 10000m</td>
<td>TW 1808.4±42.1</td>
<td>60.5s 3.3</td>
<td>&lt; 0.005</td>
<td></td>
</tr>
<tr>
<td></td>
<td>TI 1868.9±22.4</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 4. Correlation between juvenile performance (time trial) and elite performance (mean world cup position).
Figure 5. Relative performance gain from 14 y.o. to elite ages.

Table 3. Percent differences in index of speed of development (ISD) between TW and TI.
The aim of this study was to investigate if the exclusive evaluation of juvenile running and swimming performance in order to select triathlon talent is appropriate. A questionnaire was submitted for a retrospectively research on the athletes who have shown their talent. The questionnaire dealt with juvenile performance in swimming and running and it also investigated the development of the performances over the years and the volume of training per week at the different ages of development. Other parts of the questionnaire dealt with the environment of development, the support provided by family, club and federation and about how athletes and coaches approach towards training. In Figure 7 is shows an example of some differences in the gived answers between TW and TI referring at their youth ages and at their development. In particular it emphasizes the approach towards training and races. Overall the present findings indicate that the exclusive evaluation of juvenile running and swimming performance in order to select triathlon talent is not appropriate and is unable to anticipate a future peak performance. Diagnosis criteria should include several other variables that also take into account mental ability, speed of abilities development, utilization of endogenous and exogenous resources, load and stress tolerance as several leading countries have done recently.

**Figure 6.** Percent difference in training volume over the years of development between TW and TI (positive if training volume of TW is greater than TI).

**DISCUSSION**

The aim of this study was to investigate if the exclusive evaluation of juvenile running and swimming performance in order to select triathlon talent is appropriate. A questionnaire was submitted for a retrospectively research on the athletes who have shown their talent. The questionnaire dealt with juvenile performance in swimming and running and it also investigated the development of the performances over the years and the volume of training per week at the different ages of development. Other parts of the questionnaire dealt with the environment of development, the support provided by family, club and federation and about how athletes and coaches approach towards training. In Figure 7 is shows an example of some differences in the gived answers between TW and TI referring at their youth ages and at their development. In particular it emphasizes the approach towards training and races. Overall the present findings indicate that the exclusive evaluation of juvenile running and swimming performance in order to select triathlon talent is not appropriate and is unable to anticipate a future peak performance. Diagnosis criteria should include several other variables that also take into account mental ability, speed of abilities development, utilization of endogenous and exogenous resources, load and stress tolerance as several leading countries have done recently.
Above all, it is not known whether the T1 group, given the same environmental stimuli as the TW group, would have had similar results. However, bearing in mind that the environmental stimuli that world class athletes have received are those to pursue, the selection therefore should be able to evaluate if athletes are able to use and improve in the best possible way, under these particular stimuli. It is obvious therefore that talent selection is a dynamic process in constant evolution and that there should be a constant assessment using all the previously mentioned factors, comparing not only the quality but also the amount of stimuli deriving from the environment which indicates the development and growth of each athlete. This, in turn, creates the necessity to manage a model with many variables of different nature, some of which are hard to measure.

Although these variables can be obtained from a retrospective research they are difficult to integrate in the traditional models based on cut-off criteria. What is required is a model able to consider at the same time a very well measurable variables such as performance time, other measurable variables such as mental toughness, and linguistic variables such as “quite confident with swimming technique” or “high commitment in training”. Model based on Fuzzy logic allows the the integration of these linguistic variables with traditional ones and is able to give a single numerical output variable that can express the rate of talent. Despite the use of fuzzy set which seems to be less accurate, Fuzzy Logic model is used to control very complicated process and as a expectation and diagnosis tool. The model work on variables and rules with processes similar tho those intuitively used by expert operators. The Figure 8 shows a possible group of
variables to take in account as input variables. The fuzzy set with membership function for a variable such as t200, with the linguistic meaning of the top values, is showed in Figure 9. Even if the triangular function is the most used any other shape could be considered. The x-axis is defined by real values that are derived from defined criteria such as the values derived from the cumulative distribution function associated to the TW population. During the fuzzyfication (matching), crisp values are transformed into grades of membership for linguistic terms of fuzzy sets, as showed in Figure 10 where t200 is defined as “GO with rate 0.7 and VG with rate 0.3”. In this form traditional numeric variables like t200 can be easily associated to linguistic variable like "swimming skills". Fuzzy set theory defines fuzzy operators on fuzzy sets. The rules are in the form of IF-THEN statements, where the IF part is called the "antecedent" and the THEN part is called the "consequent". Through fuzzy rules and the use of defined operators the inferential process modify the antecedents combined using fuzzy operators and generate the consequents. The rules are defined by expert statement, are built on the synthesis of all the experiences and they are based on composition of proposition. For example, considering the mental toughness (MT), the training tolerance in swimming (X), the associated fuzzy set and the membership functions, the inferential process act as is showed in Figure 11 where t200 and MT are the antecedents and X is the consequent. This case has used the operator lowest and the consequent X maintains the rate 0.3 associated at the linguistic value DE. It could mean that despite a quite good swimming performance in 200m, an athlete with a quite low mental toughness is considered insufficient in tolerating high training loads in swimming. Typical fuzzy control systems have dozens of rules solved in parallel or sequentially. The results of all the rules are defuzzified to a crisp value. There are several methods to perform defuzzification but the most common considers the center of mass of the result. At the end of the process the model give the output value on a real axis (rate of talent) obtained operating according to the conceptual models followed by experienced operators.
CONCLUSION

The present findings indicate that the exclusive evaluation of juvenile running and swimming performance in order to select triathlon talent is not appropriate and is unable to anticipate the possibility of reaching a peak performance in the future. In order to avoid any waste of talent federations are forced to fix a swimming and running cut off in a not very restrictive way but this means a great waste of resources for talent promotion. Effective diagnosis criteria, on the other hand, should include several other variables that should also take into account mental ability, speed of abilities development, utilization of endogenous and exogenous resources, load and stress tolerance just as several leading countries have started to do recently. In addition talent Selection should be considered as a dynamic process where the environment of development is a crucial elemen to valuate. In order to perform a more effective talent selection the traditional cut-off model result inadequate to manage and integrate quantifiable with unquantifiable and linguistic variables. The Fuzzy Logic model is rather better as it allows the integration of all these variables and deduces the rate of talent as an output, operating according to the conceptual models followed by experienced operators.

REFERENCES

6. ERICSSON KA, KRAMPE RT, TESCH-RÖMER, C. The role of deliberate practice in the acquisition of expert performance. Psychol Rev. 1993; 100:363-406. [Full Text] [Back to text]