Reproducibility and Variability of the Conconi Test

M. Strupler¹, G. Mueller², A. Frotzler² and C. Perret¹

1. Institute of Sports Medicine, Swiss Paraplegic Centre, CH-6207 Nottwil, Switzerland
2. Clinical Trial Unit, Swiss Paraplegic Centre, CH-6207 Nottwil, Switzerland

Abstract: The Conconi test is an incremental exercise test characterised by stages of equal work. For analysis PO (power output) and HR (heart rate) are used. The deflection point of the HR/PO graph marks the point where the linear relation between PO and HR changes to a curvilinear one. HR at deflection point correlates with the anaerobic threshold. The objective of this study was to evaluate the reproducibility of HR at deflection point of tests in cycling. In literature, only one study with tests in running examined the reproducibility of the Conconi test. Nineteen trained subjects performed 4 Conconi tests on an electronically breaked cycle ergometer on different days. Finally, 68 graphs of the PO/HR relationship were analysed by computer software as well as by three different observers. HRs at deflection point of the 4 Conconi tests of every subject were compared and coefficients of variation (CVs) were calculated. Mean CVs of HR at deflection point per subject were 3.36% for the computer program and for the three observers 2.99%, 2.51% and 2.53%, respectively. CVs below 3.5% represent a good reproducibility. The agreement between the three observers was statistically good with an ICC (intraclass correlation coefficients) of 0.828. The mean range of HR at deflection point per Conconi test comparing the three observers was 4.5 bpm corresponding 2.7% of HR at deflection point.

Key words: HR, deflection point, anaerobic threshold, visual analysis.

1. Introduction

HR (heart rate) is a useful tool to guide the intensity of endurance training. Maximal HR and HR at a given intensity (i.e. at maximal lactate steady state) show large interindividual differences [1]. Therefore, exercise testing is required to optimally guide endurance training. A test proposed by Conconi et al. [2] (CT (Conconi test)) determines the relationship between speed and HR during an incremental exercise test in running. Originally, Conconi described an incremental test for running on a 400 m track with bouts of 200 m and increments of 0.5 km/h. Probst et al. [3] adapted this test for cycling on an electronically braked ergometer with bouts of equal work. CTS are often used to guide training programs for athletes of any kind of performance level. The relationship between PO (power output) and HR is linear in low to moderate workloads, but changes to a curvilinear shape at submaximal to maximal workloads. The DP (deflection point) marks the transition from the linear to the curvilinear portion of the PO/HR curve. According to Conconi et al., HR at DP (HRdp) corresponds to the anaerobic threshold and can be identified by this test in a non invasive way. In literature, the validity of the CT is discussed controversially [4-8]. Discussion is related to the methodology of the test protocol and the reproducibility of the deflection point. Often the test protocols were changed in other studies, what probably influenced test results [9].

However, for the validity of a test the reproducibility is essential. The reproducibility of HRdp was evaluated only in one study with CTS in running [10]. Therefore, the objective of this study was to evaluate: (1) the reproducibility of HRdp of 4 consecutive CTS of 19 subjects performed on a cycle ergometer using the protocol described by Probst et al. and; (2) to find out if

Corresponding author: Perret Claudio, PhD, research fields: respiratory and exercise physiology.
there is a difference of HR$\text{dp}$ when analysing the PO/HR graph either by a computer program or visually by three experienced observers.

2. Subjects and Method

2.1 Subjects

The 19 subjects (13 male; 6 female) with an endurance training volume of at least 3 times 45 min per week attended the study. Their characteristics were: age $30 \pm 8$ years, height $176 \pm 9$ cm, weight $67 \pm 9$ kg and maximal oxygen uptake $59 \pm 8$ ml min$^{-1}$·kg$^{-1}$. They accomplished 4 identical CTs on an electronically braked stationary cycle ergometer (Ergometrics 900, Ergoline, Bitz, Germany). The time between each test was at least 2 and maximally 7 days. The subjects refrained from intensive trainings the day before the test. Nutrition and caffeine intake were kept constant and recorded. The study was approved by the ethics committee of the canton of Lucerne, Switzerland. Written informed consent was obtained from each subject prior to the start of the study.

2.2 Experimental Procedure

For the CT, the adapted protocol for cycle ergometry by Probst et al. [3] was used. Shortly, after a warm-up of 5 min at 100 W, the incremental protocol started with 100 W for 2 min and continued with increments of 20 W and constant work per bout (12 kJ). The cadence could be chosen individually between 70 and 100 revolutions per min during the warm-up and was kept constant during the CT. The subjects were given strong verbal encouragement to perform to the maximum. During the tests HR was recorded every 5 sec by a Polar HR monitor (Polar S610i, Kempele, Finland). The average HR during the last 15 sec of every bout was used for analysis in relation to the corresponding PO.

2.3 Statistics

The graphs with the PO/HR relationship were analysed by a computer software program with a regression analysis (Polar Precision Performance, Polar, Kempele, Finland) and visually by three independent observers identifying DP and HR$\text{dp}$. Coefficients of variation (CV) of HR$\text{dp}$ and PO at DP (PO$\text{dp}$) were calculated for the four tests of each subject. In addition, mean CVs of HR$\text{dp}$ and PO$\text{dp}$ of the four tests were calculated. To evaluate agreement between observers on the resulting HR$\text{dp}$, ICCs (intraclass correlation coefficients) were calculated. The ranges of HR$\text{dp}$ between the three observers were calculated of each CT in absolute values and expressed as a percentage of HR$\text{dp}$ of the corresponding CT. The level of significance was set at $p = 0.05$. All statistical analyses were carried out using SPSS, Version 14.0.

3. Results

The 17 subjects with 4 CTs (68 PO/HR graphs) were included for further analysis.

Table 1 shows the CVs for HR (CV$_{HR}$) of the 4 tests of the 17 subjects analysed by the computer program and the observers.

CVs of PO$\text{dp}$ (CV$_{PO}$) of the 4 tests were higher than the CVs for HR$\text{dp}$ (Table 2).

ICC for HR$\text{dp}$ analysed visually by the three observers was $0.828$ (95% Confidence Interval 0.756-0.883; $p < 0.001$) indicating a small variance in HR$\text{dp}$ rating and thus a high agreement between the three observers.

Mean range of HR$\text{dp}$ per CT between the three observers was 4.5 bpm corresponding 2.7% of HR$\text{dp}$.

4. Discussion

4.1 Reproducibility of HR at DP (HR$\text{dp}$)

The first objective of our study was to evaluate the reproducibility of HR$\text{dp}$, because this is an important characteristic of a good exercise test. Mean CV$_{HR}$ of 2.51% to 2.99% for the observers and 3.36% for the analyses by the computer program comply with a good reproducibility. According to Jeukendrup et al. [11], test reproducibility for HR is good, when CV is below 3.5%. Going into detail by looking at each subject’s CV$_{HR}$ we


### Table 1  Coefficient of variation (CV) of HR at deflection point [%].

<table>
<thead>
<tr>
<th>Subject</th>
<th>Computer program</th>
<th>Observer 1</th>
<th>Observer 2</th>
<th>Observer 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4.86</td>
<td>5.26</td>
<td>1.81</td>
<td>1.25</td>
</tr>
<tr>
<td>2</td>
<td>2.58</td>
<td>1.30</td>
<td>1.86</td>
<td>2.08</td>
</tr>
<tr>
<td>3</td>
<td>3.34</td>
<td>1.83</td>
<td>2.88</td>
<td>1.83</td>
</tr>
<tr>
<td>4</td>
<td>2.32</td>
<td>3.07</td>
<td>2.22</td>
<td>2.73</td>
</tr>
<tr>
<td>5</td>
<td>6.49</td>
<td>4.00</td>
<td>3.39</td>
<td>3.66</td>
</tr>
<tr>
<td>6</td>
<td>3.64</td>
<td>4.76</td>
<td>2.85</td>
<td>2.52</td>
</tr>
<tr>
<td>7</td>
<td>1.38</td>
<td>0.73</td>
<td>3.25</td>
<td>2.30</td>
</tr>
<tr>
<td>8</td>
<td>1.60</td>
<td>1.60</td>
<td>1.60</td>
<td>1.60</td>
</tr>
<tr>
<td>9</td>
<td>0.51</td>
<td>3.08</td>
<td>3.02</td>
<td>3.02</td>
</tr>
<tr>
<td>10</td>
<td>1.86</td>
<td>4.80</td>
<td>0.90</td>
<td>3.37</td>
</tr>
<tr>
<td>11</td>
<td>2.77</td>
<td>3.23</td>
<td>1.48</td>
<td>1.90</td>
</tr>
<tr>
<td>12</td>
<td>5.39</td>
<td>3.12</td>
<td>1.75</td>
<td>2.28</td>
</tr>
<tr>
<td>13</td>
<td>2.31</td>
<td>2.91</td>
<td>3.85</td>
<td>2.68</td>
</tr>
<tr>
<td>14</td>
<td>3.72</td>
<td>2.51</td>
<td>2.27</td>
<td>2.08</td>
</tr>
<tr>
<td>15</td>
<td>5.35</td>
<td>1.86</td>
<td>1.87</td>
<td>1.68</td>
</tr>
<tr>
<td>16</td>
<td>3.14</td>
<td>2.53</td>
<td>3.52</td>
<td>3.25</td>
</tr>
<tr>
<td>17</td>
<td>5.91</td>
<td>4.21</td>
<td>4.21</td>
<td>4.83</td>
</tr>
<tr>
<td>mean</td>
<td>3.36</td>
<td>2.99</td>
<td>2.51</td>
<td>2.53</td>
</tr>
</tbody>
</table>

### Table 2  Coefficient of variation (CV) of PO at deflection point [%].

<table>
<thead>
<tr>
<th></th>
<th>Computer program</th>
<th>Observer 1</th>
<th>Observer 2</th>
<th>Observer 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean CV</td>
<td>6.46</td>
<td>6.92</td>
<td>6.86</td>
<td>5.49</td>
</tr>
<tr>
<td>Range of CVs persubject</td>
<td>0.11.98</td>
<td>3.17-14.99</td>
<td>2.83-18.60</td>
<td>0-10.17</td>
</tr>
</tbody>
</table>

found CV$_{HR}$ above 3.5% in 6 of the 17 subjects (35%) with the computer analyses and in 8 of 51 analyses (15.6%) by the observers. Thus, the answer to the second question of our study is that the analysis by experienced observers seems to be more reliable than the analysis by the computer program. In addition, the ICC showed a good agreement between the three observers, which also supports the good quality of visual analysis by experienced observers.

In literature, only one study exists that assessed the reproducibility of the CTs with a protocol for running [10]. Ballarin et al. [10] analysed two consecutive CTs of 75 subjects of different age and training levels performed within one week. The analysed tests had to satisfy different criteria: (1) straight-line equation of the speed/HR relationship having an $r$ higher than 0.98 and; (2) increase of HR not more than 8 bpm each minute. The speed/HR graphs were analysed by 6 different observers and the computer program. The data of the DPs for speed and HR were compared by means of linear regression analysis. The results of the two tests performed by the same subject were practically identical for speed and HR at the DP and for the straight-line equation. Comparing the analysis by computer program with the visual analysis by the observers the authors concluded that visual analysis provides information that is very similar to that obtained through computer analysis, but the accuracy of the visually obtained results varies according to the observer’s experience. The DP could be identified visually in all tests by 4 of the 6 observers. It is concluded that the test is easily repeatable in running. Because only CTs meeting the above mentioned criteria were analysed, it is possible that PO/HR graphs that showed no DP were excluded from the study in advance.

An important finding of our study is the fact that HR$_{dp}$ had a better reproducibility than PO$_{dp}$. This is consistent with the finding of Sentija et al. [12], who found a very good reproducibility for HR$_{dp}$ comparing...
different kind of modified Conconi test protocols, while speed was dependent on the test protocol. Protocols with longer duration of the bouts resulted in a slower speed at DP, without changing HR_{dp} compared with protocols with shorter bouts. In our study, also an identical test protocol leads to better reproducibility for HR than for PO. We suggest that PO has a bigger day-to-day variability than HR. Therefore, in practice HR and not PO or speed should be used for training prescriptions.

4.2 Identification of a DP

In our study, 3 observers analysed 76 HR/PO graphs for the identification of HR_{dp}. In 10 out of the 228 analyses no DP could be found (4.4%). These 10 graphs were related to two subjects.

The existence of a DP in the PO/HR relationship was discussed controversially in the last 20 years [4, 7, 13-15]. In a review article about the concept of HR_{dp} Bodner and Rhodes [15] concluded that the degree of deflection is highly dependent upon the type of protocol used and HR_{dp} appears to be reliable when a DP in the PO/HR graph can be identified. Sentija et al. [12] showed that different kind of incremental protocol may be used for determination of identical HR_{dp}, while speed at DP is protocol dependent. Lepretre et al. [16] demonstrated that in well-trained subjects HR_{dp} coincided with the optimal cardiac work for which maximal stroke volume was attained.

Probst et al. [3] were able to identify DPs in their study in all of the 687 tests performed with the described protocol. In the study of Ballarin et al. [10] two of six observers could not identify 5 out of 65 test (7.6%) and 2 out of 65 tests (3.0%) respectively. The other four observers were able to identify a DP in all the PO/HR graphs. But in this study only tests with a straight-line equation of the speed/HR relationship having an r higher than 0.98 were accepted and analysed.

According to the existing literature, we can conclude that DP can be identified in the majority of the HR/PO graphs, but in few cases (0 to 7.6%) it may be impossible to see a deflection [3, 12, 14]. In our experience, the lack of a DP mostly appears in people that are not used to perform to exhaustion and stop the test as soon as acidaemia appears. In literature, different authors deny the concept of the DP, while others confirm the concept and could correlate the DP with the anaerobic threshold and HR_{dp} with maximal stroke volume [7-13].

4.3 Ranges of HR_{dp}

In addition to the CV we calculated the ranges of HR_{dp} of each CT analysed by the observers. From a practical point of view, the average range of 4.47 bpm or 2.66% of HR_{dp} seems to be sufficient to prescribe training guidelines.

4.4 Limitation

Two of the nineteen subjects (10.5%) had to be excluded from further analysis because at least one observer could not define a HR_{dp} in at least one of the PO/HR graphs of these subjects.

5. Conclusions

The study showed that the reproducibility of HR_{dp} from CTs was good. Additionally in 15% of the subjects the CV_{HR} analysed by observers did not reach CV of < 3.5%, what was defined to be a good reproducibility for HR. For these subjects other methods for exercise testing must be applied. The visual analysis by observers seems to be more accurate than the mathematical analysis by computer. There was a good accordance for the three observers. Using the results for training prescription purposes HR seems to be more accurate than PO.

References

Reproducibility and Variability of the Conconi Test Analysis


