Reliability and Validity of the Lactate Pro 2 Analyzer

Nora May Crotty, Marie Boland, Nick Mahony, Bernard Donne & Neil Fleming

To cite this article: Nora May Crotty, Marie Boland, Nick Mahony, Bernard Donne & Neil Fleming (2021): Reliability and Validity of the Lactate Pro 2 Analyzer, Measurement in Physical Education and Exercise Science, DOI: 10.1080/1091367X.2020.1865966

To link to this article: https://doi.org/10.1080/1091367X.2020.1865966

Published online: 25 Jan 2021.
Reliability and Validity of the Lactate Pro 2 Analyzer

Nora May Crotty, Marie Boland, Nick Mahony, Bernard Donne, and Neil Fleming

ABSTRACT
This study evaluated the reliability of the portable Lactate Pro 2 analyzer (LP2), and its validity compared to a laboratory-based analyzer, YSI 1500 Sport (YSI). Blood samples (n = 258) were collected during 44 graded incremental rowing tests, with data from 17 tests used to quantify load, heart rate, and oxygen consumption at lactate threshold. Inter-LP2 reliability was high with coefficient of variation = 3.3%; Bland Altman 95% limits of agreement varied from ±0.3 mmol.L⁻¹ for blood lactate (BLa) ≤4.0 mmol.L⁻¹ to ±1.6 to +1.4 mmol.L⁻¹ for BLa >8.0 mmol.L⁻¹, with minimal systematic bias. Despite LP2 measurements being ~1.42 times higher than YSI measurements, no significant differences (P > .05) were detected when quantifying load, heart rate, or oxygen consumption data at lactate threshold. The results support using different LP2 devices interchangeably, especially when BLa ≤ 8 mmol.L⁻¹, and using the LP2 to identify training intensities at lactate threshold.

KEYWORDS
Portable lactate analyzers; lactate threshold; Bland-Altman limits of agreement; Dmax; YSI 1500 Sport

Introduction
Blood lactate (BLa) accumulation during exercise is well established as a key physiological marker in endurance athletes. It is traditionally measured during incremental exercise tests alongside other physiological variables including oxygen consumption (VO₂), heart rate (HR) and exercise workload or power. During these tests BLa measurements are used to evaluate an athlete’s endurance capacity, monitor performance, and to generate exercise prescriptions (Faude et al., 2009). The lactate threshold, also known as anaerobic threshold (Ghosh, 2004), is traditionally defined as the point beyond which there is an abrupt increase in BLa (Goodwin et al., 2007) and is of particular importance for exercise prescription (Bentley et al., 2007; Ghosh, 2004; Goodwin et al., 2007). Within elite endurance sport, an athlete’s velocity or power at lactate threshold is considered a more accurate performance indicator than their maximal oxygen uptake (Bentley et al., 2007). It is therefore of utmost importance that BLa is accurately measured and the associated lactate threshold is reliably identified by practitioners.

Traditionally, BLa has been measured by laboratory-based bench-top analyzers. In recent years, hand-held battery-operated lactate analyzers have gained increasing popularity. These hand-held devices offer a number of advantages to the coach and athlete over the traditional bench-top analyzers, especially the ability to monitor BLa in the field; they also cost less to purchase and require smaller blood samples for analysis (Buckley et al., 2003; Tanner et al., 2010). In order to use these portable analyzers for physiological assessment of athletes; however, it is important to establish their reliability (consistency of measurements) and construct validity (accuracy of measurement), especially when it comes to the identification of the lactate threshold.

The Lactate Pro 2 (LP2: Arkray, Kyoto, Japan) is one of the newer and most popular of the portable analyzers, requiring a smaller blood sample and shorter processing time than its predecessor, the Lactate Pro (LP: Arkray, Kyoto, Japan) (Arkray, 2015; Rowe & Whyte, 2016). While several studies (Baldari et al., 2009; Bonaventura et al., 2015; Buckley et al., 2003; Medbo et al., 2000; McNaughton et al., 2002; Pyne et al., 2000; Tanner et al., 2010; Van Someren et al., 2005) have evaluated the performance of the LP device, only one published study to date (Bonaventura et al., 2015) has examined the reliability of the LP2 and/or assessed its construct validity against a laboratory-based analyzer. In this study the LP2 showed the highest inter-device reliability, as assessed by the coefficient of variation (CV), of all the six portable analyzers tested (Bonaventura et al., 2015). In addition, the LP2 underscored BLa measurements by the criterion bench-top analyzer, a Radiometer ABL90 (Radiometer, Copenhagen, Denmark), throughout most of the BLa range. Previous research comparing the LP to a criterion analyzer also reported a difference in absolute BLa data between analyzers, with the magnitude of the difference often increasing with increasing BLa, implying...
heteroscedastic data (Buckley et al., 2003; Medbo et al., 2000; van Someren et al., 2005). Given the importance of the lactate threshold in exercise monitoring and prescription, some investigators have examined whether this discrepancy in absolute BLA data between the LP and a criterion analyzer significantly affects the quantification of heart rate and power at various lactate parameters (Buckley et al., 2003; van Someren et al., 2005). The results have been mixed (Buckley et al., 2003; van Someren et al., 2005). A prospective evaluation of lactate threshold-derived data from the LP2 has yet to be conducted.

One limitation of using CV to assess reliability is that it describes variation for only 68% of the test data, leaving the other 32% unaccounted for (Atkinson & Nevill, 1998). For this reason, Bland Altman (BA) 95% limits of agreement (LoA) analysis, which assesses the agreement between 95% of the test data, is a more comprehensive method for assessing absolute reliability (Atkinson & Nevill, 1998).

The primary aim of the current study, therefore, was to evaluate the inter-device reliability of the LP2 device during graded incremental testing in a large cohort of competitive rowers using BA 95% LoA. A second aim was to assess the construct validity of the LP2 device by comparing BLA data recorded with this device to those of a calibrated laboratory-based analyzer, the YSI 1500 Sport (YSI: Yellow Springs Instruments, OH, USA). The final aim was to evaluate the accuracy of LP2 device in quantifying power, heart rate (HR) and oxygen consumption ([\(\dot{V}O_2\)], at lactate threshold by comparing to the same quantification made using concurrent data from the YSI.

**Methods**

**Study design**

This cohort study was performed in accordance with the principles of the Declaration of Helsinki. Ethics approval was obtained from the Faculty of Health Sciences Research Ethics Committee at Trinity College Dublin. Data were collected in conjunction with another research project comparing physiological response to graded exercise tests (GXTs) using two commercially available rowing ergometers. Participants were recruited from the university men’s senior rowing squad and local club oarsmen. The study involved three separate visits to the Human Performance Laboratory at Trinity College Dublin. During visit 1, participants provided informed consent and were screened for contraindications to maximal exercise testing. During visits 2 and 3, each participant underwent a GXT protocol; one on the RowPerfect 3 ergometer (RP3: Care RowPerfect, Hardenberg, The Netherlands) and the other on the Concept II ergometer (C2: Concept II, VT, USA). The order of the tests was randomized.

**Test protocol**

The starting workload for each GXT was 120 W, with each stage lasting 3-min. Workload was increased by 30 W for each subsequent stage until the test was terminated. A 1-min rest period was provided between each stage, in order to facilitate blood sampling. Test termination criteria were: 1) volitional exhaustion; 2) inability to maintain target workload; or 3) a plateau in HR and [\(\dot{V}O_2\)]. Inter-test duration was 48 h and participants were advised to maintain similar sleep and eating patterns, and to avoid excessive exertion in the 24-h prior to each test.

**Data collection**

The portable analyzer used in the current study was the LP2 which analyzes a 0.3 \(\mu\)L blood sample in 15-s (Arkray, 2015; Shimojo et al., 1993). The YSI acted as the criterion analyzer for comparison with the LP2 devices. It processes a 25 \(\mu\)L blood sample in 60-s (YSI Life Sciences, 2003). Both analyzers processed non-hemolyzed blood samples, and the YSI was calibrated to a resolution of ± 0.05 mmol L\(^{-1}\) before each test.

Heart rate (HR) during the GXT was recorded continuously by radio-telemetry using a Cosmed HR monitor (Cosmed, Rome, Italy). Gas exchange variables were measured on a breath-by-breath basis using a pretest calibrated cardiopulmonary exercise testing unit (CPET) and an associated software package (Cosmed, Rome, Italy).

Blood samples were collected from the ear lobe of each participant during the rest intervals of the GXT. Prior to collection, the site was cleaned to remove any sweat or blood and then lanced with a sterile single-use lancet (Solofix, Braun, Melsungen, Germany). Approximately 40 \(\mu\)L of blood was withdrawn via capillary action into a heparinized capillary tube (Brand, Wertheim, Germany). A 25 \(\mu\)L sample was injected into the mixing chamber of the YSI and the residual blood was displaced onto a sterile glass slide. Test strips inserted into two LP2 devices (LP2a and LP2b) were then sequentially placed against the surface of this blood sample until beeps were audible, signifying the start of analysis (Arkray, 2015). The same operator collected and analyzed all blood samples throughout the study.

All blood samples collected throughout 17 of the 44 tests were analyzed by the YSI and the two LP2 devices. In the remaining 27 tests, each blood sample was analyzed by the YSI, but was only analyzed by the two LP2 devices when BLA measured by the YSI exceeded 2 mmol L\(^{-1}\). The reason for restricting analysis by the LP2 devices in these 27 tests was to limit the number of data points collected for BLA <4.0 mmol L\(^{-1}\), and, therefore, obtain a comparable
number of data points throughout the assessed BLa ranges used for the BA plots. This resulted in a total of n = 258 blood samples being measured by all three analyzers for inclusion in the final data analysis.

**Statistical analysis**

Statistical analysis was performed for three ranges of BLa measured using the LP2: ≤4.0 mmol.L⁻¹, 4.1 to 8.0 mmol.L⁻¹, and >8.0 mmol.L⁻¹ for both reliability and construct validity assessments. These three ranges were chosen as it was anticipated from previous literature that the variability between the measurements would increase with increasing magnitude of BLa and similar ranges have been used in previous research (Tanner et al., 2010). Measures of relative (Pearson’s correlation coefficient (r), coefficient of determination (r²), intraclass correlation coefficient (ICC)) and absolute (CV, relative technical error of measurement (r-TEM), 95% LoA) reliability were calculated using SPSS software (IBM, Ver 25.0, NY, USA) for repeated pairs of measurements. The ICC (3,1) was used in the current study according to the guidelines of Shrout and Fleiss (1979) and as described by Koo and Li (2016). Scedasticity of data were assessed using Levene’s test, with P < .05 inferring heteroscedastic data. When the data exhibited homoscedasticity, BA plots were used to assess for systematic bias and quantify the 95% LoA (Bland & Altman, 1999) but not if the data exhibited heteroscedasticity. To examine inter-device reliability BLa measurements for LP2a were compared to those recorded by the LP2b. To examine construct validity, measurements of BLa for YSI were compared to those of the LP2a alone, since a paired Student’s T-test detected no significant difference in BLa measurements (n = 258) between the two LP2 devices.

Finally, in order to evaluate the practical implications of variance in BLa data as measured on LP2, lactate threshold was calculated via the Dmax method using customized programs in Matlab (MathWorks, MA, USA). This calculation was only performed for the 17 tests where BLa data were measured on LP2 and YSI throughout the GXT. The Dmax method involves plotting a 3rd order polynomial curve of BLa data against workload and identifying the maximal perpendicular distance from a line drawn between the first and last data points and this curve (Cheng et al., 1992). The associated HR, (VO₂̇, BLa and workload at Dmax were then extrapolated (Figure 1). Paired Student’s T-tests detected no significant difference in the quantification of these variables at lactate threshold between the two LP2 devices (see Table 3). Therefore, subsequent paired Student T-tests, conducted to identify any significant differences between the YSI and LP2 in these lactate threshold measurements, compared the YSI to the LP2a alone. In addition, relative and absolute reliability, systematic bias, and scedasticity were assessed as described previously. For the purpose of comparing lactate threshold data between LP2 and YSI, acceptable 95% LoA were set as ±20 W for load (Mc Grath et al., 2019), limits equivalent to ±5 beats. min⁻¹ for HR and ±4 mL.kg⁻¹.min⁻¹ for (VO₂̇, at Dmax. Two of the authors, who, between them have over 45 years of experience performing exercise testing and subsequent prescription, considered these appropriate for practical use; namely, to discern biological change while mitigating against the risk of erroneous or misleading results.

**Results**

**Participant characteristics**

Twenty-two (n = 22) healthy competitive male rowers participated in the current study. A total of n = 258 blood samples collected during 44 GXTs were analyzed. Participant age ranged from 17 to 28 years, rowing experience 1 to 14 years, and all were familiar with using RP3 and C2 ergometers. Table 1 presents their anthropometric and maximal physiological data.

**Inter-device reliability**

Statistical evaluation of LP2 inter-device reliability within three ranges of BLa (≤4.0 mmol.L⁻¹, 4.1 to 8.0 mmol.L⁻¹, and >8.0 mmol.L⁻¹) are presented in Table 2 and Figure 2. Variance in data was deemed homoscedastic within each BLa range and across the total data set. Scatter plots showed high measures of relative reliability (ICC = 0.94 to 0.99; r = 0.95 to 0.99; r² = 0.91 to 0.98); BA plots showed minimal systematic bias (< 0.1 mmol.L⁻¹) and varying LoA across the three ranges, see Table 2. CV was low and stable across all ranges (3.0 to 3.6%) as was r-TEM (4.4 to 5.3%).

**Construct validity**

Statistical evaluation of LP2 construct validity within three ranges of BLa (≤4.0 mmol.L⁻¹, 4.1 to 8.0 mmol.L⁻¹, and >8.0 mmol.L⁻¹) are presented in Table 2 and Figure 3. Variance in data was deemed heteroscedastic within each BLa range and across the total data set. The LP2 device consistently over-estimated YSI assessed BLa data (Figure 3). Scatter plots showed acceptable levels of relative reliability at lower BLa ranges (ICC = 0.87); however, much poorer measures were observed in the middle and upper ranges (ICC = 0.30 at 4.1 to 8.0 mmol.L⁻¹, and ICC = 0.15 at >8.0 mmol.L⁻¹). Regression lines on the BA plots highlight the consistent bias in residuals (Figure 3). CV was very high (18.4 to 29.0%) as was r-TEM (20.6 to 28.1%), with both measures of absolute reliability increasing substantially in magnitude at the higher BLa range. Linear regression
Figure 1. Exemplary graph of Dmax calculated from YSI 1500 Sport (YSI) and Lactate Pro 2a (LP2a) data. The corresponding load and heart rate at Dmax are interpolated for YSI (dashed lines) and LP2a (dotted lines). The same procedure was used to identify the corresponding volume of oxygen consumed (\(\dot{V}O_2\)) at Dmax by replacing heart rate data with \(\dot{V}O_2\) data.

Table 1. Mean ± SD anthropometric data and physiological response to maximal exercise in the participant cohort. BMI = body mass index, Pmax = maximum power, HR max = maximum heart rate, \(\dot{V}O_2\)max = maximum volume of oxygen consumed.

<table>
<thead>
<tr>
<th>Outcome Variable</th>
<th>Measure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (y)</td>
<td>20 ± 2</td>
</tr>
<tr>
<td>Height (m)</td>
<td>1.88 ± 0.06</td>
</tr>
<tr>
<td>Body Mass (kg)</td>
<td>82.4 ± 8.7</td>
</tr>
<tr>
<td>BMI (kg.m(^{-2}))</td>
<td>23.2 ± 1.6</td>
</tr>
<tr>
<td>Body Fat (%)</td>
<td>12.8 ± 3.1</td>
</tr>
<tr>
<td>Pmax (W)</td>
<td>352 ± 27</td>
</tr>
<tr>
<td>HRmax (beats.min(^{-1}))</td>
<td>196 ± 7</td>
</tr>
<tr>
<td>(\dot{V}O_2)max (mL.kg(^{-1}).min(^{-1}))</td>
<td>57.7 ± 6.2</td>
</tr>
</tbody>
</table>

Table 2. Measures of inter-device reliability (LP2a vs. LP2b) and validity (LP2a vs. YSI) across discrete blood lactate ranges.

<table>
<thead>
<tr>
<th>Lactate concentration range (mmol.L(^{-1}))</th>
<th>≤ 4.0</th>
<th>4.1 – 8.0</th>
<th>&gt; 8.0</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reliability ICC</td>
<td>0.99</td>
<td>0.94</td>
<td>0.97</td>
<td>0.99</td>
</tr>
<tr>
<td>r-TEM (%)</td>
<td>5.3</td>
<td>5.0</td>
<td>4.4</td>
<td>5.4</td>
</tr>
<tr>
<td>CV (%)</td>
<td>3.4</td>
<td>3.6</td>
<td>3.0</td>
<td>3.3</td>
</tr>
<tr>
<td>Mean bias (mmol.L(^{-1}))</td>
<td>0.0</td>
<td>0.0</td>
<td>0.1</td>
<td>0.0</td>
</tr>
<tr>
<td>95% LoA (mmol.L(^{-1}))</td>
<td>–0.3 to 0.3</td>
<td>–0.7 to 0.7</td>
<td>–1.6 to 1.4</td>
<td>–1.0 to 1.0</td>
</tr>
<tr>
<td>Validity ICC</td>
<td>0.87</td>
<td>0.30</td>
<td>0.15</td>
<td>0.85</td>
</tr>
<tr>
<td>r-TEM (%)</td>
<td>20.8</td>
<td>20.6</td>
<td>28.1</td>
<td>33.9</td>
</tr>
<tr>
<td>CV (%)</td>
<td>18.7</td>
<td>18.4</td>
<td>24.9</td>
<td>19.8</td>
</tr>
</tbody>
</table>

LP2a = Lactate Pro 2a, LP2b = Lactate Pro 2b, YSI = YSI 1500 Sport, ICC = intraclass correlation coefficient (3,1), r-TEM = relative technical error of measurement, CV = coefficient of variation, 95% LoA = Bland Altman 95% limits of agreement.

Table 3. Group mean ± SD for load, heart rate, \(\dot{V}O_2\), and BLa at Dmax measured with the Lactate Pro 2a (LP2a), Lactate Pro 2b (LP2b) and YSI 1500 Sport (YSI) analyzers.

<table>
<thead>
<tr>
<th></th>
<th>YSI</th>
<th>LP2a</th>
<th>LP2b</th>
</tr>
</thead>
<tbody>
<tr>
<td>Load (W)</td>
<td>257 ± 19</td>
<td>259 ± 25</td>
<td>260 ± 23</td>
</tr>
<tr>
<td>Heart rate (beats.min(^{-1}))</td>
<td>175 ± 9</td>
<td>176 ± 8</td>
<td>176 ± 8</td>
</tr>
<tr>
<td>(\dot{V}O_2) (mL.kg(^{-1}).min(^{-1}))</td>
<td>49 ± 5</td>
<td>49 ± 5</td>
<td>49 ± 5</td>
</tr>
<tr>
<td>BLa (mmol.L(^{-1}))</td>
<td>2.6 ± 0.7***</td>
<td>3.3 ± 0.5</td>
<td>3.3 ± 0.7</td>
</tr>
</tbody>
</table>

\(\dot{V}O_2\) = volume of oxygen consumed, BLa = blood lactate concentration.

Asterisk symbol (*) infers significantly higher BLa data at Dmax comparing LP2a and YSI, *** infers \( P < 0.001 \).

(y = 1.433x – 0.278) predicted 95% of the variance in data (Figure 4).

Identification of lactate threshold

Measures of load, HR, \(\dot{V}O_2\) and BLa at lactate threshold (Dmax) interpolated using LP2 and YSI data are presented in Table 3. Paired Student’s T-tests detected no significant differences \( (P > .05) \) between YSI and LP2 for the interpolation of the load, heart rate or \(\dot{V}O_2\) at Dmax; however, BLa at Dmax was significantly higher using LP2 data \((P < .001)\). Reliability measures of Dmax variables are presented in Table 4. These indicated high relative reliability and acceptable absolute reliability between analyzers for interpolation of load, HR and \(\dot{V}O_2\) at Dmax, indicating...
strong agreement in the determination of lactate threshold for these variables. However, relative reliability of BLa data at Dmax was very low (ICC = 0.33).

Discussion

Portable lactate analyzers offer a number of advantages over laboratory-based devices; however, their reliability and construct validity needs to be established before they can be depended upon to reliably assess athletic performance. We, therefore, set out to investigate both the inter-device reliability of a newer portable analyzer, the LP2, and its construct validity when compared to a calibrated laboratory-based analyzer, the YSI. The study produced three main findings: (1) the inter-device reliability of the LP2 analyzer was high, especially

Figure 2. Scatter plots of LP2a and LP2b data in (a) ≤ 4.0 mmol.L\(^{-1}\), (c) from 4.1 to 8.0 mmol.L\(^{-1}\) and (e) > 8.0 mmol.L\(^{-1}\) ranges. Regression lines (solid black) and lines of equivalence (dashed red) are presented along with respective \(r\), \(r^2\) and ICC data. Bland-Altman plots of LP2a and LP2b data in (b) ≤ 4.0 mmol.L\(^{-1}\), (d) from 4.1 to 8.0 mmol.L\(^{-1}\) and (f) > 8.0 mmol.L\(^{-1}\) ranges. Linear regression lines (solid black), mean bias (dashed blue), upper and lower 95% LoA (dotted red) are presented along with their respective values. LP2a = Lactate Pro 2a; LP2b = Lactate Pro 2b, BLa = blood lactate, LoA= limits of agreement.
in the range of (LP2) BLa ≤8 mmol.L⁻¹; (2) BLa measurements were consistently higher compared to corresponding measurements recorded by the YSI (by a factor of ~1.4) and (3) the discrepancy between absolute LP2 and YSI measurements had little effect on the quantification of physiological variables at the lactate threshold using the Dmax method.

**Inter-device reliability of LP2 analyzer**

All assessed relative reliability indices, and the absolute reliability indices of CV and r-TEM indicated high interdevice reliability of the LP2 device throughout the range of BLa measurements. However, when we assessed absolute reliability using 95% LoA, the results were less clear-cut. While there is no precedent for what constitutes
a clinically relevant change in blood lactate with training, we detected what we considered acceptable 95% LoA for BLa ≤ 4 mmol.L\(^{-1}\) (−0.3 to +0.3 mmol.L\(^{-1}\)), and BLa 4.1 to 8 mmol.L\(^{-1}\) (−0.7 to +0.7 mmol.L\(^{-1}\)), and a negligible (0.0 mmol.L\(^{-1}\)) mean bias for both of these ranges. From an exercise testing standpoint, the lactate thresholds used for exercise prescription and performance monitoring occur when BLa data is ≤8 mmol. L\(^{-1}\) (Faude et al., 2009; Goodwin et al., 2007). These results, and the high reliability indicated by the other measures, would therefore suggest that different devices of the same LP2 analyzer may be used interchangeably for the majority of testing purposes. For BLa > 8 mmol.L\(^{-1}\), while indices of relative reliability remained good, the 95% LoA increased to −1.6 to +1.4 mmol.L\(^{-1}\). Despite these wider limits, given the exponential nature of lactate accumulation and clearance kinetics, and the minimal (0.1 mmol.L\(^{-1}\)) systematic bias observed, it may also be reasonable to assume the use of LP2 devices interchangeably in this range to monitor test response and/or post-exercise recovery. It is also noteworthy that the two LP2 devices did not differ significantly when used for the calculation of load, HR or (\(\dot{V}O_2\)) at Dmax, suggesting the wide LoA detected above 8 mmol.L\(^{-1}\) did not impact upon the identification of lactate threshold.

We observed an overall CV of 3.3% (range 3.0 to 3.4%) for inter-device reliability of the LP2 in the current study. This was considerably higher, for reasons that are unclear, than the CV of 0.0 to 1.0% reported in the only other study to date that has examined the inter-device reliability of the LP2 (Bonaventura et al., 2015). However, the CV calculated in that study was based on repeated measures using the same 22 blood samples which might not be directly comparable to the 258 samples used in the current study. Of note, a CV of 3.3% was better than that calculated for the LP in previous research (CV = 3.6 to 5.7%) (Bonaventura et al., 2015; Mc Naughton et al., 2002;
Tanner et al., 2010), suggesting that the LP2 is more reliable than its predecessor, the LP.

**Construct validity of the LP2 analyzer**

LP2 BLa measurements were overall considerably (~1.42 times) higher than those of the YSI 1500 criterion analyzer, with the magnitude of the difference increasing with increasing BLa concentrations. This differed from previous research in which the LP2 measurements underscored those of the criterion laboratory-based analyzer, Radiometer ABL 90 (Bonaventura et al., 2015) and from another study where LP measurements underscored those of a YSI 2300 analyzer (Pyne et al., 2000). The results were similar, however, to studies comparing the LP to the YSI 1500, where the variance data were deemed heteroscedastic, and LP measurements were ~1.5 times higher than the corresponding YSI measurements (Buckley et al., 2003; Medbo et al., 2000). These discrepancies in BLa measurements between the handheld device and the criterion analyzer have been attributed to differences in the type of sample analyzed (Buckley et al., 2003; Medbo et al., 2000). In the case of YSI analyzers, unless the blood sample is hemolyzed prior to analysis, it measures the concentration of plasma lactate in the blood sample (Goodwin et al., 2007; YSI Life Sciences, 2003). In contrast, the Radiometer ABL 90, the LP, and the LP2, all measure whole blood lactate (Arkay, 2015; Bonaventura et al., 2015; Goodwin et al., 2007). In the case of the previous investigation using the YSI 2300 (Pyne et al., 2000) the blood samples were hemolyzed using a lysis solution prior to analysis, resulting in the release of lactate from red blood cells and therefore higher BLa measurements than would have been obtained with non-lysed samples.

**Identification of lactate threshold**

Despite the large difference in absolute BLa measurements between the YSI and LP2 analyzers, there were no statistically significant differences (P > .05) detected comparing analyzers in the quantification of load, HR, or (\(\dot{V}O_2\)) at lactate threshold using the Dmax method. The 95% LoA were very good and excellent for HR (\(\pm 4 \text{ beats.min}^{-1}\)) and (\(\dot{V}O_2\)) (\(\pm 2 \text{ mL.kg}^{-1}.\text{min}^{-1}\)), respectively, at Dmax, although only equivocal for load at Dmax (\(\pm 19 \text{ W}\)). Nonetheless, the LoA for load, HR and (\(\dot{V}O_2\)) at Dmax were within *a-priori* acceptable ranges suggesting that the analyzers can be used interchangeably for the calculation of these Dmax variables; however, caution would be advised when using them interchangeably for load at Dmax. To this author’s knowledge, no previous publication has assessed the calculation of lactate threshold using direct BLa measurements from the LP2. Several authors, using retrospective data, have modeled the estimated lactate thresholds via linear regression equations (Bonaventura et al., 2015; Rowe & Whyte, 2016) and observed strong agreement to criterion analyzers. Our calculation of Dmax variables, based on a direct comparison of LP2 and YSI BLa measurements, were consistent with these findings. Based on the small differences detected in mean data and acceptable 95% LoA for computed physiological variables at Dmax, we conclude that using the LP2 for the assessment of lactate threshold will not differ from a calibrated laboratory-based analyzer.

The finding in the current study of poor agreement between the LP2 and YSI analyzers for absolute BLa data, yet similar data for load, HR and (\(\dot{V}O_2\)) at lactate threshold calculated using the Dmax method, concurs strongly with previous research (Buckley et al., 2003). In this study of GXT data on kayak and rowing ergometers, the measurements of load and HR at Dmax were slightly lower than in the current study. However, the mean difference in load at Dmax between the two analyzers was similar at ~1.4 W compared to ~2 W in the current study; likewise the difference in HR at Dmax was ~0 beats.min\(^{-1}\) compared to ~1 beats.min\(^{-1}\). Not surprisingly, a significant difference between the laboratory criterion and portable analyzers was detected when comparing lactate derived indices at fixed blood lactate concentrations (Buckley et al., 2003; van Someren et al., 2005).

**Study limitations**

Before interpreting the current results, several study limitations must be considered. Firstly, this study was performed in a cohort of competitive male rowers. Whether the results are applicable to the wider population of endurance athletes remains to be clarified. Secondly, the criterion analyzer used in this study was a YSI 1500 Sport. While this analyzer has been widely used in exercise laboratories for many years (Buckley et al., 2003), it is no longer commercially available. However, because the critical hardware components, calibration technique, and measurement technology have remained unchanged in newer generation YSI analyzers, it seems reasonable to conclude that the study results would have been similar if we had used a newer model of the YSI to assess construct validity of the LP2. In addition, while one study has questioned the reliability of the YSI 1500 (Medbo et al., 2000), the large discrepancy in measurements reported between the two YSI 1500 analyzers in that study suggests reexamination of the data may be warranted. Furthermore, other authors have concluded that the
YSI was very reliable (Buckley et al., 2003). Finally, it has been suggested by Bonaventura et al. (2015) that to accurately evaluate inter-device reliability, up to six devices should be tested against one another. While this approach may be optimal, most previous research examining the reliability of hand-held lactate analyzers have used only two (Bonaventura et al., 2015; Pyne et al., 2000; Tanner et al., 2010), or three devices (Baldari et al., 2009). Due to practical constraints, we were limited to testing two devices in the current study.

**Conclusion**

This study comprehensively assessed the inter-device reliability of the LP2, and observed that the device was highly reliable when analyzing BLa ≤8 mmol.L⁻¹ but was less reliable at higher BLa concentrations. Furthermore, despite consistently higher measures of absolute BLa when compared to the criterion YSI, the calculation of load, HR and (𝑉𝑉̇O₂) at lactate threshold (using Dmax protocol) did not differ between the analyzers. To the best of our knowledge, this is the first study to directly assess quantification of physiological variables at lactate threshold using the LP2 device. Our results support using different LP2 devices interchangeably for exercise monitoring and prescription, especially at lactate threshold and sub-threshold training intensities, and using the LP2 to identify training intensities at lactate threshold.

**Acknowledgments**

The authors would like to thank all the rowers who participated in the study and Mr. Richard Ruggieri, the coach of the men’s senior rowing team at Trinity College Dublin.

**ORCID**

Nora May Crotty http://orcid.org/0000-0003-3344-9878

**Conflicts of interest**

The authors declared no conflicts of interest.

**References**


