MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

„Comparison of Two Metabolic Gas Analyzers, and the Validity of Near Infrared Spectroscopy as a Method to Determine Anaerobic Threshold and Respiratory Compensation Point"

verfasst von / submitted by

Keltey Kathleen McGirr, Bakk.rer.nat.

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Master of Science (MSc)

Wien, 2019 / Vienna 2019

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	A 066 826
Studienrichtung lt. Studienblatt / degree programme as it appears on the student record sheet:	Masterstudium Sportwissenschaft
Betreut von / Supervisor:	Univ.-Prof. Mag. Dr. Harald Tschan

Foreword

"Science is not purely an intellectual thing: Like any human enterprise it depends upon the human factors of goodwill, and comprehension, on help frankly and freely given and received, on sympathy and fellowship in the needs and projects of others, on the satisfaction and delight in their discoveries and success."

-Archibald Vivian Hill
(The ethical dilemma of science, p. 9)

These words, written by the scientist who discovered maximal oxygen uptake, resonated with me as I thought back on my studies, mainly in the course of composing this thesis. Without the goodwill, comprehension, help, sympathy, and fellowship of a number of people, it would have been impossible to complete it.

I am thankful for my dad, who was always encouraging, helpful, understanding and open for a good discussion on various matters concerning my studies. My brothers Coren, Darren, and Benen were supportive during the intense times of my studies, and were also often able to pull me away from my work to give me a much needed break. I want to thank my grandparents for their support throughout my studies, and the many encouraging conversations.

I want to give a special thanks to Christoph Triska, who has helped me through the whole process of this thesis, from teaching me how to use the devices, to helping with the statistical analyses, and putting in the time to correct my writings, as well as always taking time to answer any questions. Further thanks goes out to my supervisor Harald Tschan, who has always been so supportive. I also want to thank Barbara Wessner and Benedikt Mitter for their help with the data analyses, as well as Gabi Maier and Wolfgang Reith for their support during the testing period.

Last but not least, I want to thank my friends who have made my time as a student unforgettably amazing.

Abstract

The aim of this thesis was to compare two commercially available metabolic gas analyzers, namely the MetaLyzer 3B (Cortex Biophysik GmbH., Leipzig, Germany) (ML 3B), and the MasterScreen CPX (Viasys Healthcare, Höchberg, Germany) (MS CPX), as well as validate near infrared spectroscopy (NIRS) as a method to determine the anaerobic threshold (AT) and the respiratory compensation point (RCP).

20 healthy recreationally active sport students (m = 14, w = 6) performed two graded exercise tests (GXT) on the treadmill in an alternate order with either the ML 3B or the MS CPX, with the NIRS applied to the gastrocnemius lateralis muscle during the GXT with ML 3B. NIRS derived AT (NdAT) was determined in the oxygenation index (OI = $\Delta[O_2Hb] - \Delta[HHb]$) and NIRS derived RCP (NdRCP) was determined in the concentration change of deoxyhemoglobin ($\Delta[HHb]$).

Mean maximal running speed was not significantly different between the two GXTs (P = 0.919). Mean $\dot{V}O_{2peak}$ of ML 3B was measured significantly higher than MS CPX (3,482 ± 904 mL·min^{-1} versus 3,246 ± 789 mL·min^{-1}, respectively; P < 0.001), but was significantly correlated (r = 0.982; P < 0.05). NdAT is significantly different than AT (P = 0.014) with a CoV of 24.6% and NdRCP is not significantly different than RCP (P = 0.789) with a CoV of 16.6%.

$\dot{V}O_2$ of ML 3B and MS CPX can be made comparable using an equation that demonstrates a small error of the estimate: $\dot{V}O_{2\ MS\ CPX} = (\dot{V}O_{2\ ML\ 3B} - 6.009) / 1.077$. According to this study NIRS cannot be used to determine AT and RCP validly.

Zusammenfassung

Das Ziel dieser Arbeit war es die Spiroergometriesysteme MetaLyzer 3B (Cortex Biophysik GmbH. Leipzig, Germany) (ML 3B), und MasterScreen CPX (Viasys Healthcare, Höchberg, Germany) (MS CPX) zu vergleichen und die Nahinfrarotspektroskopie (NIRS) als Möglichkeit die anaerobe Schwelle (AT) und den respiratorischen Kompensationspunkt (RCP) zu bestimmen, zu validieren.

20 gesunde Sportstudenten (m = 14, w = 6) absolvierten zwei Stufenleistungstests (GXT) am Laufband mit dem ML 3B oder dem MS CPX, mit dem NIRS auf dem m. gastrocnemius lateralis während des GXT mit dem ML 3B. Die AT wurde mit dem NIRS mittels Oxygenierungsindex (OI; $\Delta[O_2Hb] - \Delta[HHb]$) (NdAT) und der RCP wurde über den Konzentrationsunterschied im Desoxyhämoglobin bestimmt ($\Delta[HHb]$) (NdRCP).

Es gab keinen signifikanten Unterschied der durchschnittlichen maximalen Laufgeschwindigkeit der beiden GXTs (P = 0.919). Die durchschnittliche $\dot{V}O_{2max}$ von ML 3B ist signifikant höher als die der MS CPX (3,482 ± 904 mL·min^{-1} bzw. 3,246 ± 789 mL·min^{-1}; P < 0.001), jedoch ergab sich ein signifikanter Zusammenhang. (r = 0.982; P < 0.05).

Es gibt einen signifikanten Unterschied zwischen NdAT und AT (P = 0.014) mit einem CoV von 24.6%. Es gibt keinen signifikanten Unterschied zwischen NdRCP und RCP (P = 0.789), aber eine CoV von 16.6%.

$\dot{V}O_2$ der ML 3B und des MS CPX können durch eine Formel mit einem kleinen Standardfehler vergleichbar gemacht werden: $\dot{V}O_{2\,MS\,CPX} = (\dot{V}O_{2\,ML\,3B} - 6.009) / 1.077$. Laut dieser Studie ist NIRS keine valide Methode die AT und den RCP zu bestimmen.

Table of contents

1. Introduction

Over the last decades, equipment for performance assessment has drastically improved. This has been of major value for athletes and their trainers and has taken training control to a new level, allowing deeper insight than ever before into the metabolic processes in an athlete's body. With this increase in possibilities and the production of equipment by many different manufacturers, comes the challenge of assessing performance accurately within and across various methods (Hollmann, Strüder, Predel, & Tagarakis, 2006).

As performance assessment is a requirement for optimal training control, it is necessary to evaluate and athlete's capabilities precisely. For this reason, performance assessment demands valid and reliable methods (Hohmann, Lames, & Letzelter, 2007; Hottenrott & Neumann, 2010; Weineck, 2010). Validity is described as the ability to produce same results using either two different methods of measurement or applying the same method with different equipment (Hohmann et al., 2007). In addition, it is advisable to practice procedures which are economical and inflict the least amount of discomfort as possible on the athlete (Weineck, 2010).

There are various methods to determine an athlete's individual endurance capacity, including the determination of lactate thresholds (e.g. Brooks, 1985), ventilatory thresholds (e.g. Beaver, Wasserman, & Whipp, 1986), near infrared spectroscopy (NIRS) derived thresholds (e.g. Bellotti, Calabria, Capelli, & Pogliaghi, 2013; Van Der Zwaard et al., 2016) or critical power (e.g. Poole, Burnley, Vanhatalo, Rossiter, & Jones, 2017).

The method of best practice for measuring metabolic gas exchange goes back to Douglas in 1911 who created a portable metabolic gas analyzer. When using this apparatus the subject breathes into a gas bag, and the air in this bag is analyzed (Douglas, 1911). With some alterations and perfections, this method is today called the Douglas bag method and is considered the most precise way to measure metabolic gas expenditure and is therefore regarded as the gold standard (Hollmann et al., 2006). Derived from this original method of measuring metabolic gas exchange, various automated metabolic gas analyzers have been created and are currently used in performance assessment for various groups of people (e.g., elite athletes, recreational athletes, people with disease) (Macfarlane, 2001).

In previous studies, metabolic gas analyzers have shown to be valid and reliable (e.g., Foss & Hallén, 2005; Jensen, Jørgensen, & Johansen, 2002; Rietjens, Kuipers, Kester, & Keizer, 2001). For example, Foss and Hallén (2005) validated the Oxycon-Pro® (Oxycon Pro, Erich Jaeger, Germany) against the Douglas bag. No significant differences were found in the measurement of oxygen uptake ($\dot{V}O_2$) (P > 0.05), carbon dioxide output ($\dot{V}CO_2$) (P > 0.05), and minute ventilation ($\dot{V}E$) (P > 0.05) at given exercise intensities. Jensen et al. (2002) compared a metabolic cart (AMIS 2001, Innovision, Odense, Denmark) with the Douglas bag and resulted in no significant differences in $\dot{V}O_2$ and $\dot{V}E$ between the two systems (with P > 0.05). Also Rietjens et al. (2001) compared the Oxycon-Pro® with the Douglas bag at various exercise intensities. Results indicated no significant differences in $\dot{V}E$ (with P < 0.001), $\dot{V}O_2$ (with P < 0.001) or $\dot{V}CO_2$ (with P < 0.001).

However, in the past, other studies have been conducted revealing that these automated systems have not always proven to measure metabolic gases accurately, only making them valid by using an equation to correct the error (e.g., Díaz et al., 2008; Macfarlane, 2001; Perret & Mueller, 2006). For example, Diaz et al. (2008) showed significant differences in the measurement of $\dot{V}O_2$ at certain set intensities between the Oxycon Pro® (Erich Jaeger, Viasys Healthcare, Germany) and the Oxycon Mobile® (Erich Jaeger, Viasys Healthcare, Germany) (P = 0.09), with Oxycon Pro® producing (non-significantly) higher values than Oxycon Mobile®. In Perret and Mueller (2006), Oxycon Mobile® significantly underestimated $\dot{V}O_2$ compared to the Oxycon Pro® at certain exercise intensities (P < 0.05). Additionally, in a review, Macfarlane (2001) draws attention to variation in measurement in different metabolic gas analyzers. He suggests, that caution needs to be taken when using metabolic gas analyzers interchangeably, as not all equipment produce the same values. Since these studies have shown a discrepancy in the measurement of either $\dot{V}O_2$, $\dot{V}CO_2$, or $\dot{V}E$, or even all of these parameters, it seems necessary to compare metabolic gas analyzers when planning to use them interchangeably.

As mentioned above, NIRS can be used to determine an athletes individual endurance capacity (e.g. Bellotti et al., 2013; Wang, Tian, Zhang, & Gong, 2009). NIRS has previously been proven to be a reliable method to determine muscle oxygen saturation (Austin et al., 2005) and muscle oxygenation and deoxygenation (Van Der Zwaard et al., 2016), as well as a valid method to establish ventilatory thresholds (e.g. Bellotti et al., 2013; Van Der Zwaard et al.,

2016). According to, for example, van der Zwaard et al. (2016) or Wang et al. (2009), the anaerobic threshold (AT) can be found at the breakpoint of the oxygenation index (OI). Among others, Bellotti et al. (2013) and Fontana et al. (2015) have found that the breakpoint in deoxyhemoglobin represents the respiratory compensation point (RCP). In order to find out whether or not these thresholds are valid, the NIRS derived thresholds are compared to the threshold of ventilatory parameters measured with a metabolic gas analyzer, which is considered the gold standard for determining thresholds (Bhambhani, 2004). Some studies also compare the NIRS derived thresholds to lactate thresholds (LT) (B. Wang, Tian, & Zhang, 2012).

However, some of these studies also found significant differences between NIRS derived thresholds and metabolic thresholds (B. Wang, Tian, et al., 2012; B. Wang, Xu, et al., 2012). Wang, Xu, et al. (2012) for example revealed significant differences between the AT determined from NIRS parameters compared to the AT determined from ventilatory parameters ($P < 0.001$). Also Wang, Tian, et al. (2012) showed significant differences between the thresholds ($P < 0.001$)

2. Aims and Goals of the Thesis

The goal of this thesis is to provide an insight on metabolic gas analyzers and the differences between devices. In addition, this thesis should provide an understanding of how NIRS can be applied when used in an applied setting and should offer information on the use of NIRS in performance assessment. Literature on these topics is analyzed and discussed. Furthermore, a study was conducted, in order to provide additional specific data on this subject matter.

The first aim of the study is to compare two routinely used devices, namely the MetaLyzer 3B (Cortex Biophysik GmbH., Leipzig, Germany) (ML 3B), and the MasterScreen CPX (Viasys Healthcare, Höchberg, Germany) (MS CPX). Observations at the Austrian Institute of Sports Medicine (ÖISM) have indicated that these apparatus might reveal a difference in measurement and therefore it is necessary to evaluate them. For athletes, who want to interchange systems it is essential to know of any discrepancies in order to assess performance accurately. If this study results in inconsistencies between the measurements of the metabolic gas analyzers, another aim of this study is to create an equation with which the systems can be made interchangeable.

The third aim of the study is to evaluate the validity of NIRS to determine the first and second threshold. NIRS has shown to be a valid method to establish thresholds (e.g. Bellotti et al., 2013; Fontana et al., 2015; Van Der Zwaard et al., 2016). However, due to pilot research conducted in the laboratory, indicating that this might not be the case, it was considered necessary to revalidate NIRS as a method to validly determine thresholds.

3. Rationale of the Thesis

The main reason for this thesis is that the results are essential to athletes who commonly use either of the two metabolic gas analyzers. Due to a substitution of the device an athlete who has always used the MS CPX is now going to use the ML 3B and therefore it is necessary to be able to compare parameters between the two devices. If the metabolic gas analyzers result in no significant difference, the athlete can assume a change for example in the maximal oxygen uptake ($\dot{V}O_{2peak}$) is due to a training intervention. However, if the devices produce different values, the athlete can attribute a difference in $\dot{V}O_{2peak}$ to a certain extent to the metabolic gas analyzers. If this is the case, it would be beneficial for an athlete to be able to have an equation, with which the values can be compared.

Concerning the use of NIRS as a tool to determine exercise intensities, there are some gaps in the literature. For example, very few studies have been conducted with an incremental running test (Karatzanos, Paradisis, Zacharogiannis, Tziortzis, & Nanas, 2010; Snyder & Parmenter, 2009), most studies have included incremental cycling tests (e.g. Fontana et al., 2015; Miura et al., 1998). There also seems to be little agreement in the literature on what parameter is optimal for determining the AT.

Therefore it is necessary to validate NIRS as a method to determine the AT and the respiratory compensation point during an incremental running test.

4. Performance Assessment

Performance assessment is the foundation of training control (Weineck, 2010). It is necessary before and in the course of a training cycle in order to determine strengths and weaknesses of an athlete and to adapt the training program. It is also crucial when comparing an athlete's current state with an optimal or desirable state (Hohmann et al., 2007; Schnabel, Harre, & Borde, 1997). Performance assessment is possible in all different sports, however, for this thesis, the assessment of endurance capacity is of interest. There are many different methods to establish an athlete's individual endurance capacity, including the determination of lactate thresholds (Brooks, 1985), ventilatory thresholds (Beaver et al., 1986), NIRS derived thresholds (Bellotti et al., 2013; Van Der Zwaard et al., 2016) or critical power (Poole et al., 2017). However, no matter what method is used, the requirement for any test is quality criteria.

5. Quality Criteria

Performance assessment requires quality criteria to ensure meaningful results with little error. Quality criteria include objectivity, reliability, and validity (Hohmann et al., 2007; Hottenrott & Neumann, 2010; Weineck, 2010).

5.1. Objectivity

Methods of measurement are considered objective, when the execution, the analysis and the interpretation is standardized. Measurements need to be independent of the people performing the test, as well as independent of the people who are analyzing and interpreting the results (Hohmann et al., 2007; Weineck, 2010).

5.2. Reliability

Reliability describes the accuracy of a test. The less reliable a test is, the larger the error of measure and therefore there is a reduced ability to compare results between two tests (Hohmann et al., 2007; Weineck, 2010).

5.3. Validity

Validity of a method of assessment is the degree to which it measures what is being asked of it to measure. Also, the method of measurement can be considered valid, when it produces same results as another method, which measures the same parameter (Hohmann et al., 2007; Weineck, 2010).

6. The Metabolic Gas Analyzer

6.1. History of Metabolic Gas Analyzers

First experiments of measuring gas exchange go back to the 1790s and early 19 hundreds. However, the systems used were very inaccurate and not comparable with today's metabolic gas analyzers (Hollmann et al., 2006). In 1911 Douglas developed the first true metabolic gas analyzer system, which was able to analyze the components of breath. In the original system created by Douglas, expired air was collected in a bag, which was worn while exercising. A rubber tube led over the head with a support system to the mouth, as can be seen in Figure 1. The mouthpiece was connected to a valve, with which one could regulate when expired air flowed into the bag. The expired collected air could then be analyzed (Douglas, 1911).
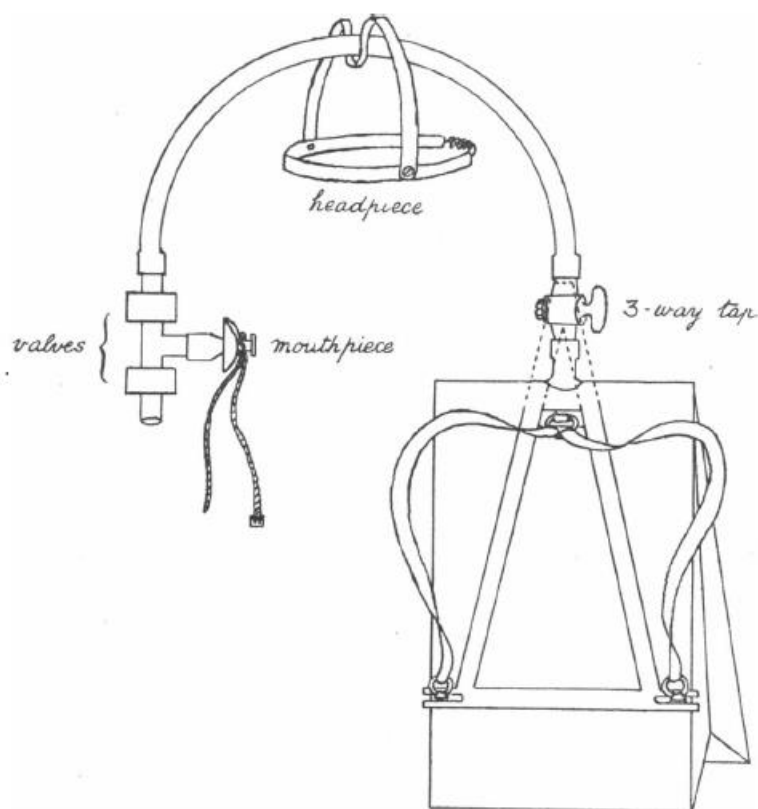

Figure 1. Image of a Douglas bag (Douglas, 1911).

Following this, Hill discovered $\dot{V}O_{2peak}$ in 1924. He also defined the terms "O_2-deficit", "steady state", and "O_2-debt" (Hill, Long, & Lupton, 1924). At the same time, Knipping (1924; as cited in Hollmann et al., 2006) developed a spirometer system, which could measure breathing frequency, breathing depth, respiratory minute volume, O_2 uptake, and CO_2 output, which was soon used worldwide. Knipping combined spirometry with work load and compared values at set difficulties (Knipping, 1924; as cited in Hollmann et al., 2006). By 1929 the term

"spiroergometry" was commonly known in the field and used among athletes as well as in clinical settings. For this reason, 1929 can be considered the true beginning of performance assessment (Hollmann et al., 2006).

Originally, the only reliable value measured was $\dot{V}O_{2peak}$. In 1959, Hollmann attempted to measure submaximal values, resulting in the ability to determine the so called "aerobic-anaerobic transition" and the "oxygen maximal steady state" (Hollmann et al., 2006). Wasserman and McIlroy (1964) used the same method to determine the threshold and named it "anaerobic threshold". This term became the commonly used one.

In the 1960s, breath-by-breath systems were developed, and data was recorded digitally. Since then, portable metabolic gas analyzers have been produced, enabling testing in the field (Hollmann et al., 2006).

6.2.   Types of Metabolic Gas Analyzers

Concerning the method of measurement, there are two different types of metabolic gas analyzers, namely the mixing chamber and the breath-by-breath method. In a metabolic gas analyzer, where a mixing chamber is used, a certain known amount of expired air is collected in a mixing chamber and the air components are measured. In a breath-by-breath system, the components of the inspired and expired air are measured at each individual breath directly at the mouth (Beijst, Schep, Breda, Wijn, & Pul, 2013; Hodges, Brodie, & Bromley, 2005).

Metabolic gas analyzers can also be categorized according to their size and mobility. One can discern between a metabolic cart, a stationary, often also called mobile metabolic gas analyzer, and the portable gas analyzer. The metabolic cart is cumbersome when it comes to moving it around. It is a big cart with a computer and the machine, and can only be moved within a room, but cannot be transported into the field. Therefore it is only useable for laboratory tests, and not field tests. The stationary, or mobile, metabolic gas analyzer can also only be used in the laboratory. It is not as big or cumbersome as a metabolic cart, as it is quite small and can be connected to a laptop. It can be transported to different places easily, but is stationary and cannot be used in the field. The portable metabolic gas analyzer is the smallest of them all and is connected to the software via Bluetooth, enabling it to be transported with the athlete and permitting it to be used in the field. It is lightweight and optimally should not influence the

athlete in his or her performance.

6.3. Measured Parameters

Breath-by-breath metabolic gas analyzers measure three components, of which other parameters can be calculated. The measured parameters are $\dot{V}O_2$, $\dot{V}CO_2$, and minute ventilation ($\dot{V}E$). Using these parameters, calculations can be done to obtain other parameters of interest, such as respiratory exchange ratio (RER; i.e. $\dot{V}CO_2/\dot{V}O_2$), equivalents for O_2 (i.e. $\dot{V}E/\dot{V}O_2$) and CO_2 (i.e. $\dot{V}E/\dot{V}CO_2$), and end-tidal oxygen tension ($PETO_2$) and end-tidal carbon dioxide tension ($PETCO_2$) (Macfarlane, 2001). Following, each one of these parameters will be explained.

6.3.1. $\dot{V}O_2$

The main parameter of interest for athletes is $\dot{V}O_2$. $\dot{V}O_2$ describes the oxygen uptake per min which is measured at a given intensity and is commonly measured in $L \cdot min^{-1}$. When performing a GXT, the $\dot{V}O_2$ at maximal running speed is called the $\dot{V}O_{2peak}$, which is a determinant of an athlete's endurance capacity. $\dot{V}O_{2peak}$ was first determined by Hill et al. (Hill et al., 1924; Hill & Lupton, 1923). It is considered the largest oxygen uptake per minute which is reachable during dynamic work under use of large muscle groups or whole body exercise.

$\dot{V}O_{2peak}$ is traditionally used as the gold standard for determination of cardiorespiratory fitness, as this is the only parameter which defines the cardiovascular and respiratory capacity. $\dot{V}O_{2peak}$ indicates an athlete's potential for endurance performance and is an indicator of training status. $\dot{V}O_{2peak}$ can be influenced through training, as well as detraining or physical inactivity. It is a fixed parameter at a given intensity which varies according to the method of training, whether it is running, cycling, rowing or any other endurance activity (Eston & Reilly, 1996). As $\dot{V}O_{2peak}$ is significantly positively related to body mass, maximal oxygen uptake can be set in relation to bodyweight and therefore can be made comparable between people. The $\dot{V}O_{2peak}$ of an average person is 40 $mL \cdot kg^{-1} \cdot min^{-1}$ for women and 55 $mL \cdot kg^{-1} \cdot min^{-1}$ for men. An elite athlete's $\dot{V}O_{2peak}$ can exceed 80 $mL \cdot kg^{-1} \cdot min^{-1}$ (Eston & Reilly, 1996). There is a significant negative correlation between $\dot{V}O_2$ and height if body mass is excluded. This means that when comparing two men with the same body mass but different height, the taller person will have

a lower $\dot{V}O_{2peak}$ (Kroidl, Schwarz, & Lehnigk, 2010).

6.3.2. $\dot{V}CO_2$

$\dot{V}CO_2$ is the amount of CO_2, which is measured in the expired air in $L \cdot min^{-1}$. As the intensity increases, the body uses more carbohydrates as energy supply. The oxidation of carbohydrates produces more CO_2 than the oxidation of fat. For this reason, $\dot{V}CO_2$ inclines over proportionally as work load increases (Kroidl et al., 2010).

6.3.3. $\dot{V}E$

The components of minute ventilation are breathing frequency and breathing depth. During physical activity, first breathing depth increases and then breathing frequency. $\dot{V}E$ of an average man is 120-140 $L \cdot min^{-1}$ at the point of $\dot{V}O_{2peak}$. An endurance trained athlete can reach up to 250 $L \cdot min^{-1}$ at $\dot{V}O_{2peak}$. Breathing frequency hardly ever exceeds 60 breaths per min. $\dot{V}E$ increases linearly to $\dot{V}O_2$ at light to moderate exercise and has a relatively larger increase during heavier exercise. When analyzing $\dot{V}E$ during a GXT, the RCP can be determined around the point where $\dot{V}E$ increases at a larger rate than before (i.e. non-linear increase) (Hollmann et al., 2006).

Endurance training has an influence on $\dot{V}E$, namely in the way that an endurance trained athlete has a reduced $\dot{V}E$ at a given intensity than an untrained or not as well trained person. The reason for this is that the tidal volume increases and therefore, more $\dot{V}O_2$ is higher per breath. In addition, the air stays in the lungs longer, enabling a more efficient O_2 exchange (Eston & Reilly, 1996)

6.3.4. $PETO_2$ and $PETCO_2$

$PETO_2$ and $PETCO_2$ stand for end-tidal oxygen tension and end-tidal carbon dioxide tension, respectively, and describes the pressure of O_2 and CO_2 in the exhaled gas. It is measured in mmHg. Generally, during a GXT, $PETO_2$ usually decreases first, and after it plateaus it increases again. $PETCO_2$ will first increase, then plateau before it starts decreasing again. (Beaver et al., 1986; Kroidl et al., 2010). In Figure 2, $PETO_2$ and $PETCO_2$ are presented for
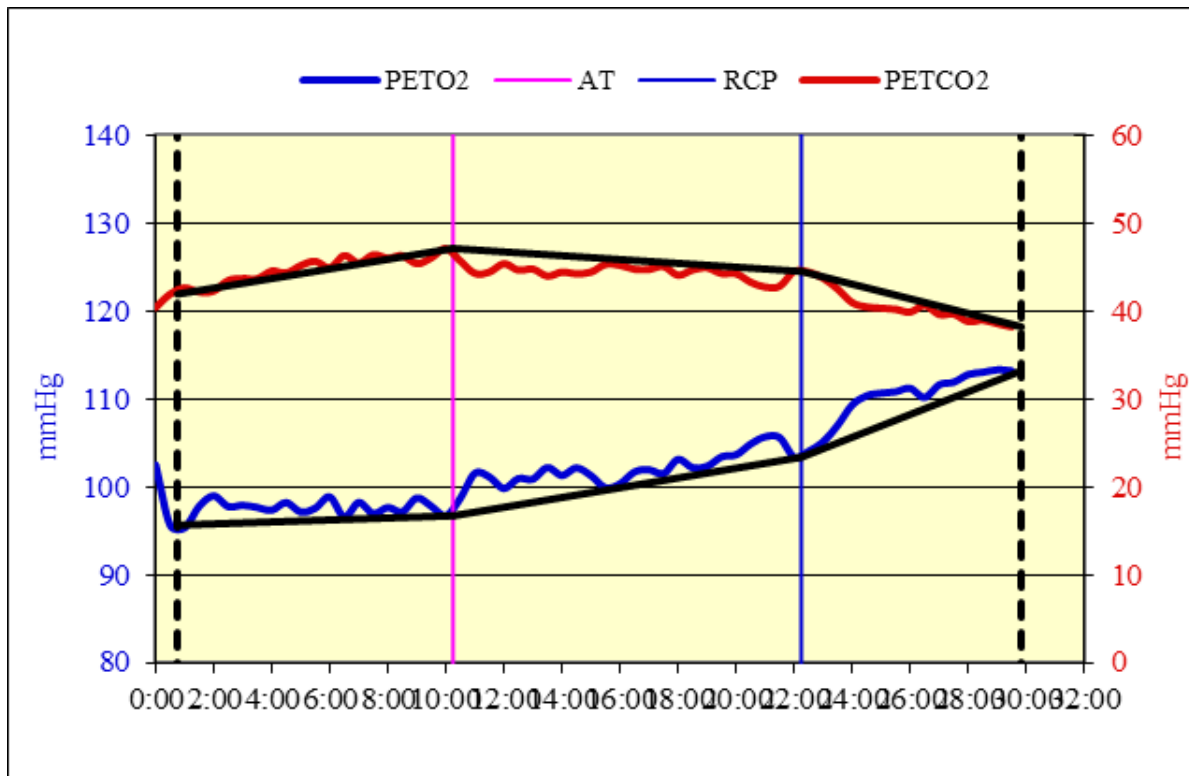
subject 9.

Figure 2. Graph of PETO₂ and PETCO₂ for subject 9 to determine AT and RCP. Notes: The vertical dotted black lines indicate beginning and end of GXT, the red line is PETCO₂, the blue line is PETO₂, the vertical pink line is AT and the vertical blue line is RCP.

6.3.5. EqO₂ and EqCO₂

The equivalents of O_2 and CO_2 describe the amount of O_2 and CO_2 which are in the expired air and is calculated accordingly: $\dot{V}E/\dot{V}O_2$ and $\dot{V}E/\dot{V}CO_2$. $\dot{V}E$ and $\dot{V}O_2$ are both measured in $L \cdot min^{-1}$ and EqO₂ and EqCO₂ do not have units. EqO₂ and EqCO₂ are a determinant for how intense a workload is perceived. Depending on age, sex and fitness level, a person will be exhausted when EqO₂ starts reaching values larger than 30 to 35, and athletes can reach values of up to 40 or 50 (Kroidl et al., 2010).

When preforming a GXT, as the load starts increasing, EqO₂ decreases until it reaches a minimum, which indicates that breathing economy is most efficient at this point. After the minimum is reached, the value begins to increase linearly until RCP, at which point EqO₂ and EqCO₂ increase at a quicker rate (Beaver et al., 1986; Kroidl et al., 2010).

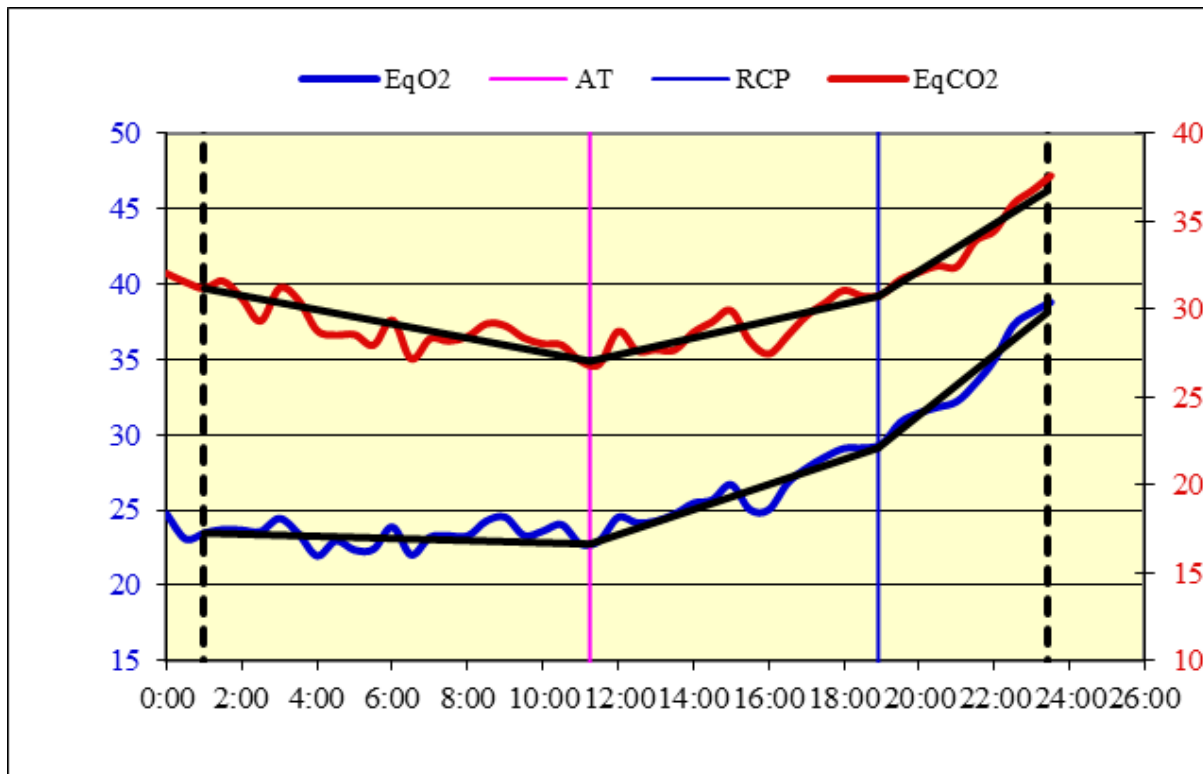Figure 3 shows the trend of EqO₂ and EqCO₂ during a GXT in subject 10.

Figure 3. Graph of EqO₂ and EqCO₂ for subject 10 to determine AT and RCP. Notes: The vertical dotted black lines indicate beginning and end of GXT, the red line is EqCO₂, the blue line is EqO₂, the vertical pink line is AT and the vertical blue line is RCP.

6.3.6. RER

The RER (Respiratory exchange ratio), also known as RQ (respiratory quotient), shows the relationship between $\dot{V}CO_2$ and $\dot{V}O_2$ and is calculated as follows: $\dot{V}CO_2$ (L·min^{-1}) / $\dot{V}O_2$ (L·min^{-1}). With an increase in load during a GXT, $\dot{V}CO_2$ increases resulting in a larger RER. At RCP, the RER is around 1 and at exhaustion RER is larger than 1. Endurance athletes are able to reach values around 1.15 (Beaver et al., 1986; Hollmann et al., 2006; Kroidl et al., 2010). Figure 4 shows an example of subject 14 of how $\dot{V}CO_2$ and $\dot{V}O_2$ cross each other at RCP, indicating that RER is 1.
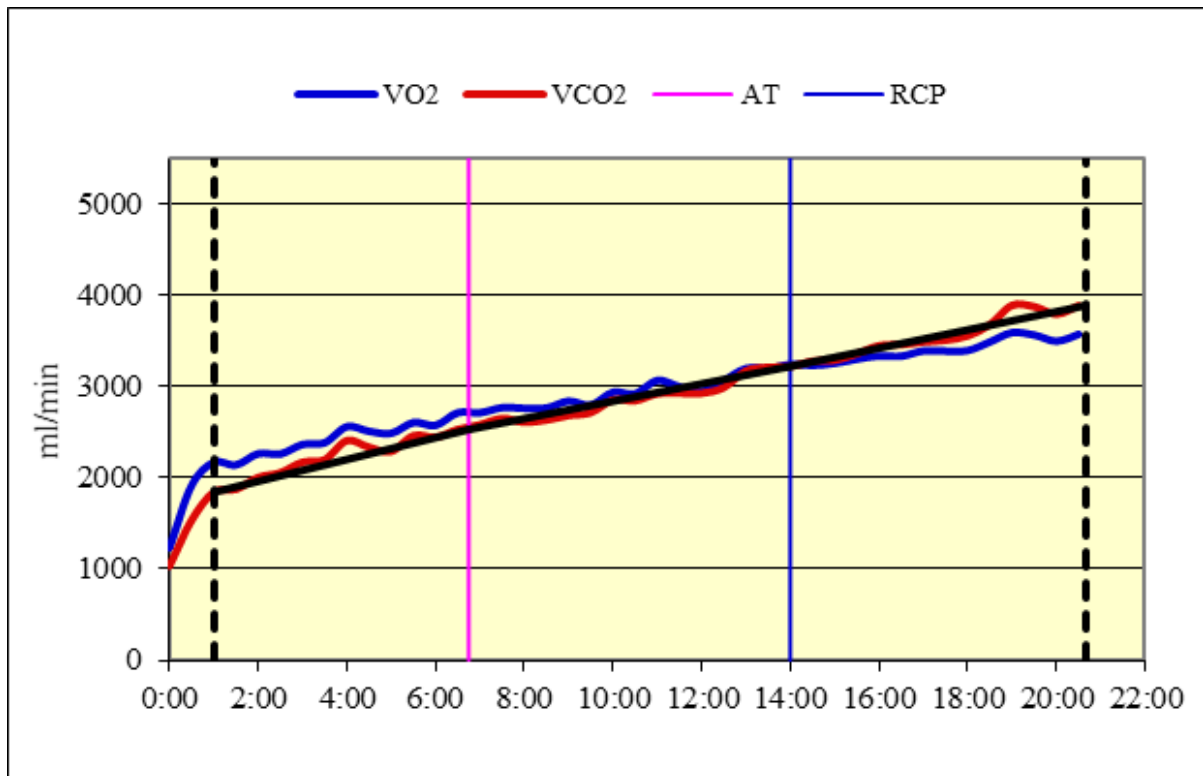
Figure 4. Graph of $\dot{V}O_2$ and $\dot{V}CO_2$ for subject 14 to determine AT and RCP. Notes: The vertical dotted black lines indicate beginning and end of GXT, the red line is $\dot{V}CO_2$, the blue line is $\dot{V}O_2$, the vertical pink line is AT and the vertical blue line is RCP.

6.4. Ventilatory Thresholds

Based on these measured parameters, conclusions can be drawn on an athlete's current endurance capacity and training intensities can be determined. When talking about performance assessment, peak values and maximal performance are important, but also submaximal data is of interest. Especially data at certain intensities are significant, as these values are used in training control. In order to understand this, it is necessary to give a brief overview of energy supply during exercise and the role these intensities play in an athletes training.

6.4.1. Energy Supply

The three energy substrates are carbohydrates, fats and proteins, which can be broken down to release energy. This energy is stored in the form of adenosine triphosphate (ATP). There are three different processes to produce ATP. The simplest and quickest system is with the use of creatine phosphate, which regenerates ATP. However, this can only provide energy for 3-5 s.

If workloads last longer, the glycolytic system and the oxidative system come into play. The glycolytic system produces ATP by breaking down glucose or glycogen. Although this energy system produces less ATP, it is able to supply energy under limited oxygen resources, which is used predominantly during the early phases of high-intensity exercise. The end product is lactic acid, which inhibits further glycogen breakdown. When exercise lasts more than about two min, athletes rely on the energy supply of the oxidative system. It is considered an aerobic process, as oxygen is used to break down the substrates. Carbohydrates provide the quickest supply of energy and are the main source of energy at higher intensities, producing an over proportional amount of CO_2 when oxidized. Fatty acid oxidation is a slower process, however it can produce limitless amounts of ATP. This is beneficial for lower intensities of long duration. Protein can also be used, if it is converted into glucose beforehand, however, carbohydrates and fatty acids are the main source of energy (McArdle, Katch, & Katch, 2001; Wilmore, Costill, & Kenney, 2008).

6.4.2. Anaerobic Threshold

The anaerobic threshold (AT) is the transition point from aerobic to anaerobic energy supply. This is an important point to determine, as it is the point of optimal efficiency. It is a steady state which can be maintained over a very long time (Kroidl et al., 2010). This transition point is clearly visible when $\dot{V}E$ and $\dot{V}CO_2$ increase in a non-linear fashion. This is the sign, that more CO_2 is being produced, as the use of carbohydrates produces more CO_2 than the use of fat. Also in relation to $\dot{V}O_2$, $\dot{V}CO_2$ increases over proportionally at this point.

Kroidl et al. (2010) (including information from Beaver et al. (1986)) describe different methods to determine the AT:

- The point at which EqO_2 begins to rise, after beforehand declining or being flat while at the same time $EqCO_2$ declines or stays constant is considered the AT.
- When looking at the graph of VCO_2 against $\dot{V}E$, the first breakpoint is the AT.
- Often, in the $\dot{V}E$ curve, two breakpoints are visible, of which the first one is the AT (Figure 5).
- At the point where $PETO_2$ quits decreasing and begins to rise and $PETCO_2$ quits increasing or increases slowly, the AT can be set.

In addition, when examining the graphs for the AT, it is crucial to make sure, that the points, which are being set, take place before RER reaches one, as any breakpoint after this point

would indicate a high use of carbohydrates, as these produce more CO_2 and therefore exceed O_2 uptake.
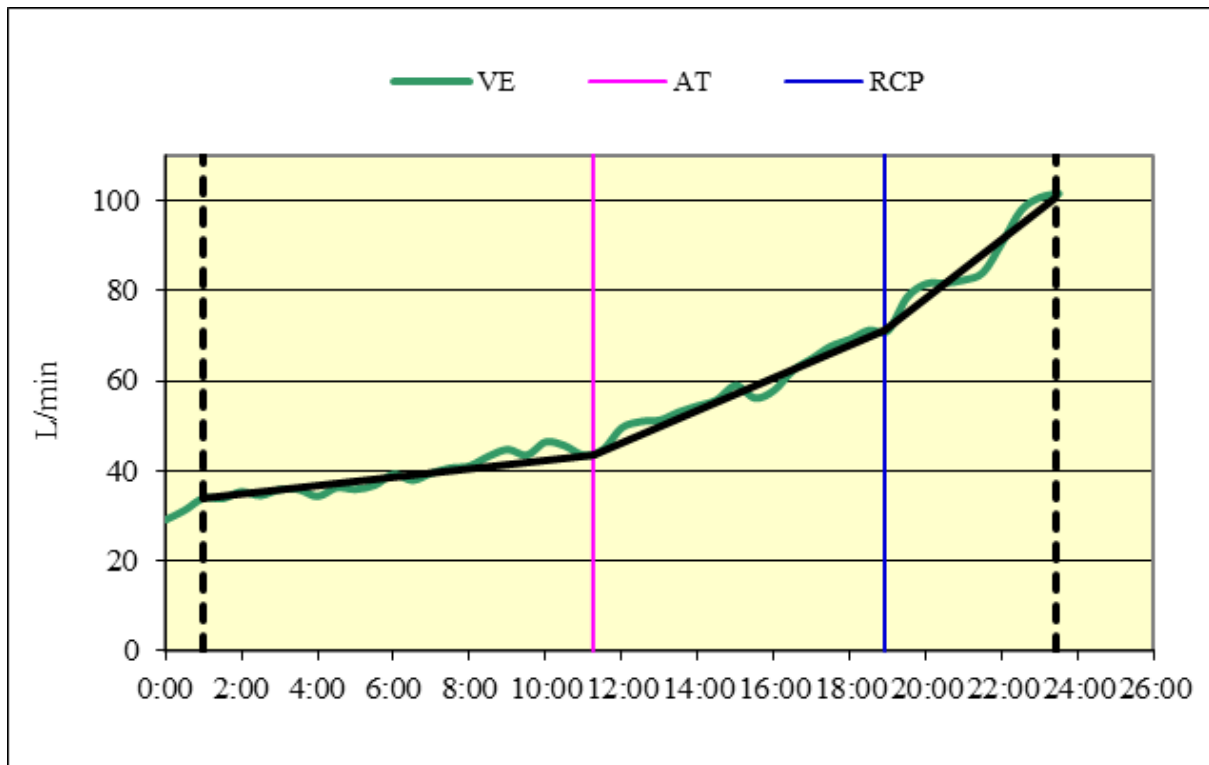

Figure 5. Graph of $\dot{V}E$ of subject 10 to determine AT and RCP. Notes: The vertical dotted black lines indicate beginning and end of GXT, the green line is $\dot{V}E$, the vertical pink line is AT and the vertical blue line is RCP.

6.4.3. Respiratory Compensation Point

The respiratory compensation point is the point at which a maximal O_2 uptake with minimal respiratory work is achieved and athletes are able to maintain this intensity for a long time period. Hollmann et al (2006) describe this as the "point of optimal efficiency of respiration". This threshold can also be termed the maximal steady state, as it is the highest intensity which can still be accomplished aerobically. This point also coincides with the point at which lactate begins to increase exponentially, and is therefore also called the maximal lactate steady state. According to Wasserman (1984) and there are different methods to determine the RCP when using metabolic gas analyzers.

- When $\dot{V}CO_2$ is plotted against $\dot{V}O_2$, there is a point at which CO_2 expiration increases more than O_2 uptake, resulting in a change in linearity. This point can be taken to determine RCP.
- With increased intensity, EqO_2 increases in a characteristic way. As there is more CO_2

being expired at RCP, $\dot{V}E$ also increases, resulting in the numerator increasing more in relation to the denominator which means the EqO_2 increases over proportionally. At this point the RCP can be placed.

- An increase in $PETO_2$ is also a determinant for RCP. The reason for this is the hyperventilation which occurs in order to breathe out more CO_2.

- The increase in RER can also be taken as the point at which RCP can be determined. When RER reaches 1, this is usually a determinant for RCP. However this value can be influenced by intentional hyperventilation. At this point $\dot{V}O_2$ and $\dot{V}CO_2$ is the same and beyond this point, $\dot{V}CO_2$ increases more than $\dot{V}O_2$, resulting in an RER larger than 1.

- Also, when $\dot{V}E$ begins to increase exponentially, RCP can be placed, as seen in Figure 6 for subject 4.
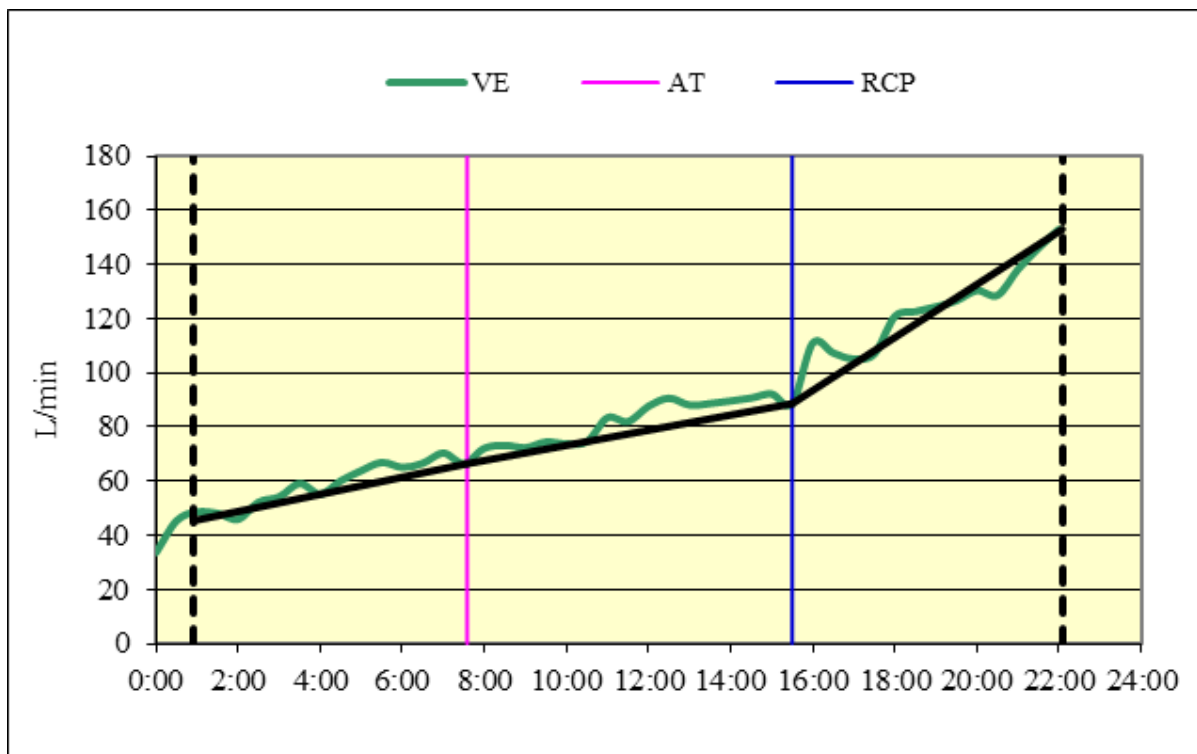

Figure 6. Graph of $\dot{V}E$ of subject 4 to determine and RCP. Notes: The vertical dotted black lines indicate beginning and end of GXT, the green line is $\dot{V}E$, the vertical pink line is AT and the vertical blue line is RCP.

7. Studies Comparing Metabolic Gas Analyzers

Many different commercially available metabolic gas analyzers have been validated against a criterion (often Douglas bag or a metabolic cart), e.g. Oxycon pro (Erich Jaeger GmbH, Höchberg, Germany) (e.g. Díaz et al., 2008; Rietjens, Kuipers, Kester, & Keizer, 2001), MetaMax 3B (Cortex Biophysik, Leipzig, Germany) (e.g. Brandes, Klein, Ginsel, & Heitmann, 2015; D. J. Macfarlane & Wong, 2012), or the K4b² (Cosmed, Rome, Italy) (e.g. Duffield, Dawson, Pinnington, & Wong, 2004; Leprêtre et al., 2012), to name a few. Some systems have produced acceptable results (e.g. Foss & Hallén, 2005; Jensen, Jørgensen, & Johansen, 2002; Rietjens et al., 2001), while other systems have not proven to be valid (e.g. Díaz et al., 2008; Perret & Mueller, 2006).

However, the goal of this thesis is not to give an overview on all metabolic gas analyzers with their faults and advantages. This has already been done in numerous reviews (Atkinson, Davidson, & Nevill, 2005; Hodges et al., 2005; Macfarlane, 2001; Meyer, Davidson, & Kindermann, 2005). Nevertheless, as many validation studies comparing metabolic gas analyzers have been conducted, the methods can be analyzed and compared to the current study, to provide a better understanding.

7.1.   Commonly used Protocols

Most studies compared one or more metabolic gas analyzers to a criterion, which is either the Douglas bag, or another metabolic system, which has previously produced valid results when compared to the Douglas bag (e.g. Beltrami et al., 2014; Carter & Jeukendrup, 2002; Crouter et al., 2006). Other studies aim to compare two devices in order to find out whether or not there are discrepancies (e.g. Brandes et al., 2015; Díaz et al., 2008; Leprêtre et al., 2012).

When looking at the studies validating a certain metabolic system, the methods can be consolidated into different categories. Concerning the exercise protocol, studies were either based on one or more GXTs (e.g. Beltrami, Froyd, Mamen, & Noakes, 2014; Díaz et al., 2008; Leprêtre et al., 2012), on trials at certain given intensities to test validity at lower and higher workloads (e.g. Carter & Jeukendrup, 2002; Jensen et al., 2002; McLaughlin, King, Howley, Bassett, & Ainsworth, 2001), or both methods were used (e.g. Larsson, Wadell, Jakobsson, Burlin, & Henriksson-Larsén, 2004).

Regarding the method of assessing metabolic parameters, three different approaches can be found. Some studies used multiple metabolic gas analyzers simultaneously by connecting the mouthpieces in series (e.g. Beltrami et al., 2014; Foss & Hallén, 2005; Perret & Mueller, 2006). Perret and Mueller (2006) even created a special facemask to enable simultaneous assessment. Other studies used the systems consecutively during the same test, by applying one apparatus after the other (e.g. Brandes et al., 2015; Carter & Jeukendrup, 2002; D. J. Macfarlane & Wong, 2012). In these cases, the apparatus were applied and interchanged during a steady state, ensuring little variance in measured parameters. In addition, the systems were used in a randomized order. Further studies examined the metabolic gas analyzers in two separate GXTs or trials at certain intensities, separated by at least a day. This means that one test is performed with one system and the other test with another system (e.g. Díaz et al., 2008; Jensen et al., 2002; Vogler, Rice, & Gore, 2010). And then some studies used a combination of these methods (e.g. Crouter, Antczak, Hudak, DellaValle, & Haas, 2006; Duffield et al., 2004; Larsson et al., 2004).

To sum it up, studies using different metabolic gas analyzers have shown that some systems produce valid data, while other devices are more inaccurate. The Oxycon pro has proven to be a valid system when compared to the Douglas bag ($P > 0.05$) (Rietjens et al., 2001), whereas the Oxycon mobile seems to generally underestimate $\dot{V}O_2$ (Díaz et al., 2008; Perret & Mueller, 2006). The MetaMax 3B often overestimates $\dot{V}O_2$ compared to the Douglas bag ($P < 0.01$) (D. J. Macfarlane & Wong, 2012; Vogler et al., 2010), and also the K4b² system has shown to produce higher $\dot{V}O_2$ values than the Douglas bag ($P < 0.05$) (McLaughlin et al., 2001).

8. Overview of the Literature Determining NIRS Derived AT and NIRS Derived RCP

In order to set the stage for the empirical part of this thesis, literature on the topic of determining the anaerobic threshold and respiratory compensation point with NIRS will be retrieved and analyzed. This should give an overview of the different methods to determine the threshold and should make clear, why the methods for the study in this thesis were chosen.

8.1. Inclusion and Exclusion Criteria

All studies which were published online in the database Pubmed (https://www.ncbi.nlm.nih.gov/pubmed/) by February 2019 were considered for this thesis. In order to find the literature, various combinations of the following keywords were used:

Near infrared spectroscopy, anaerobic threshold, ventilatory threshold, respiratory compensation point, muscle tissue deoxygenation.

Additionally, any relevant literature which was found in the references of these articles and in meta analyses and reviews was looked at as well. Studies which included healthy physically active and healthy sedentary participants were chosen. If the study included healthy participants as well as patients with disease, the study was included, but only paid attention to the healthy subjects. Also, any studies, which specifically concentrated only on the elderly or children as subjects were excluded, but here again, if other participants were also included, the study was considered as eligible. Studies had to include a comparison of NIRS thresholds and ventilatory or lactate thresholds in order to be considered. It was not sufficient if the NIRS threshold was determined, but not compared with other methods of determining thresholds.

8.2. Studies

Following are two tables giving an overview of the studies chosen to be analyzed. The articles are divided into two tables according to whether NIRS derived AT (NdAT) or NIRS derived RCP (NdRCP) was determined. Altogether, 14 studies met the inclusion criteria, seven of which analyzed AT, five deal with RCP, and two include both AT and RCP.

Table 1. Overview of all the studies determining NdAT and comparing NdAT with other thresholds. In the column "methods" the tests, the placement of the NIRS, and the method for NdAT determination is described.

Study	Aim	Subjects	Methods	Results
Coquart et al., 2017	Determine NdAT in $\Delta[O_2Hb]$ compared to metabolic AT	14 trained male cyclists (20.6 ± 2.9 years) $\dot{V}O_{2peak}$: 65.1 ± 6.0 mL·min^{-1}·kg^{-1}	Incremental cycling test, starting at 150 W, increments of 50 W every 4 min until 300 W, then 25 Watt every 2 min Metabolic gas analyzer (breath-by-breath) and NIRS on m. vastus lateralis Visually seen inflection point in $\Delta[O_2Hb]$ to determine NdAT	No significant difference between AT and NdAT ($P > 0.05$)
Grassi et al., 1999	Determine NdAT in OI compared to lactate threshold	5 well trained male mountain climbers (32.8 ± 5.4 years) $\dot{V}O_{2peak}$: 51.0 ± 4.2 mL·min^{-1}·kg^{-1}	Incremental cycling test, starting at 60 W, increments of 30 W every 3 min Metabolic gas analyzer (breath-by-breath), lactate and NIRS on m. vastus lateralis Visual inspection of OI to determine NdAT	Only regression and correlation calculations. Onset of blood lactate accumulation significantly correlated with onset of muscle deoxygenation (r = 0.95)

Study	Aim	Subjects	Methods	Results
Miura et al., 1998	Determine NdAT in $\Delta[O_2Hb]$ in 2 groups (healthy active, healthy sedentary) compared to metabolic threshold	Total of 27 6 male active normal subjects (23 $\pm$ 3 years) $\dot{V}O_{2peak}$: 41.4 $\pm$ 6.0 mL·min^{-1}·kg^{-1} 21 sedentary normal subjects (14 males, 7 females, 27 $\pm$ 4 years) $\dot{V}O_{2peak}$: 27.4 $\pm$ 3.5 mL·min^{-1}·kg^{-1}	Incremental cycling test starting at 20 W, increments of 15 or 20 W·min^{-1} for sedentary, 30 W·min^{-1} for active subjects Metabolic gas analyzer (breath-by-breath) and NIRS on m. vastus lateralis Manually placed regression lines to determine NdAT in $\Delta[O_2Hb]$.	No significant difference in AT and NdAT in active and sedentary subjects (with $P < 0.05$)
Racinais et al., 2014	Determine NdAT in $\Delta[O_2Hb]$ and $\Delta[HHb]$ compared to metabolic AT	25 trained cyclists (37 $\pm$ 8 years) $\dot{V}O_{2peak}$: 53.0 $\pm$ 8.0 mL·min^{-1}·kg^{-1}	Incremental cycling test, with increments of 25 W·min^{-1} Metabolic gas analyzer (breath-by-breath) and NIRS on m. vastus lateralis NdAT determined with regression lines with the smallest sum of errors in $\Delta[O_2Hb]$ and $\Delta[HHb]$	NdAT was significantly different than AT and therefore not determinable ($P < 0.05$)

Study	Aim	Subjects	Methods	Results
Raleigh et al., 2018	Breakpoint in tissue saturation index (TSI) compared to lactate threshold (LT) at aerobic-anaerobic transition (AAT)	31 male cyclists and triathletes (29 ± 9 years) $\dot{V}O_{2peak}$: 62.5 ± 8.1 mL·min^{-1}·kg^{-1}	GXT cycling, starting at 100 Watt, increments of 25 W every 3 min Metabolic gas analyzer (breath-by-breath), lactate and NIRS on m. vastus lateralis Two regression lines with lowest sum of squared residuals to determine TSI breakpoint	TSI threshold not different than lactate threshold (P > 0.05) but AAT occurred at higher workload than both LT and TSI threshold (P < 0.05)
van der Zwaard et al., 2016	Determine NdAT in OI compared to metabolic AT across sexes and training status	40 subjects: 10 trained male cyclists (23 ± 3 years), $\dot{V}O_{2peak}$: 60.0 ± 6.6 mL·min^{-1}·kg^{-1}, 10 trained female cyclists (24 ± 4 years), $\dot{V}O_{2peak}$: 53.6 ± 4.5 mL·min^{-1}·kg^{-1}, 11 endurance trained males (23 ± 2 years), $\dot{V}O_{2peak}$: 60.8 ± 5.5 mL·min^{-1}·kg^{-1} 9 recreationally trained males(24 ± 2 years), $\dot{V}O_{2peak}$: 48.7 ± 5.3 mL·min^{-1}·kg^{-1}, classified according to $\dot{V}O_{2peak}$	Incremental cycling test, starting at 1.5 W·kg^{-1} and increased by 0.5 W·kg^{-1} every 3 min Metabolic gas analyzer (breath-by-breath) and NIRS on m. vastus lateralis Double linear regression method to determine breakpoint in OI	In all subjects, breakpoint in OI was not significantly different from AT (with P < 0.05)

Study	Aim	Subjects	Methods	Results
Wang et al., 2009	Determine NdAT in OI compared to lactate threshold (LT)	8 female fin swimmers (19.8 ± 2.4 years) $\dot{V}O_{2peak}$: 53.4 ± 12.2 mL·min^{-1}·kg^{-1}	Incremental cycling test starting at 40 W, increments of 30 W every 3 min Lactate and NIRS on m. vastus lateralis Manually placed regression lines in OI to determine breakpoint	No significant difference between OI breakpoint and (LT) (with P < 0.05)
Wang, Xu, et al., 2012	Determine NdAT in OI compared to metabolic AT and lactate threshold between two muscles, namely m. vastus lateralis and m. gastrocnemius lateralis	31 active college students (12 men, 19 females) (19.7 ± 1.8 years) $\dot{V}O_{2peak}$: 51.2 ± 3.4 mL·min^{-1}·kg^{-1}	Incremental cycling test, starting at 100 W for men and 40 W for women, increments of 30 W every 3 min Metabolic gas analyzer (breath-by-breath), lactate and NIRS on m. vastus lateralis and m. gastrocnemius lateralis Linear regression with least squares method to determine NdAT in OI	Significant difference between vastus lateralis and gastrocnemius lateralis in OI threshold (P < 0.001), breakpoint in vastus lateralis appeared significantly earlier than breakpoint in gastrocnemius lateralis, ventilatory, and lactate threshold. But all thresholds were well correlated.

Study	Aim	Subjects	Methods	Results
Wang, Tian, et al., 2012	Determine NdAT in OI with compared to metabolic AT and LT	16 female middle and long distance swimmers (19 ± 0.5 years) $\dot{V}O_{2peak}$: 53.4 ± 2.2 mL·min^{-1}·kg^{-1}	Incremental cycling test, starting at 40 W, increments of 30 W every 3 min Metabolic gas analyzer (breath-by-breath), lactate , and NIRS on m. vastus lateralis Two regression lines with lowest sum of squared residuals to determine NdAT in OI	NdAT appeared slightly earlier than lactate threshold, but not significantly different (P = 0.063), NdAT and lactate threshold both appeared significantly earlier than AT (P < 0.001)

Table 2. Overview of all the studies determining NdRCP and comparing NdRCP with other thresholds. In the column "methods" the tests, the placement of the NIRS, and the method for NdRCP determination is described.

Study	Aim	Subjects	Methods	Results
Bellotti et al., 2013	Determine NdRCP with Δ[HHb] compared to metabolic RCP	32 healthy, sedentary males (48 ± 17 years) $\dot{V}O_{2peak}$: 39.4 ± 11.4 mL·min^{-1}·kg^{-1}	Incremental cycling test, starting at 50 W and increasing 10-30 W·min^{-1} Metabolic gas analyzer (breath-by-breath) and NIRS on m. vastus lateralis Double linear regression that minimized the squared sum of the residuals in HHb to determine threshold	$\dot{V}O_2$ and heartrate (HR) at NdRCP is no different than $\dot{V}O_2$ and heartrate at RCP ($P = 0.74$)
Fontana et al., 2015	Determine NdRCP with Δ[HHb] compared to metabolic RCP	118 males (47 ± 19 years) $\dot{V}O_{2peak}$: 40.0 ± 12.0 mL·min^{-1}·kg^{-1}	Incremental cycling test increasing 10-30 W·min^{-1} Metabolic gas analyzer (breath-by-breath) and NIRS on m. vastus lateralis Double linear regression that minimized the squared sum of the residuals in HHb to determine threshold	RCP and NdRCP were not significantly different (with $P < 0.05$)

Study	Aim	Subjects	Methods	Results
Karatzanos et al., 2010	NIRS TSI to determine NdRCP compared to metabolic RCP	17 subjects, 11 males, 6 females (22.5 ± 1.0 years) $\dot{V}O_{2peak}$: 48.0 ± 5.7 mL·min^{-1}·kg^{-1}	Incremental running test, starting at 2.22 m·s^{-1} and increase of 0.14 m·s^{-1} every minute. Metabolic gas analyzer and NIRS on m. gastrocnemius lateralis. TSI to determine NdRCP. 2 models to determine NdRCP: analyzing curve visually and calculation	High correlation between NdRCP and metabolic RCP. Visual curve analysis method was not significantly different ($P > 0.05$), calculated method significantly different ($P = 0.01$)
Keir et al., 2015	Comparison of $\dot{V}O_2$ associated with critical power, RCP, MLSS, and Δ[HHb] breakpoint as a boundary between heavy and very heavy exercise domains	12 males (25 ± 2 years) $\dot{V}O_{2peak}$: 49.3 ± 8.7 mL·min^{-1}·kg^{-1}	Ramp incremental cycling test, starting at 20 W for 4 min and then increasing 25 W·min^{-1}. Metabolic gas analyzer (breath-by-breath) and NIRS on m. vastus lateralis. Linear least square regression analysis to determine breakpoint in Δ[HHb].	$\dot{V}O_2$ and HR at NdRCP in Δ[HHb] is not significantly different than RCP, MLSS and critical power (with $P < 0.05$)

Study	Aim	Subjects	Methods	Results
Miura et al., 1998	Determine NdRCP in $\Delta[O_2Hb]$ and $\Delta[HHb]$ in 2 groups (healthy active, healthy sedentary) compared to metabolic RCP	Total of 27; 6 male active normal subjects (23 ± 3 years); $\dot{V}O_{2peak}$: 41.4 ± 6.0 mL·min^{-1}·kg^{-1} 21 sedentary normal subjects (14 males, 7 females, 27 ± 4 years) $\dot{V}O_{2peak}$: 27.4 ± 3.5 mL·min^{-1}·kg^{-1}	Incremental cycling test starting at 20 W, increments of 10 W·min^{-1} for heart failure, 15 or 20 W·min^{-1} for sedentary, 30 W·min^{-1} for active subjects Metabolic gas analyzer (breath-by-breath) and NIRS on m. vastus lateralis Manually placed regression lines to determine NdRCP in $\Delta[O_2Hb]$. If NdRCP could not be found in O2Hb, then $\Delta[HHb]$ was used.	No significant difference in RCP and NdRCP in both groups (with $P < 0.05$)
Racinais et al., 2014	Determine NdRCP in $\Delta[O_2Hb]$ and $\Delta[HHb]$ compared to metabolic RCP	25 trained cyclists (37 ± 8 years) $\dot{V}O_{2peak}$: 53.0 ± 8.0 mL·min^{-1}·kg^{-1}	Incremental cycling test, with increments of 25 W·min^{-1} Metabolic gas analyzer (breath-by-breath) and NIRS on m. vastus lateralis NdRCP determined with regression lines with the smallest sum of errors in $\Delta[O_2Hb]$ and $\Delta[HHb]$	NdRCP is determinable and is not significantly different from RCP (with $P < 0.15$)

Study	Aim	Subjects	Methods	Results
Snyder & Parmenter, 2009	NIRS to detect breakpoint in TSI compared to metabolic and lactate thresholds	16 well-trained runners and triathletes, 9 men (32 ± 6 years) $\dot{V}O_{2peak}$: 64.9 ± 4.9 mL·min^{-1}·kg^{-1} 7 women (31 ± 9 years) $\dot{V}O_{2peak}$: 50.8 ± 7.0 mL·min^{-1}·kg^{-1}	Incremental treadmill running test, 6 min stages with speed set individually Metabolic gas analyzer (breath-by breath), lactate, and NIRS on m. gastrocnemius lateralis Breakpoint in StO2 determined visually	Running speed at TSI breakpoint was not significantly different than running speed at lactate or metabolic thresholds (with $P < 0.05$)

8.3.   Summary of Literature Defining NdAT

All studies included in this overview to determine NdAT used an incremental cycling test to determine the threshold. In all but three studies, stages of 3 min were chosen (Grassi et al., 1999; Raleigh, Donne, & Fleming, 2018; Van Der Zwaard et al., 2016; B. Wang, Tian, et al., 2012; B. Wang et al., 2009; B. Wang, Xu, et al., 2012), in two studies 1 min stages were applied (Miura et al., 1998; Racinais, Buchheit, & Girard, 2014), and one study used 4 min stages (Coquart et al., 2017). The beginning workload varied from one study to the other and increments were set between 25 and 50 W. Additionally, all subjects cycled with the NIRS placed on the m. vastus lateralis, except for in one study, in which the subjects cycled with two NIRS devices, namely one on the m. vastus lateralis and one on the m. gastrocnemius lateralis (B. Wang, Xu, et al., 2012). NdAT was attempted to be determined in OI in five cases (Van Der Zwaard et al., 2016; B. Wang, Tian, et al., 2012; B. Wang et al., 2009; B. Wang, Xu, et al., 2012), in tissue saturation index (TSI) in a single study (Raleigh et al., 2018), in $\Delta[O_2Hb]$ in two cases (Coquart et al., 2017; Miura et al., 1998) and one study used $\Delta[HHb]$ and $\Delta[O_2Hb]$ (Racinais et al., 2014).

Four studies revealed that there is no significant difference between NdAT and AT or NdAT and LT (Coquart et al., 2017; Miura et al., 1998; Van Der Zwaard et al., 2016; B. Wang et al., 2009), two studies showed a significant correlation between NdAT and AT or NdAT and LT (Grassi et al., 1999; B. Wang, Xu, et al., 2012), and three studies established that NdAT is significantly different than AT (Racinais et al., 2014; Raleigh et al., 2018; B. Wang, Tian, et al., 2012).

In short, all studies used a cycling GXT and determined NdAT in either OI, TSI or $\Delta[HHb]$ and $\Delta[O_2Hb]$. Results varied from finding significant differences to establishing no significant difference and high correlation between NdAT and AT.

8.3.1. NdAT in OI

Certain studies (e.g. Grassi et al., 1999; Wang, Tian, et al., 2012) state the importance of a constant $\Delta[tHb]$ during the GXT to be able to establish NdAT. $\Delta[tHb]$ is an indicator of blood volume changes (B. Wang, Tian, et al., 2012). A general pattern can be recognized in $\Delta[tHb]$ during a GXT. During lower intensities, $\Delta[tHb]$ increases slightly, but not significantly, and at higher work intensities, it either stays unchanged or decreases slightly (Figure 7).
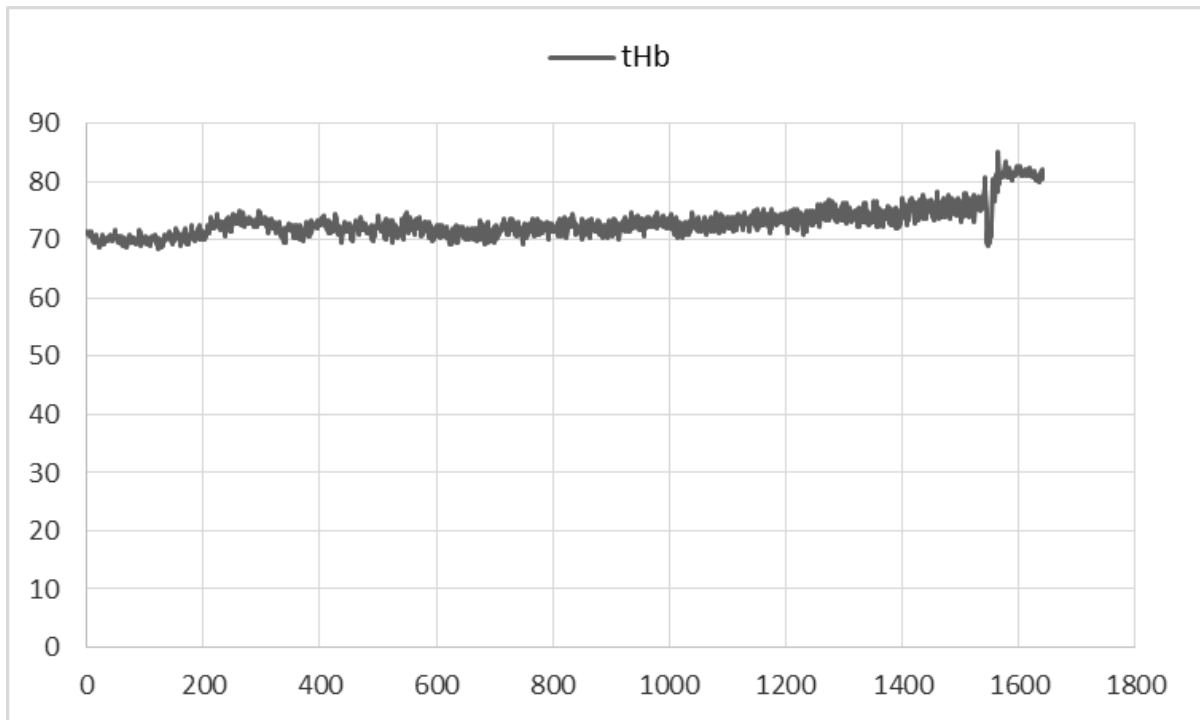
Figure 7. Graph of Δ[tHb] of subject 4 to determine constancy. Notes: The grey line is the graph of the Δ[tHb] data during the entire GXT, including 180s baseline.

Also in OI a general pattern can be seen, the main finding being that at NdAT, OI quickly declines (Grassi et al., 1999; B. Wang, Tian, et al., 2012; B. Wang et al., 2009) (Figure 8).
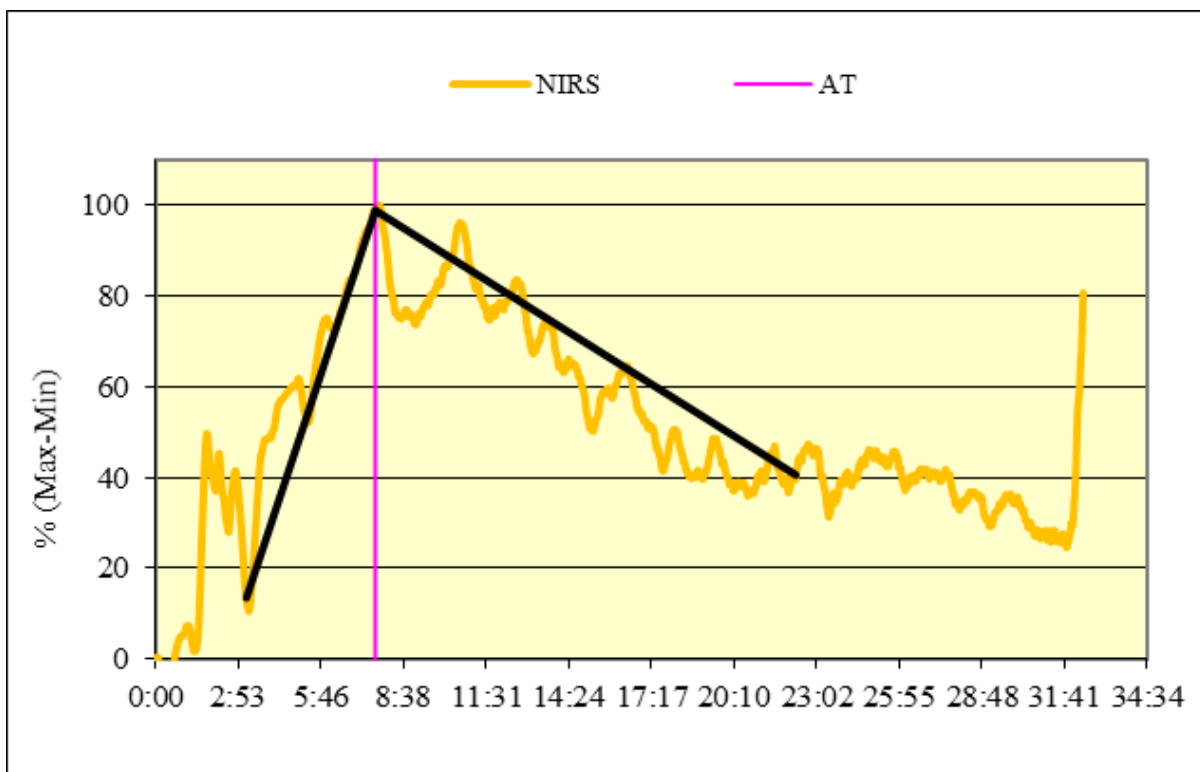

Figure 8. Graph of OI of subject 2 to determine NdAT. Notes: The yellow line is the smoothed OI data, the pink line indicates the NdAT, and the black lines are the manually places regression lines.

In the studies, where OI was used to determine NdAT, four of them paid attention to a constant $\Delta[tHb]$ (Grassi et al., 1999; B. Wang, Tian, et al., 2012; B. Wang et al., 2009; B. Wang, Xu, et al., 2012). Van der Zwaard et al. (2016), establish NdAT in OI, though they do not mention the fact, that a constant $\Delta[tHb]$ is required for this method to be valid. However, as they came up with no difference between NdAT and AT ($P > 0.05$), either $\Delta[tHb]$ was constant without their knowledge, or the method was successful, even without a constant $\Delta[tHb]$. Grassi et al. (1999) clearly state that $\Delta[tHb]$ needs to be constant and also describe the $\Delta[tHb]$ curve precisely. During intensities up to about 60-65% of the maximal work load, $\Delta[tHb]$ increased slightly, before staying unchanged. Therefore, the point at which NdAT was found in OI occurred during constant $\Delta[tHb]$. Grassi et al. (1999) did not perform a paired sample t-test, but they resulted in a high correlation between NdAT and AT ($r = 0.95$; $P = 0.005$). Wang, Xu, et al. (2012) also mentioned the importance of a constant $\Delta[tHb]$ to determine NdAT in OI. They observed differences in NdAT established in the m. vastus lateralis and the m. gastrocnemius lateralis, with the m. vastus lateralis showing the breakpoint sooner ($P < 0.001$). There were significant correlations between the breakpoint in m. gastrocnemius lateralis and the lactate as well as the ventilatory threshold. Wang et al. (2009) also determined NdAT in OI with a constant $\Delta[tHb]$, and revealed no significant difference between NdAT and LT, whereas Wang, Tian, et al. (2012) result in significant differences between NdAT and AT in OI in the presence of a constant $\Delta[tHb]$.

8.3.2. NdAT in $\Delta[O_2Hb]$

The common graph of $\Delta[O_2Hb]$ during a GXT is constant or begins to decrease slightly at onset of work load, at NdAT it declines and at NdRCP it declines more rapidly (Coquart et al., 2017; Miura et al., 1998). Coquart et al. (2017) and Miura et al. (1998) found no significant difference between NdAT and AT (with $P < 0.05$). However, Racinais et al. (2014) were not able to locate a breakpoint in $\Delta[O_2Hb]$ for NdAT, the only infliction point they found corresponded to RCP.

8.3.3. NdAT in TSI

Tissue saturation index describes the ration between $\Delta[O_2Hb]$ and $\Delta[tHb]$. According to Raleigh et al. (2018), during a GXT this index declines before flattening out, at which point the NdAT can be set. However, while NdAT in TSI corresponded to LT, both thresholds were significantly different than AT.

8.4.   Summary of Literature Defining NdRCP

Of the studies, which were included in this overview for NdRCP establishment, in five an incremental cycling test was performed (Bellotti et al., 2013; Fontana et al., 2015; Keir et al., 2015; Miura et al., 1998; Racinais et al., 2014) and in two the subjects participated in an incremental running test (Karatzanos et al., 2010; Snyder & Parmenter, 2009). The incremental cycling tests in all the studies were performed with 1 min stages, with increments anywhere between 10 and 30 W. All subjects, which carried out the cycling test, had the NIRS applied to the m. vastus lateralis. In the two studies which included an incremental running test, the NIRS was applied to the m. gastrocnemius lateralis (Karatzanos et al., 2010; Snyder & Parmenter, 2009). Stage length and increments varied between and within the two studies.

NdRCP was attempted to be established in $\Delta[HHb]$ in four studies (Bellotti et al., 2013; Fontana et al., 2015; Keir et al., 2015; Racinais et al., 2014), one study used $\Delta[HHb]$ and $\Delta[O_2Hb]$ (Miura et al., 1998), and in two cases TSI was used (Karatzanos et al., 2010; Snyder & Parmenter, 2009). All studies concluded there is no difference between NdRCP and RCP, however one study resulted in a significant difference when the threshold was calculated, but not when it was visually determined (Karatzanos et al., 2010).

In short, two studies conducted running GXTs while the others used cycling GXTs. NdRCP was determined in $\Delta[HHb]$ and TSI. All studies resulted in no significant difference between NdRCP and RCP.

8.4.1. NdRCP in $\Delta[HHb]$

The common trend of $\Delta[HHb]$ during a GXT shows a very slight increase in the warm-up phase, before a linear incline during the incremental work load phase. At high-intensity workloads, the curve has a breakpoint, at which $\Delta[HHb]$ increase slows down or even plateaus. At this infliction point, NdRCP can be set (Figure 9).

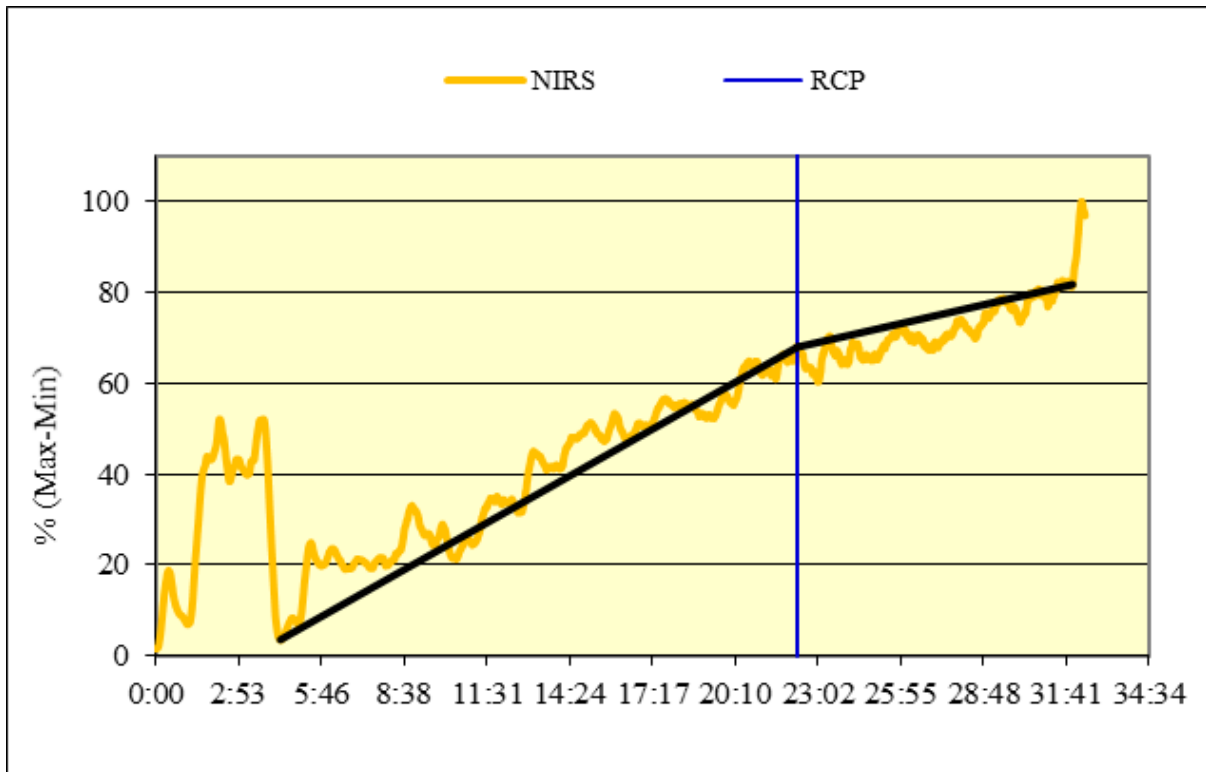*Figure 9. Graph of Δ[HHb]of subject 2 to determine RCP. The yellow line is Δ[HHb], the black lines are the manually placed regression lines and the vertical blue line is the RCP.*

All studies using Δ[HHb] to determine NdRCP find no difference between NdRCP and RCP (Bellotti et al., 2013; Fontana et al., 2015; Keir et al., 2015; Racinais et al., 2014). As mentioned above, Miura et al (1998) attempted to find NdAT and NdRCP in Δ[O₂Hb], and found that the breakpoint coincided with NdRCP.

8.4.2. NdRCP in TSI

As mentioned before, Raleigh et al. (2018) described the TSI curve as a decline with a subsequent plateau. Karatzanos et al. (2010) however, observed the TSI trend differently, namely either as an inverted U-shape or a flatter pattern, where TSI decreases slightly and then declines quicker at higher workloads. Also Snyder and Parmenter (2009) showed that the shape of the TSI curve varied between subjects. Both Karatzanos et al. (2010) and Snyder and Parmenter (2009) were able to determine NdRCP which was not significantly different than RCP. However, Snyder and Parmenter (2009) did result in significant differences when NdRCP was calculated versus visually determined.

8.5.   Outlook on the Empirical Study

When looking at the literature on the topic of determining thresholds using NIRS, there is already a lot of knowledge which has been obtained. However, certain areas are not as well scoped out as others. A clear statement can be made, that NdRCP is determinable during an incremental cycling test, when analyzing Δ[HHb]. Enough studies have proven this to be the case, also indicating that Δ[HHb] is the best parameter to use when doing so (Wang et al., 2006). However, this field is lacking studies on NdRCP determined in incremental running tests. It is important that performance assessment is executed in the athlete's specific sport, meaning runners should perform a GXT while running and cyclist should perform a GXT while cycling (Eston & Reilly, 1996). For this reason it is necessary to establish whether or not NIRS derived thresholds determined during a running incremental test are also valid, so that this method can be used among athletes and their coaches as a surrogate for invasive testing methods.

In addition, there seems to be little consensus on the method of determination of NdAT within the literature. Also, the outcomes of these studies vary, indicating the need for further research in this area. As with NdRCP, studies performing an incremental running test are lacking and therefore, there is need for studies to determine NdAT during a running GXT.

For this reason, the following study was carried out with subjects performing an incremental running test with the NIRS placed on the m. gastrocnemius lateralis, and NdAT and NdRCP were attempted to be determined.

9. Methods

9.1. Subjects

Twenty healthy recreationally active sport students (m = 14, w = 6; age: 22.4 ± 4.9 years; body mass: 67.3 ± 12.4 kg; height: 1.73 ± 0.12 m) volunteered to participate in this study. Each subject was informed of all experimental procedures prior to the study, was notified of the risks that would be taken when participating, and, after having filled out a health questionnaire, gave their written informed consent. All procedures were conducted in accordance with the declaration of Helsinki (1976) and were approved by the Ethics committee of the University of Vienna (Reference Number: 00325).

9.2. Study Design

The subjects completed two incremental graded exercise tests (GXT) on separate days. Each test was performed at least 48 h and at the most 7 days apart from each other, at similar times of the day (±1 h) in order to avoid effects of residual fatigue, training or detraining, or effects of the diurnal rhythm. The GXTs were performed on a treadmill (Saturn, h/p/cosmos Sport and Medical, Traunstein, Germany). Before the first GXT, subjects were weighed and measured (M877 (Seca Beam Scale and Stadiometer, Birmingham, UK)).The participants were asked to avoid vigorous exercise and alcohol preceding the day of the GXTs and were instructed to have eaten their last meal three hours before the GXTs and arrive in the laboratory well hydrated. The subjects were also advised to consume the same food before both tests. Furthermore, the participants were asked to refrain from caffeine and sports beverages within three hours of the GXTs.

9.3. Equipment

MetaLyzer 3B

The MetaLyzer 3B (Cortex Biophysik GmbH., Leipzig, Germany) is a stationary breath by breath metabolic gas analyzer. 30 min prior to using it, the system was turned on and once it was warmed up it was calibrated according to the manufacturer's manual. Flow and volume were calibrated using a 3,000 mL calibration syringe once a day and gas analyzers were calibrated once a month with known gas concentrations (5% CO_2, 15% O_2, Cortex Biophysik GmbH, Leipzig, Germany). A two meter long sampling tube was used to measure O_2

(electrochemical cell) and CO_2 (infrared analyzer). The software used is MetaSoft® Studio 5.8.2 (Cortex Biophysik GmbH., Leipzig, Germany).

Master Screen CPX

The Master Screen CPX (Viasys Healthcare, Höchberg, Germany) is a stationary breath by breath metabolic gas analyzer in form of a metabolic cart. 30 min prior to using it, the system was turned on and when warm, calibrated according to the manufacturer's manual. Flow, volume, and gas analyzers were calibrated automatically (5.03% O_2, 15.94% CO_2) prior to each test. O_2 and CO_2 were measured by the differential paramagnetic principle and by infrared absorption respectively. The Software used is LabManager (version 5.32.0.5; CareFusion Germany 234 GmbH, Höchberg, Germany).

Near Infrared Spectroscopy

The PortaMon (Artinis Medical Systems, Elst, Netherlands) is a wireless and portable continuous wave NIRS system with two wavelengths (760 and 850 nm) and a sampling rate of 10 Hz. The NIRS uses the modified Lambert-Beer Law and spatially resolved spectroscopy. It is 83.8 x 42.9 x 17.2 mm (WxHxD) in size and 88 g in weight (including battery). This device is small and light enough to not hinder the athlete while running when applied to the muscle. The NIRS measures changes in oxyhemoglobin ($\Delta[O_2Hb]$), deoxyhemoglobin ($\Delta[HHb]$), total hemoglobin ($\Delta[tHb]$) and regional tissue saturation index at capillary level. The distances between the receiver and the three transmitters are 30, 35 and 40 mm, which results in a measurement depth of up to about 15 mm. The software used is Oxysoft 3.0.95 (Artinis Medical Systems, Elst, Netherlands).

9.3.1. Equipment Application

The mask of the respective metabolic gas analyzer (Hans Rudolph Inc, Shawnee, Kansas, USA) was attached to the head covering mouth and nose a couple of minutes before the subjects began running, so that they were able to get used to the feeling. Mask size was chosen by testing air tightness. If the GXT was performed with the ML 3B, the NIRS was applied. The NIRS was first wrapped in plastic wrap to keep the device dry. Then it was placed on the muscle belly of the lateral gastrocnemius of the right leg, parallel to the longitudinal axis of the lower leg. The NIRS was held in place by a strip of adhesive tape, then covered with two pieces of black medical tape to prevent light from intruding and was finally wrapped with an elastic bandaging

to prevent any slipping of the device. Once all devices were applied, the subject stepped on to the treadmill and the turbine was inserted into the mask.

9.4. Protocol

Pulmonary ventilation and gas exchange were measured continuously breath by breath during all tests. In a randomized order, the subjects ran one of the GXTs with the ML 3B and the other with the MS CPX. Additionally, during the GXT with the ML 3B, $\Delta[O_2Hb]$, $\Delta[HHb]$, and $\Delta[tHb]$ were continuously obtained via NIRS.

After a three-minute warm-up of walking at 1.39 $m \cdot s^{-1}$, the test commenced at either 1.67 $m \cdot s^{-1}$ or 2.22 $m \cdot s^{-1}$ depending on the self-reported fitness level of the subjects, and the speed was increased by 0.28 $m \cdot s^{-1}$ every three minutes until volitional exhaustion. The beginning speed was chosen enabling the subjects to complete a minimum of five stages, which resulted in the GXT lasting at least 15 min (Midgley, Bentley, Luttikholt, Mcnaughton, & Millet, 2008). The treadmill incline was set to 1%. The runners were strongly verbally encouraged to run until exhaustion. At this point the treadmill was stopped, the mask was removed, and the participants then continued walking at 1.39 $m \cdot s^{-1}$ until the individual desired to quit.

9.5. Data Analyses

Obtained parameters which are of interest for answering the research questions include $\dot{V}O_2$, $\dot{V}O_{2peak}$, $\dot{V}CO_2$, $\dot{V}E$, $\Delta[O_2Hb]$, $\Delta[HHb]$, and $\Delta[tHb]$, as well as maximal running speed (Vmax). Of these parameters, further parameters were calculated, including EqO_2, $EqCO_2$, $PETO_2$, $PETCO_2$ (metabolic gas analyzer) and oxygenation index (OI) (NIRS).

Vmax was compared between the two tests. If the last stage was not completed, maximal running speed was calculated using the following equation (Kuipers, Rietjens, Verstappen, Schoenmakers, & Hofman, 2003):

$$V_{max} = V_{completed} + t/T \times \text{speed increment} \tag{1}$$

where V_{max} is the maximal running speed, $V_{completed}$ is the running speed of the last completed stage ($m \cdot s^{-1}$), t is the duration of the uncompleted stage (s), T is the length of the stage (s), and speed increment is the increase in speed between each stage ($m \cdot s^{-1}$).

In order to determine the NIRS derived anaerobic threshold (NdAT), the difference between $\Delta[O_2Hb]$ and $\Delta[HHb]$ was calculated to create the OI. First, the OI data was smoothed by calculating the moving average with an interval of $\pm$ 10 data points using a Microsoft Excel® spreadsheet. This dataset was then relativized using the following equation:

$$Sr = [(\text{instantaneous value})-(\text{mean baseline})] / [(\text{array maximum})-(\text{mean baseline})] \qquad (2)$$

with Sr being the smoothed relative data, instantaneous value being the value given at a certain time of the smoothed data, mean baseline being the smoothed mean of the data acquired before the GXT started, and array maximum being the largest value of the smoothed data. This dataset was then plotted against time (Grassi et al., 2018; Van Der Zwaard et al., 2016; Wang, Tian, & Zhang, 2012; Wang, Xu, et al., 2012).

A breakpoint in the graph was visually determined with the help of manually placed regression lines. According to Wang, Xu, et al. (2012) NdAT can only be determined if $\Delta[tHb]$ is constant. A decrease in OI with a simultaneous non-decreasing $\Delta[tHb]$ indicates deoxygenation (B. Wang, Xu, et al., 2012). In order to establish constancy, two methods were used the analyze the graph of $\Delta[tHb]$. Firstly, the whole graph of $\Delta[tHb]$ for the length of the GXT was visually inspected. Naturally, during and shortly after the baseline, $\Delta[tHb]$ increased. If it reached a plateau and stayed at that value, $\Delta[tHb]$ was considered constant.

Secondly, it was attempted to determine constancy more accurately. Based on the assumption, that $\Delta[tHb]$ needs to be constant only at the point of NdAT and not necessarily during the entire GXT, an interval was created in which $\Delta[tHb]$ was inspected. This interval was based on the error of measurement of the ML 3B, which is considered to be 150 ml·min^{-1}. The 150 ml·min^{-1} were transferred into seconds, by calculating $\dot{V}O_2$ in 1 s and then multiplying this by 150, which is the error of measurement. This resulted in 259 s. $\Delta[tHb]$ was then visually analyzed in the interval NdAT $\pm$ 259 s, which is an interval of 519 s.

Both methods were applied separately, and then compared with each other. If the methods did not result in two similar answers (meaning the one method showed $\Delta[tHb]$ constancy and the other method determined $\Delta[tHb]$ was not constant) both graphs were looked at side by side and a decision was made.

If any graphs result in Δ[tHb] not being constant, it was still attempted to determine the NdAT, however this will be critically discussed, as this has been done in another study (Grassi et al., 1999).

When detecting the NIRS derived respiratory compensation point (NdRCP), the data of Δ[HHb] was analyzed, as this has been proven to be the most accurate parameter to detect NdRCP (Wang et al., 2006). The data was smoothed and relativized using the same method mentioned above and in equation 2 and then plotted against time. A breakpoint in the graph was visually determined with the help of manually placed regression lines.

Data from the metabolic gas analyzers were recorded breath by breath and the mean of 30-second intervals was plotted. AT and RCP were determined by visual inspection of $\dot{V}O_2$, $\dot{V}CO_2$, $\dot{V}E$, RER, EqO_2 and $EqCO_2$, as well as $PETO_2$ and $PETCO_2$. AT was set at the point where EqO_2 begins to rise, after beforehand declining or being flat, while at the same time $EqCO_2$ decreases. Also, the first breakpoint in the $\dot{V}E$ curve was used to establish AT. RCP was determined at the point where $EqCO_2$ began to increase over proportionally, as well as the second breakpoint in the $\dot{V}E$ curve. In addition, the point at which RER reaches 1 was taken into consideration (Beaver et al., 1986; Wasserman, 1984; Wasserman, Whipp, Koyal, & Beaver, 1973).

9.6.   Statistical Analyses

All data were checked for normality by analyzing the histogram, the boxplot, the Q-Q diagram and the Shapiro-Wilk test. Differences in measures of Vmax, $\dot{V}O_{2peak}$, and $\dot{V}O_2$, $\dot{V}CO_2$, and $\dot{V}E$ at each stage, as well as NdAT, AT, NdRCP and RCP were assessed with the paired sampled t-test, as all data were distributed normally. Agreement between the two tests were analyzed via Bland-Altman method with the limits of agreement set at 95% (Bland & Altman, 1968) and the correlation was calculated using the Pearson correlation coefficient. Standard error of estimate was calculated and the coefficient of variation of the log transformed data was obtained (Hopkins, Schabort, & Hawley, 2001). Level of significance was set at $P < 0.05$. Statistical analysis were performed with IBM SPSS Statistics, Version 24 (IBM SPSS statistics, SPSS Inc. Chicago, USA).

10. Results

10.1.   Results of ML 3B and MS CPX

Mean V_{max} and mean $\dot{V}O_{2peak}$ of the GXT for each device is presented in Table 3. There were no significant differences in V_{max} between the two tests (P = 0.919), however, there was a significant difference in $\dot{V}O_{2peak}$ between the two devices (P < 0.001), with the $\dot{V}O_{2peak}$ measured by ML 3B being higher by 236 mL·min^{-1}.

Table 3. Mean maximal running speed and mean $\dot{V}O_{2peak}$ in the two tests and metabolic gas analyzers.

V_{max} (m·s^{-1})		$\dot{V}O_{2peak}$ (mL·min^{-1})	
ML 3B	MS CPX	ML 3B	MS CPX
3.90 ± 0.45	3.90 ± 0.42	3,482 ± 904	3,246 ± 789*

*Notes. V_{max} = maximal running speed; $\dot{V}O_{2peak}$= maximal oxygen uptake; ML 3B = MetaLyzer 3B; MS CPX = MasterScreen CPX; *significantly different at P < 0.001*

Mean $\dot{V}O_2$, $\dot{V}CO_2$, and $\dot{V}E$ for each device at each stage is shown in Table 4.

Table 4. Mean $\dot{V}O_2$, $\dot{V}CO_2$ and $\dot{V}E$ of the metabolic gas analyzers in mL·min^{-1}. Data is presented in mL·min^{-1}.

n	m·s^{-1}	$\dot{V}O_2$ ML 3B	$\dot{V}O_2$ MS CPX	$\dot{V}CO_2$ ML 3B	$\dot{V}CO_2$ MS CPX	$\dot{V}E$ ML 3B	$\dot{V}E$ MS CPX
6	**1.67**	1,392.5	1,274.3*	1,186.5	1,076.7*	35.6	32.9*
6	**1.94**	1,613.9	1,485.9*	1,435.7	1,333.7*	41.3	39.4
20	**2.22**	2,115.6	1,964.3*	1,865.5	1,735.1	50.0	48.8
20	**2.5**	2,411.2	2,207.4**	2,230.4	2,070.1**	59.9	57.7
20	**2.78**	2,654.0	2,436.9**	2,507.0	2,350.3*	68.2	66.9
19	**3.06**	2,887.4	2,639.3**	2,785.8	2,604.5*	77.3	76.3
19	**3.33**	3,075.0	2,864.9**	3,069.1	2,908.0*	89.4	87.8
15	**3.61**	3,382.7	3,180.7**	3,429.5	3,326.5	105.4	107.0
10	**3.89**	3,688.8	3,466.3**	3,839.2	3,631.9	122.8	154.2
4	**4.17**	4,119.9	3,782.5**	4,242.4	4,000.7*	125.1	722.2
3	**4.44**	4,368.8	4,118.8[#]	4,634.9	4,459.0[#]	150.2	145.3[#]
1	**4.72**	5,512.2	5,136.0[#]	6,035.9	5,851.8[#]	194.7	200.0[#]

*Notes. n = number of subjects; $\dot{V}O_2$= oxygen uptake; $\dot{V}CO_2$ = Carbon dioxide output; $\dot{V}E$ = minute ventilation; ML 3B = MetaLyzer 3B; MS CPX = MasterScreen CPX; *significantly different than values of ML 3B at P < 0.05; **significantly different than values of ML 3B at P < 0.001; [#]no statistical test performed due to number of subjects.*

P-values, Pearson correlation coefficient, and degrees of freedom (t-values) for each increment for $\dot{V}O_2$, $\dot{V}CO_2$, $\dot{V}E$ are presented in Table 5, Table 6, and Table 7, respectively.

Table 5. P-values, correlation (r) and degrees of freedom (t) presented for each stage for $\dot{V}O_2$.

n	m·s⁻¹	P	r	t_5	t_{19}	t_{18}	t_{14}	t_9	t_3
6	**1.67**	0.022	0.680	3.266					
6	**1.94**	0.002	0.904	5.974					
20	**2.22**	0.035	0.828		2.273				
20	**2.5**	<0.001	0.968		6.249				
20	**2.78**	<0.001	0.961		5.732				
19	**3.06**	<0.001	0.966			5.906			
19	**3.33**	<0.001	0.976			6.610			
15	**3.61**	<0.001	0.972				4.993		
10	**3.89**	<0.001	0.980					6.598	
4	**4.17**	0.026	0.961						4.129

Table 6. P-values, correlation (r) and degrees of freedom (t) presented for each stage for $\dot{V}CO_2$.

n	m·s⁻¹	P	r	t_5	t_{19}	t_{18}	t_{14}	t_9	t_3
6	**1.67**	0.01	0.802	4.079					
6	**1.94**	0.009	0.914	4.129					
20	**2.22**	0.063	0.712		1.975				
20	**2.5**	<0.001	0.943		4.359				
20	**2.78**	0.002	0.926		3.495				
19	**3.06**	0.001	0.938			3.882			
19	**3.33**	0.001	0.949			3.802			
15	**3.61**	0.062	0.923				2.031		
10	**3.89**	0.25	0.928					2.667	
4	**4.17**	0.03	0.965						3.912

Table 7. P-values, correlation (r) and degrees of freedom (t) presented for each stage for $\dot{V}E$.

n	m·s⁻¹	P	r	t_5	t_{19}	t_{18}	t_{14}	t_9	t_3
6	**1.67**	0.001	0.919	6.8					
6	**1.94**	0.118	0.797	1.884					
20	**2.22**	0.471	0.509		0.735				
20	**2.5**	0.052	0.848		2.072				
20	**2.78**	0.403	0.793		0.885				
19	**3.06**	0.307	0.816			1.051			
19	**3.33**	0.406	0.776			0.85			
15	**3.61**	0.687	0.772				-4.411		
10	**3.89**	0.187	0.947					-0.927	
4	**4.17**	0.879	0.859						-0.166

In stages 2.5-3.89 m·s⁻¹ (i.e. 9-14 km·h⁻¹) the statistical significance of $\dot{V}O_2$ was P < 0.001, and in the other stages, P-values were between P = 0.002 and P = 0.035, indicating significant differences in measurement of $\dot{V}O_2$ between the devices in all stages, with mean $\dot{V}O_2$ measured by ML 3B being significantly higher in all stages compared to MS CPX. This results in an

overall significant difference in $\dot{V}O_2$ between the devices ($P < 0.001$), as well as a significant correlation (r = 0.982, $P < 0.001$). $\dot{V}CO_2$ demonstrated significant differences in all stages except for stages 2.22 $m \cdot s^{-1}$ ($P = 0.063$), 3.61 $m \cdot s^{-1}$ ($P = 0.062$), and 3.89 $m \cdot s^{-1}$ ($P = 0.250$). Correlation was was high (r ranges from r = 0.712 to r = 0.965. $\dot{V}E$ only demonstrated a significant difference in stage 1.67 $m \cdot s^{-1}$. Correlation was moderate to high (r ranges from r = 0.509 to r = 0.947). Figure 10 represents the correlation between $\dot{V}O_2$ of MS CPX and ML 3B. The formula of the regression line can be used as an equation to calculate $\dot{V}O_2$ from one device to the other and is as follows:

$$\dot{V}O_{2 \text{ MS CPX}} = (\dot{V}O_{2 \text{ ML 3B}} - 6.009) / 1.077 \tag{3}$$

with $\dot{V}O_{2 \text{ MS CPX}}$ being $\dot{V}O_2$ measured with MS CPX at any given speed and $\dot{V}O_{2 \text{ ML 3B}}$ being $\dot{V}O_2$ measured at the same speed with ML 3B.
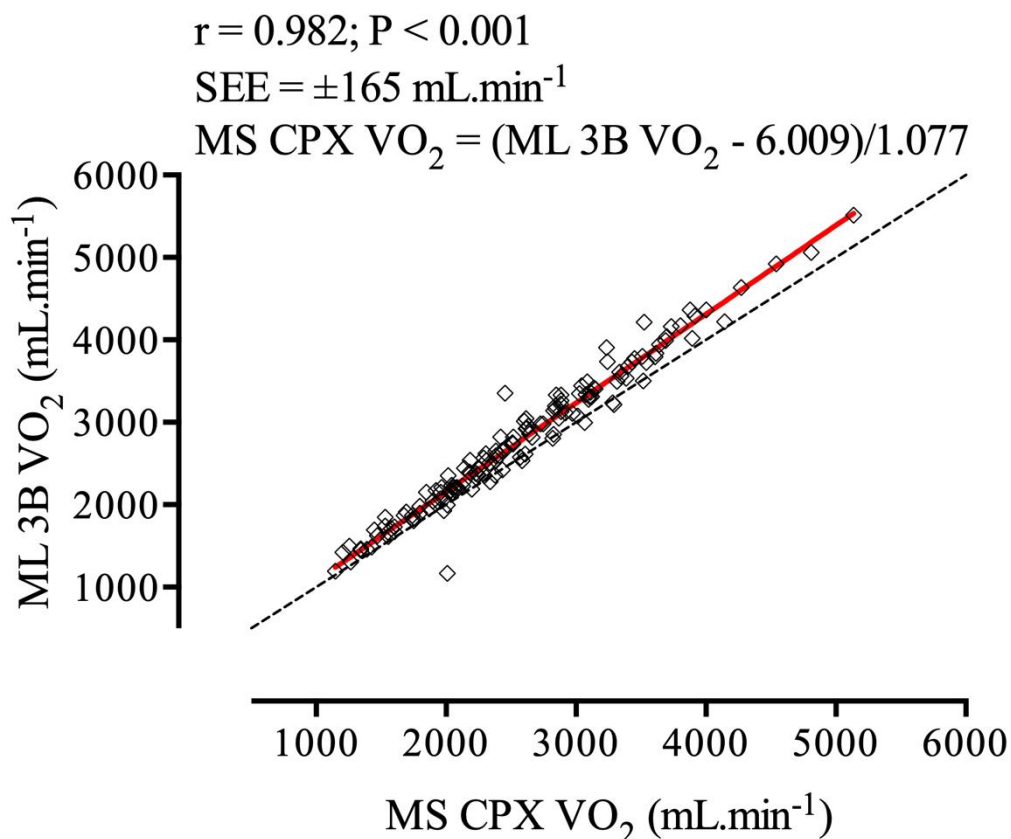
Figure 10. Correlation between $\dot{V}O_2$ of MS CPX and ML 3B. The squares represent $\dot{V}O_2$ of MS CPX and ML 3B for each subject, the red line is the regression line and the dotted black line is the line of identity.

Figure 11 shows the differences between the means of $\dot{V}O_2$ of MS CPX and ML 3B. The bias is -203 ± 176 mL·min^{-1} (95% LoA: -547 to 141 mL·min^{-1}).
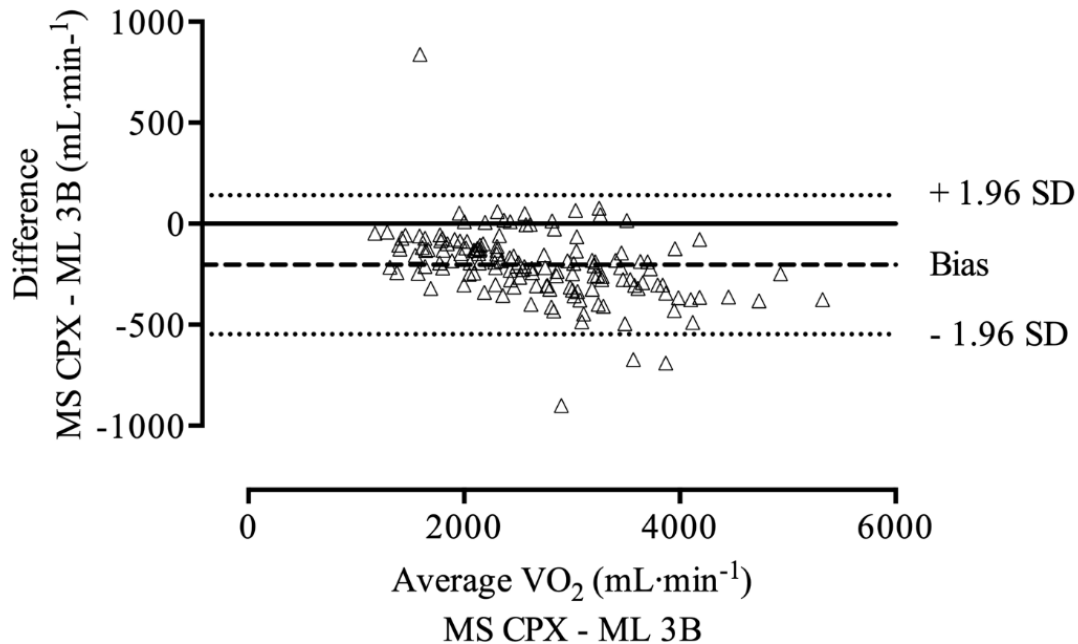
Figure 11. Bland-Altman plot of the difference between the means of $\dot{V}O_2$ of MS CPX and ML 3B. The triangles are the difference of the means of $\dot{V}O_2$ of MS CPX and ML 3B. The middle fat dotted line represents the mean bias, and the two thin dotted black lines represent the 95% limits of agreement.

Tables 8, 9, and 10 show the coefficient of variation (CoV) and the standard error of the estimate (SEE) in both raw units and standardized for each increment between the devices for $\dot{V}O_2$, $\dot{V}CO_2$, $\dot{V}E$, respectively.

Table 8. Coefficient of variation (CoV) and standard error of the estimate (SEE) in raw units and standardized presented for each stage for $\dot{V}O_2$.

n	m·s^{-1}	CoV (%)	SEE (raw units)	SEE (standardized)
6	1.67	6.4	81.3	1.08
6	1.94	3.3	49.7	0.47
20	2.22	12.0	192.8	0.86
20	2.5	5.2	101.7	0.26
20	2.78	5.3	126.8	0.29
19	3.06	5.0	130.6	0.27
19	3.33	4.1	115.8	0.22
15	3.61	4.3	127.8	0.24
10	3.89	3.3	108.9	0.20
4	4.17	5.5	191.4	0.29

Table 9. Coefficient of variation (CoV) and standard error of the estimate (SEE) in raw units and standardized presented for each stage for $\dot{V}CO_2$.

n	m·s^{-1}	CoV (%)	SEE (raw units)	SEE (standardized)
6	1.67	4.2	46.3	0.74
6	1.94	4.3	54.8	0.44
20	2.22	13.4	195.6	0.99
20	2.5	6.3	128.7	0.35
20	2.78	7.6	170.4	0.41
19	3.06	6.9	171.1	0.37
19	3.33	5.8	161.8	0.33
15	3.61	7.1	216.4	0.42
10	3.89	6.2	208.6	0.40
4	4.17	5.2	188.4	0.27

Table 10. Coefficient of variation (CoV) and standard error of the estimate (SEE) in raw units and standardized presented for each stage for $\dot{V}E$.

n	m·s-1	CoV (%)	SEE (raw units)	SEE (standardized)
6	1.67	2.7	0.9	0.43
6	1.94	6.4	2.5	0.76
20	2.22	13.4	5.9	1.69
20	2.5	8.2	4.8	0.63
20	2.78	9.7	6.8	0.77
19	3.06	8.9	7.2	0.71
19	3.33	9.4	8.4	0.81
15	3.61	10.0	11.3	0.82
10	3.89	30.9	74.8	0.66
4	4.17	8.8	11.5	0.60

CoV of $\dot{V}O_2$ and $\dot{V}CO_2$ is low to moderate, indicating little variability (ranging from 3.3% to 12% and from 4.2% to 13.4% for $\dot{V}O_2$ and $\dot{V}CO_2$, respectively). Also the SEE of these parameters is considered low (ranging from 0.2% to 1.08% for $\dot{V}O_2$ and from 0.27% to 0.99% for $\dot{V}CO_2$). CoV of $\dot{V}E$ is moderate to high (ranging from 2.7% to 30.9%) and SEE is low (ranging from 0.43% to 1.69%).

10.2.   Results of NIRS Derived Thresholds and Ventilatory Thresholds

NIRS data recordings of two subjects failed, resulting in the analysis of 18 datasets. Of these 18 datasets, in three subjects NdAT could not be determined and in four subjects NdRCP could not be established. Therefore, NdAT was defined in 15 subjects and NdRCP was determined in 14 subjects. AT and RCP determined in the data of the metabolic gas analyzer were visible in all subjects.

Mean NdAT was measured at 526 ± 117 s, and mean AT was measured at 419 ± 142 s, resulting in a significant difference (P = 0.014), but a significant correlation (r = 0.555; P = 0.021). The typical error of the estimate is 104 s and the CoV is 24.6%. The Bland-Altman 95% LoA are ± 254 s, with a mean difference of -86 ± 130 s.

Δ[tHb] was considered constant in eight subjects and not constant in nine subjects. Hence, the NdAT data can be divided into two datasets, namely one with constant Δ[tHb] and one with non-constant Δ[tHb]. NdAT with constant Δ[tHb] (508 ± 87 s) compared to AT (407 ± 67 s) resulted in a significant difference (P = 0.039), and a non-significant correlation (r = 0.088; P = 0.836). The typical error of the estimate is 101 s and the CoV is 23.7%. The Bland-Altman 95% LoA are ± 221 s, with a mean difference of -101 ± 113 s. NdAT with non-constant Δ[tHb] (503 ± 140 s) compared to AT (470 ± 187 s) revealed no significant difference (P = 0.180) and a non-significant correlation (r = 0.651; P = 0.058). The typical error of the estimate is 157 s and the CoV is 45.6%. The Bland-Altman 95% LoA are ± 291 s, with a mean difference of 73 ± 148 s.

Mean NdRCP was recorded at 1010 ± 173 s and mean RCP was recorded at 974 ± 202 s. This revealed no significant differences (P = 0.789) and a significant correlation (r = 0.594; P = 0.015). The typical error of the estimate is 150 s and the CoV is 16.6%. The Bland-Altman 95% LoA are ± 359 s, with a mean difference of -13 ± 183 s.

11. Discussion

The aims of this study were to compare the ML 3B and the MS CPX, to come up with an equation to calculate from one device to the other, and to compare NIRS derived thresholds to metabolic thresholds. As it is not the goal of this study to identify which metabolic gas analyzer measures correctly and which system is inaccurate, statements indicating that one device produces higher or lower values, should not be seen in relation to reference values. Also, such statements should not favor one system over the other, they are merely meant to provide information.

11.1. ML 3B and MS CPX

The findings of this study show that ML 3B produces significantly higher $\dot{V}O_2$ values throughout the entire GXT, with the discrepancy at $\dot{V}O_{2peak}$ being 236 mL·min^{-1}. The mean difference of $\dot{V}O_2$ at each stage is presented in Table 11. In the present study, $\dot{V}CO_2$ was overestimated by MM 3B in some stages ($P < 0.05$), whereas $\dot{V}E$ was only significantly different at the first stage ($P < 0.05$).

In order to compare the two systems, participants performed two GXTs while running with either the one or the other metabolic gas analyzer. This method was also applied in other studies (e.g. Díaz et al., 2008; Jensen, Jørgensen, & Johansen, 2002; Vogler, Rice, & Gore, 2010). The reason for this was on the one hand the inability to simultaneously use both apparatus, as this would have required the connection of two turbines, which did not seem feasible for this study. On the other hand, consecutive measurements of both devices during one GXT was not optimal in this case. This would have required longer stages in order for the participants to reach a steady state and then remain at that stage while systems were interchanged. However, the goal was to gain data over a wide range of intensities, for which shorter stages are necessary. Furthermore, this would have also been impossible due to the fact that the adapter for the turbine to the mask is different for each device and in addition it would have been difficult to insert the turbine while having the subjects run. Other studies using this method had the participants cycle (e.g. Brandes, Klein, Ginsel, & Heitmann, 2015; Carter & Jeukendrup, 2002). Stage length was chosen to be 3 min to give the participants time to adapt to the given intensity and enter into a short steady state before workload is increased. A steady state is reached after about 1 to 2 min at lower stages. At higher intensities a steady state is not attained (Draper & Hodgson, 2008).

Table 11. Mean difference of $\dot{V}O_2$ measured by ML 3B and MS CPX at each stage in mL·min^{-1}.

n	m·s^{-1}	$\dot{V}O_2$ (ML 3B-MS CPX)
6	**1.67**	118.2
6	**1.94**	128.0
20	**2.22**	151.3
20	**2.5**	203.8
20	**2.78**	217.1
19	**3.06**	248.1
19	**3.33**	210.1
15	**3.61**	202.0
10	**3.89**	222.5
4	**4.17**	337.4
3	**4.44**	250.0
1	**4.72**	376.2

Notes. n = number of subjects; $\dot{V}O_2$ = oxygen uptake; ML 3B = MetaLyzer 3B; MS CPX = MasterScreen CPX

Figure 12 shows the difference of $\dot{V}O_2$ between ML 3B and MS CPX at each stage (r = 0.868). It can clearly be seen that the discrepancy between the two devices increases with higher intensity, i.e. higher $\dot{V}O_2$ values. This can also be observed in Figure 10, where the regression line increases its deviation from the line of identity at higher $\dot{V}O_2$ values. This trend can be recognized in the Bland-Altman plot (Figure 11) as well.
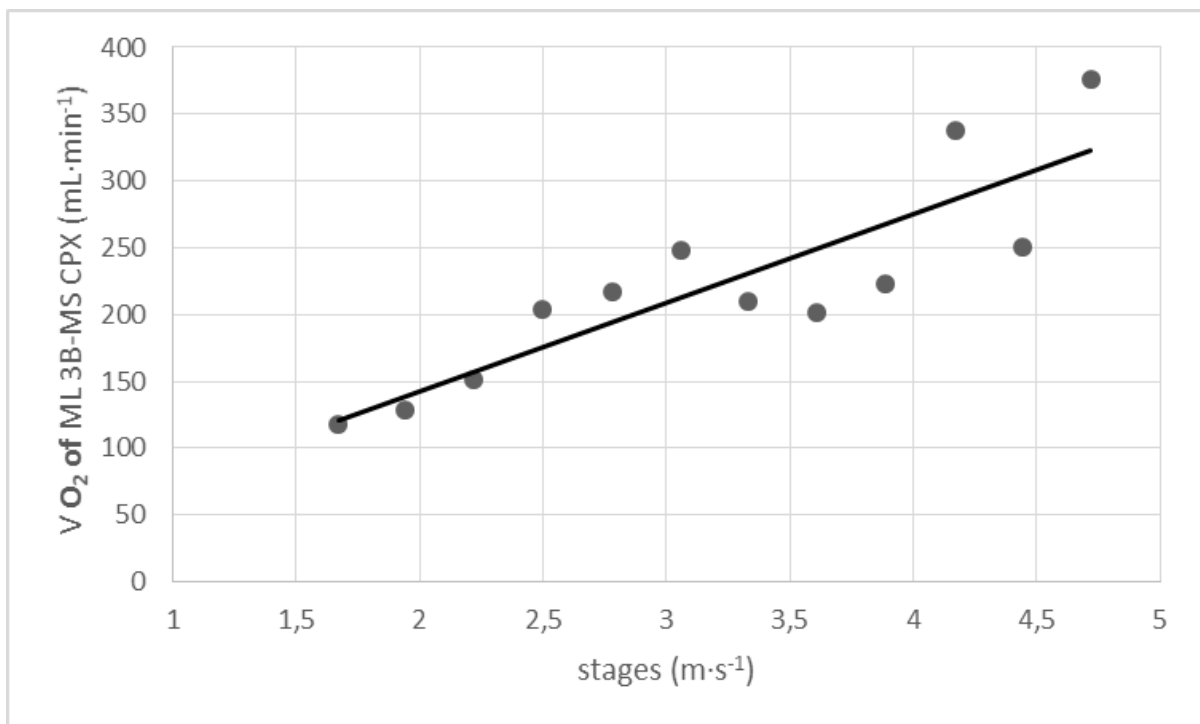
Figure 12. Difference of mean $\dot{V}O_2$ at each stage. The black dots are the mean $\dot{V}O_2$ at each stage and the black line is the regression line. $\dot{V}O_2$ = oxygen uptake; ML 3B = MetaLyzer 3B; MS CXP = MasterScreen CPX.

Since $\dot{V}O_2$ of ML 3B and MS CPX are highly correlated (r = 0.982, P < 0.001), which can be seen in Figure 10, it is possible to create an equation to calculate $\dot{V}O_2$ from one device to the other. The equation ($\dot{V}O_{2\,MS\,CPX} = (\dot{V}O_{2\,ML\,3B} - 6.009) / 1.077$) is the formula of the regression line in Figure 10. Using this equation, $\dot{V}O_2$ measured with the two devices can be made comparable. The standard error of the estimate (SEE) of the overall regression is 165mL·min^{-1}, which is interpreted as *small*.

This has been done in other studies as well. Diaz et al. (2008) made the Oxycon mobile and the Oxycon pro comparable through an equation ($\dot{V}O_{2\,(Oxycon\,pro)} = -508.639 + 1.281*\dot{V}O_{2\,(Oxycon\,mobile)}$), as their study revealed differences in the measurement of $\dot{V}O_2$ (P < 0.05). Also Duffield et al. (2004) came up with significant differences in $\dot{V}O_2$ (P < 0.05) and created an equation to correct the $\dot{V}O_2$ values of the K4b² system, which was validated against a metabolic cart ($\dot{V}O_{2\,(Metabolic\,cart)} = 0.926*(\dot{V}O_{2\,(K4b²)} - 0.227)$). No studies so far have compared MS CPX to a different system. Only one reliability study was found on ML 3B, indicating that the device is highly reliable with an intraclass correlation coefficient (ICC) of 0.984 (Meyer, Georg, Becker, & Kindermann, 2001). However, no validity study could be found using the ML 3B.

Therefore, the results of this study cannot be compared to results of other studies. Furthermore, statements cannot be made on the accuracy of both devices, as they have not been validated against the Douglas bag. However, studies have been conducted with systems developed by the same companies. MS CPX, Oxycon pro, Oxycon Alpha, and Oxycon mobile are all metabolic gas analyzers produced by Viasys Healthcare (Höchberg, Germany), formerly called Erich Jaeger GmbH. ML 3B, MetaMax 3B, MetaMax II, and MetaMax I are produced by Cortex Biophysik GmbH. (Leipzig, Germany).

For example, Rietjens et al. (2001) validated the Oxycon pro against the Douglas bag, and resulted in no significant difference of $\dot{V}O_2$, $\dot{V}CO_2$, or $\dot{V}E$ between the two devices (P > 0.05) and a significant correlation (r = 0.996, P < 0.001 for $\dot{V}E$; r = 0.957, P < 0.001 for $\dot{V}O_2$; and r = 0.980, P < 0.001 for $\dot{V}CO_2$). Foss and Hallén (2004) also compared the Oxycon pro to the Douglas bag. The Oxycon pro can be used in the mixing chamber mode or in the breath-by-breath mode (Macfarlane, 2001), and in this case the mixing chamber mode was used. This study resulted in a slightly lower estimation of $\dot{V}O_2$ by 0.8 % (P < 0.05). Perret and Mueller (2006) validated the Oxycon mobile against the Oxycon pro, based on the assumption that

Oxycon pro produces valid results as shown in Rietjens et al. (2001). The Oxycon mobile showed significantly lower $\dot{V}O_2$ values at 200 and 250 W during a cycling GXT ($P < 0.05$). No significant differences were found for $\dot{V}CO_2$. Also Diaz et al. (2008) compared the Oxycon mobile to the Oxycon pro as it has proven to be valid (Foss & Hallén, 2005; Rietjens et al., 2001). Likewise, Oxycon mobile measured significantly lower at 12 and 17 km·h^{-1} during a running GXT ($P < 0.05$). $\dot{V}CO_2$ and $\dot{V}E$ were not significantly different between the devices. Macfarlane & Wong (2012) validated the MetaMax 3B against the Douglas bag. They found significant differences in $\dot{V}O_2$ and $\dot{V}CO_2$ during moderate and vigorous exercise ($P < 0.01$), where MetaMax 3B overestimated these values by 10%. Also Vogler et al. (2010) validated the MetaMax 3B against the Douglas bag, showing that the MetaMax 3B overestimated $\dot{V}O_2$, $\dot{V}CO_2$, and $\dot{V}E$ by 4%, 7%, and 4% respectively ($P < 0.01$). When looking at these few studies the conclusion can be drawn that the Oxycon pro produces valid results when compared to the Douglas bag, while the Oxycon mobile seems to underestimate values and the MetaMax 3B seems to generally overestimates.

However, caution must be taken when analyzing these studies side by side to the data collected in the present study. Even though these systems were produced by the same manufacturer, they can hardly be compared to each other.

11.2.   NdAT

The findings of this study show that NdAT cannot be determined with NIRS validly ($P = 0.014$), with NdAT occurring 86s earlier than AT, with the typical error of the estimate and the CoV being interpreted as *very high* (104s, and 24.6%, respectively).

As mentioned before, various authors point out the importance of a constant Δ[tHb] during the GXT (e.g. Wang, Tian, & Zhang, 2012; Wang, Tian, Zhang, & Gong, 2009; Wang, Xu, et al., 2012). If Δ[tHb] is constant with an increase in Δ[HHb] and a decrease in Δ[O$_2$Hb], this is an indicator of true deoxygenation. If the GXT results in alterations of Δ[tHb], then this means there are changes in blood volume (B. Wang et al., 2009). OI is the difference between Δ[O$_2$Hb] and Δ[HHb], and an increase in Δ[tHb] in the course of the GXT can influence this parameter. As mentioned in an earlier chapter, in this study two methods were applied to evaluate Δ[tHb]. The first method, which was also applied in various other studies (Grassi et al., 1999; B. Wang,

Xu, et al., 2012) relied on visual inspection of Δ[tHb] throughout the entire GXT. An example of this for subject 4 is shown in Figure 13.
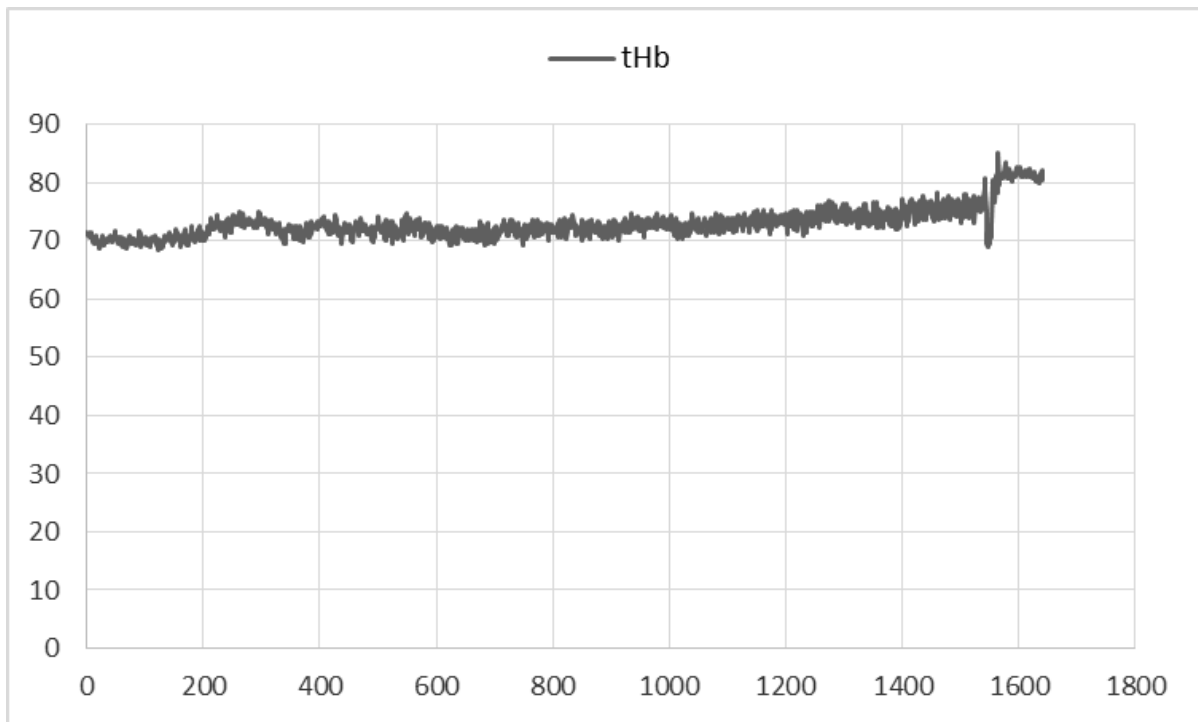
Figure 13. Graph of Δ[tHb] of subject 4 to determine constancy. Notes: The grey line is the graph of the Δ[tHb] data during the entire GXT, including 180s baseline.

For the second method, the assumption was made that Δ[tHb] only needs to be constant at the point of NdAT. Therefore, as described before, an interval was created in which Δ[tHb] was observed. Figure 14 shows Δ[tHb] in the interval of 529 s for subject 4.
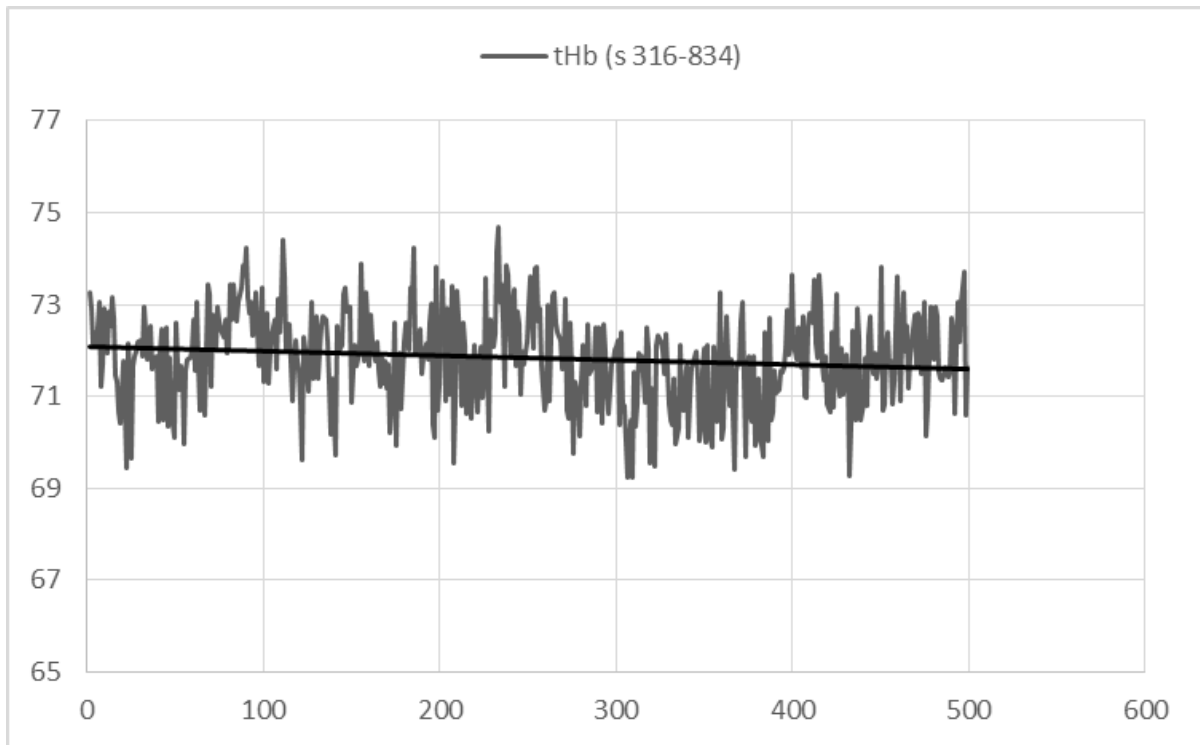
Figure 14. Graph of Δ[tHb] of subject 4 in the interval of 316-834 s to determine constancy. Notes: The grey line is the graph of the Δ[tHb] data during within in the 519-s interval, the black line is the regression line.

Since other studies (Van Der Zwaard et al., 2016; B. Wang et al., 2009) resulted in no difference between NdAT and AT when using OI, this parameter was taken to establish NdAT in the present study. Δ[tHb] was constant in 8 of 17 subjects in which a breakpoint in OI was found. Since the graph of OI was examined in relation to the baseline, the curves of OI as a function of time ended up looking very different between the subjects. It depended on whether or not the OI was positive or negative relative to the baseline and therefore, each curve appeared differently. An example of the OI curve of subject 4 is shown in Figure 15.
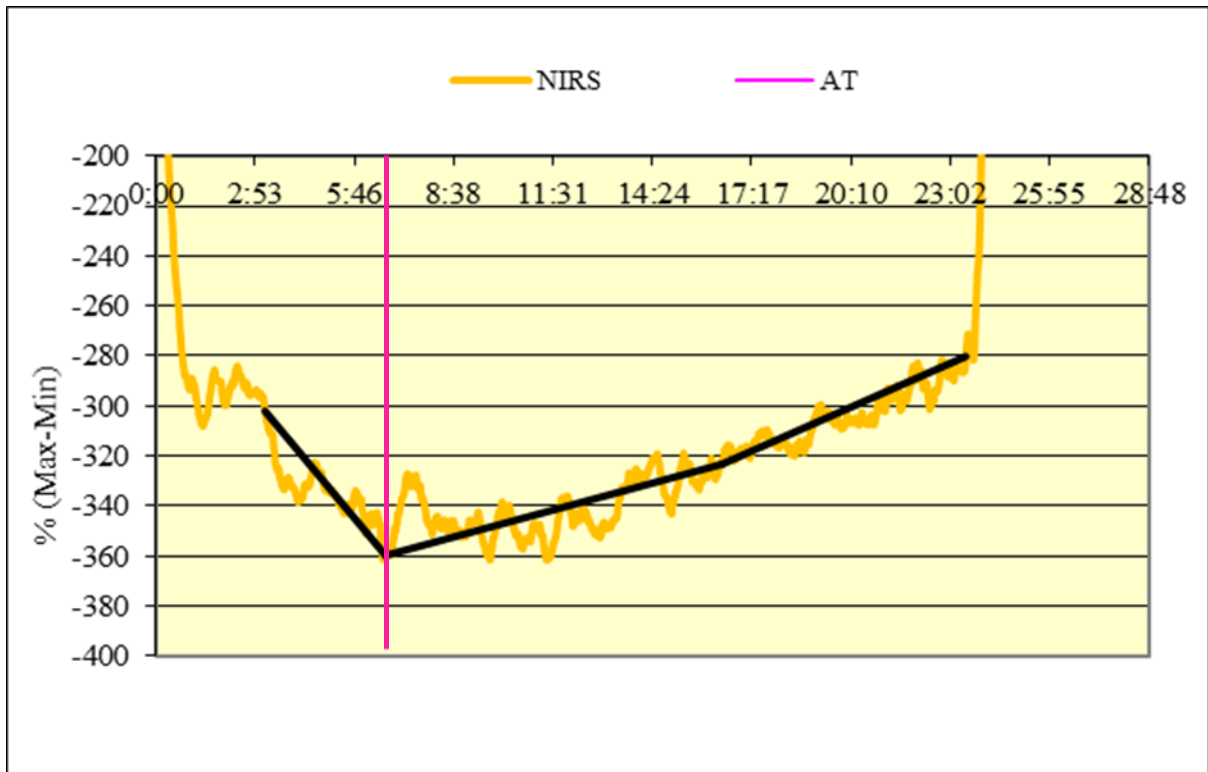
Figure 15. Graph of OI of subject 4 to determine NdAT. Notes: The yellow line is the smoothed OI data, the pink line indicates the NdAT, and the black lines are the manually places regression lines.

In contrast to Figure 15, OI of subject 2 in Figure 16 looks entirely different.
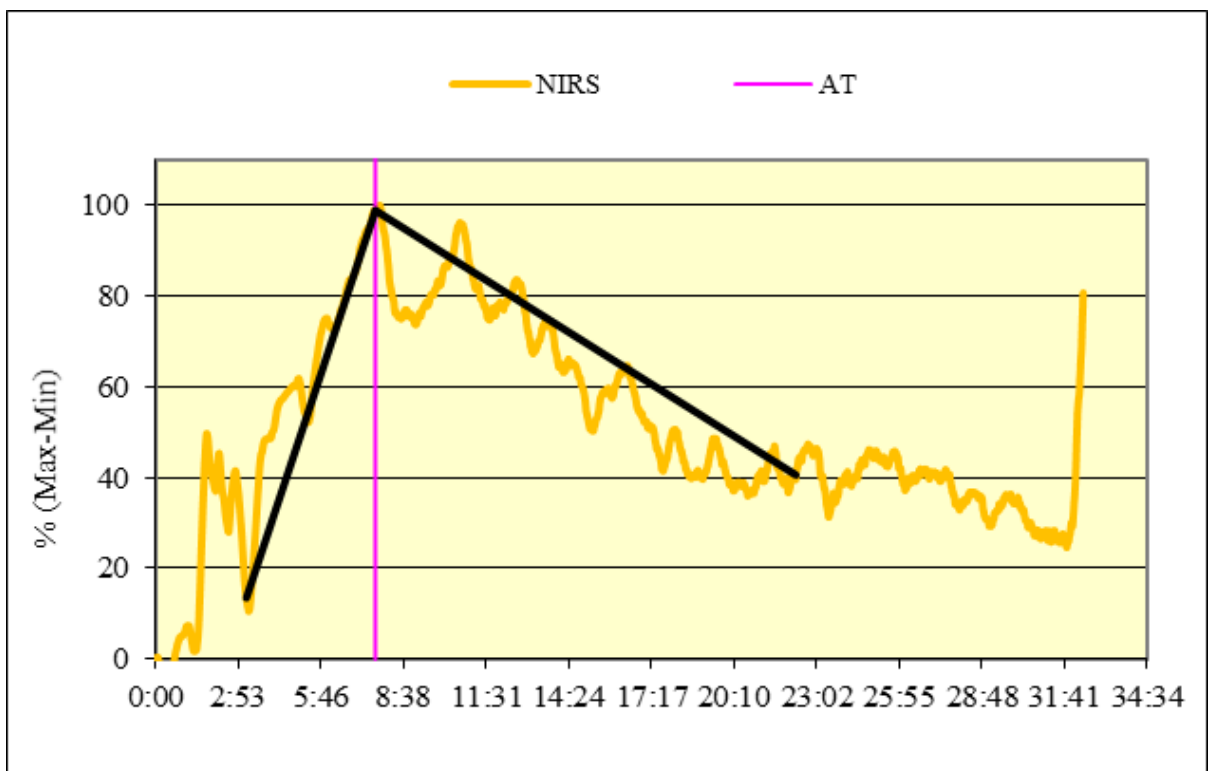
Figure 16. Graph of OI of subject 2 to determine NdAT. Notes: The yellow line is the smoothed OI data, the pink line indicates the NdAT, and the black lines are the manually places regression lines.

The present study resulted in a significant difference between NdAT and AT (P = 0.014) when analyzing combined datasets of both constant and non-constant Δ[tHb]. The mean difference between the two methods is 86 s, with AT measured earlier than NdAT. Also, the typical error of the estimate is interpreted as *very high* (104 s), as well as the CoV (24.6%).

The mean difference of NdAT and AT with constant Δ[tHb] is 101 s (P = 0.039). Also here, the typical error of the estimate is interpreted as *very high* (101 s) and the CoV is 23.7%, which is also considered very high. The mean difference of NdAT and AT with non-constant Δ[tHb] is 73 s, which is not significantly different (P = 0.18). But the typical error of the estimate is 157 s and CoV is 45.6%, which is again very high.

All these findings indicate that NIRS is not a valid method to determine AT, also when Δ[tHb] is constant. This is in consistence with Wang, Xu, et al. (2012), Wang, Tian, et al. (2012), and Racinais et al. (2014). Wang, Xu, et al. (2012) and Wang, Tian, et al. (2012) both attempted to determine NdAT in OI, with NdAT appearing earlier than AT in both cases. However, in the current study, NdAT appeared after AT. Even though Wang, Xu, et al. (2012) resulted in an earlier NdAT, they also concluded that the breakpoint of the m. gastrocnemius lateralis appeared significantly later than the breakpoint in the m. vastus lateralis. This could possibly be compared to the results of the present study, where measurements were done on the m. gastrocnemius lateralis. In contrast to Wang, Tian, et al (2012) and Wang, Xu, et al. (2012), Racinais at al. (2014) used Δ[O$_2$Hb] and Δ[HHb] to determine NdAT, but also resulted in a significant difference from AT. Raleigh et al. (2018) found no significant difference in NdAT and LT, however both thresholds were significantly different than AT. This was also the case in Wang, Tian, et al. (2012), where NdAT and LT were not significantly different, but both thresholds appeared earlier than AT.

In contrast to the present study, Coquart et el. (2017), Miura et al. (1998), van der Zwaard et al. (2016), and Wang et al. (2009) revealed that NIRS is a valid to establish NdAT. However, these studies were performed during a cycling GXT with the NIRS on the m. vastus lateralis. For this reason, these studies might not be comparable to this one, as the subjects in the current study performed a running GXT.

In the present study, OI was examined, but as mentioned, in some subjects it was not possible to define a breakpoint. This could be a problem for the application of NIRS for performance assessment, besides the fact that it is not valid. If athletes only rely on NIRS to determine the thresholds versus combining it with other methods, they may end up with no results at all, if the breakpoint cannot be established.

11.3.   NdRCP

The findings of the current study show no significant difference between NdRCP and RCP (P = 0.789), however the typical error of the estimate and the CoV is interpreted as *very high* (150s, and 16.6%, respectively). This reveals that NIRS is not able to determine RCP validly.

For the determination of NdRCP, Δ[HHb] was the analyzed parameter. The reason for this is that Wang et al. (2006) conducted a study to find out which variable measured by NIRS reflected RCP most accurately, by assessing Δ[HHb], Δ[O_2Hb], and tissue saturation index for highest correlation with RCP. They resulted in the workload and $\dot{V}O_2$ at NdRCP in Δ[HHb] being significantly correlated with workload and $\dot{V}O_2$ at RCP (r = 0.99, P < 0.0001 and r = 0.987, P < 0.0001, respectively) (Wang et al., 2006).

The findings of the present study reveal that there is no significant difference between NdRCP and RCP (P = 0.789), with the difference between the two methods being 36 s. However, the typical error of the estimate is 150 s, which is interpreted as *very large*. In addition, CoV is 16.6%, which indicates high intraindividual variability between the two methods. The high 95% LoA also signify inconsistencies between the two methods. Figure 17 shows the relationship between NdRCP and RCP. It is clearly visible, that the dots data stray far from the regression line, which also shows variability. Therefore, though there is no significant difference between NdRCP and RCP, this study revealed that NIRS is not a valid tool to determine RCP.
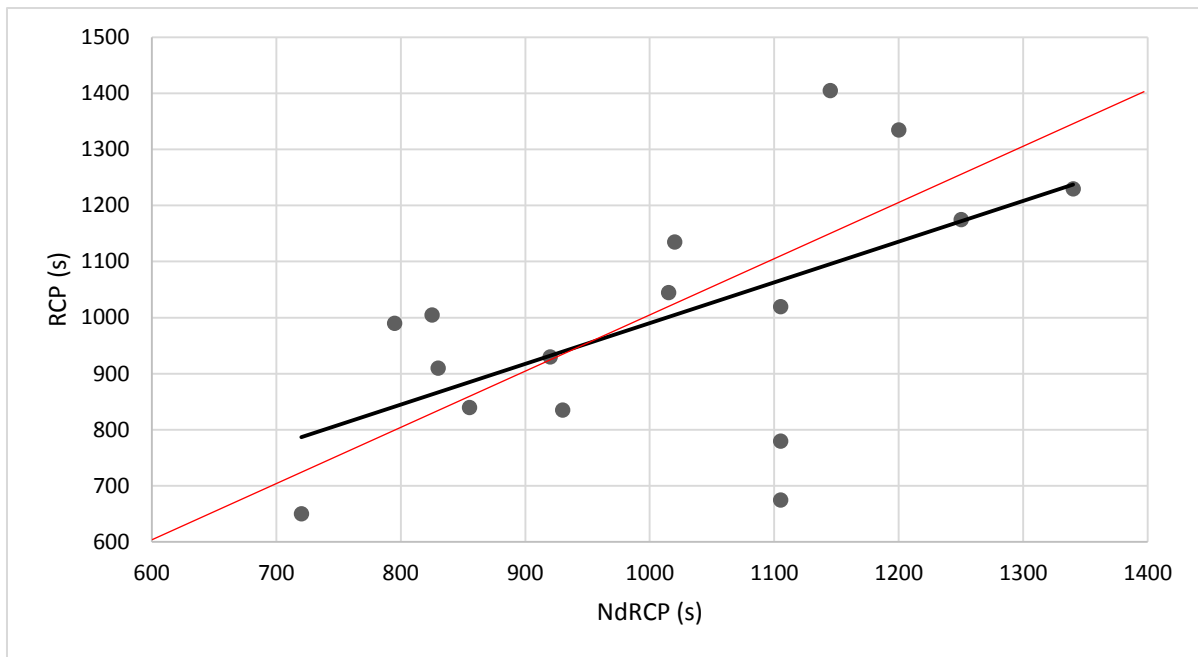
Figure 17. Relationship between NdRCP (s) and RCP (s). The black dots indicate the thresholds, the black line is the regression line and the red line is the line of identity. NdRCP = NIRS derived respiratory compensation point, RCP = respiratory compensation point

Two common trends of Δ[HHb] were observed among the participants. The first trend is reflected in the literature, namely a linear incline and at NdRCP, Δ[HHb] plateaus or increases at a slower rate, which can be seen in subject 2 (Figure 18).
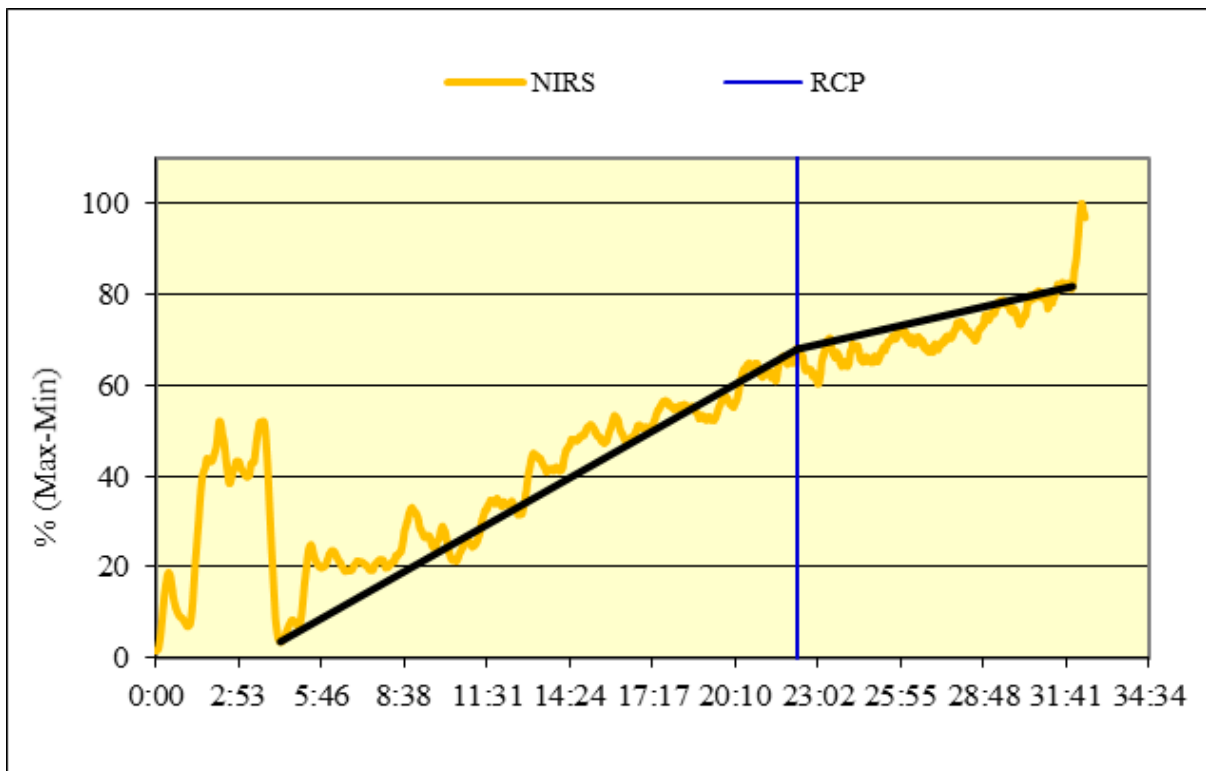
Figure 18. Graph of Δ[HHb]of subject 2 to determine RCP. The yellow line is Δ[HHb], the black lines are the manually placed regression lines and the vertical blue line is the RCP.

The second trend represents a relative constant Δ[HHb] with a subsequent decline at NdRCP, shown in subject 16 in Figure 19.
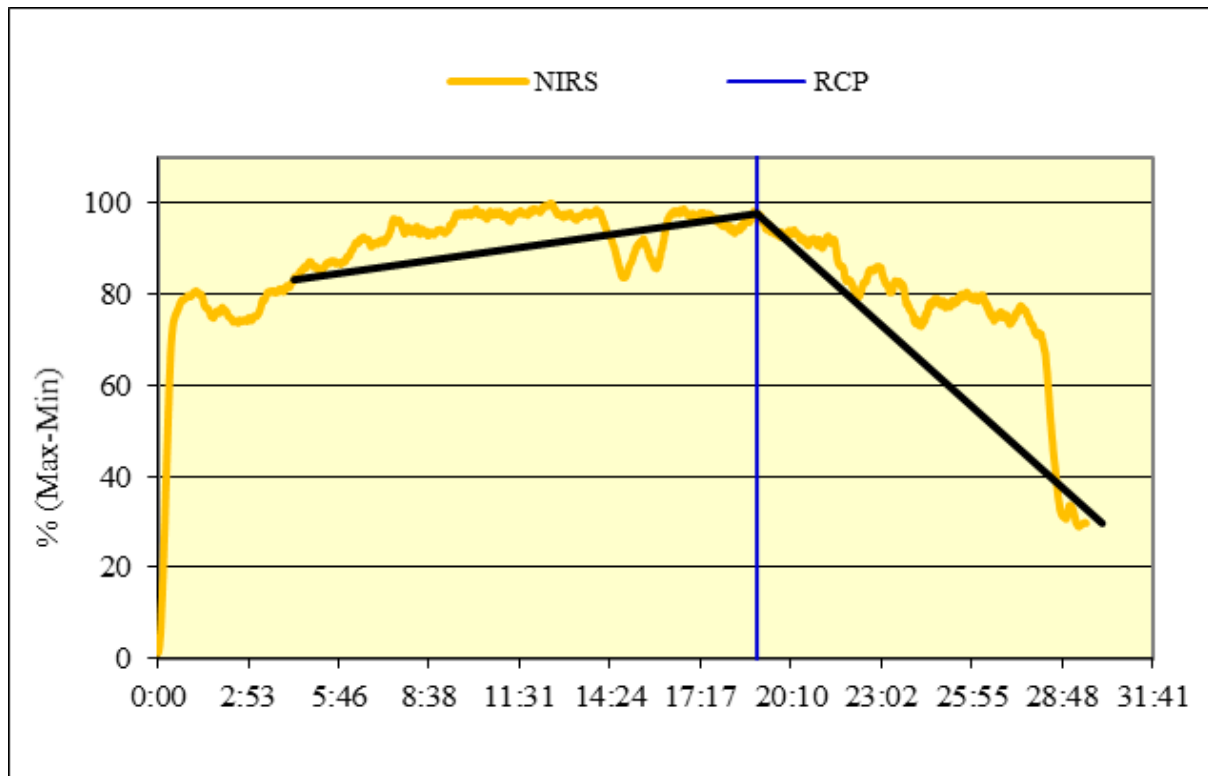
Figure 19. Graph of Δ[HHb]of subject 16 to determine RCP. The yellow line is Δ[HHb], the black lines are the manually placed regression lines and the vertical blue line is the RCP.

The present study is in line with all other studies, which determined NdRCP in Δ[HHb], as there is no significant difference between NdRCP and RCP (Bellotti et al., 2013; Fontana et al., 2015; Keir et al., 2015; Miura et al., 1998; Racinais et al., 2014). However, since the CoV, the standard error of the estimate and the 95% LoA are so high, NIRS cannot be seen as a valid tool to determine RCP. This is in contrast to all other studies, which agree that NIRS is valid in establishing RCP. Also the two studies which implemented a running GXT came to this conclusion (Karatzanos et al., 2010; Snyder & Parmenter, 2009).

One encountered problem was the fact, that of the 18 datasets, in 4 subjects a breakpoint in Δ[HHb] could not be found. This would be problematic for athletes during performance assessment.

12.Limitations

12.1.   Criterion

One of the limitations of this study is that the ML 3B and the MS CPX were not validated against a criterion. A criterion is considered to be the Douglas bag, or a metabolic cart, which has been validated against the Douglas bag and has produced correct values. Therefore, the current results cannot state which system measures more accurately. Furthermore, these data cannot be compared to reference values. Though it was not the goal of the study to do so, including a criterion would have provided more results.

12.2.   Separate GXTs

As mentioned before, it was not possible to apply the systems simultaneously. This is clearly a limitation to the study. The human body is not able to produce the exact same values in two separate GXTs, as factors such as nutrition, previous activity and sleep influence performance (Hollmann et al., 2006). A simultaneous measurement during the same GXT would have eliminated this factor.

12.3.   Adipose Tissue Thickness

Adipose tissue thickness (ATT) can have an effect on the NIRS signal (Van Beekvelt, Borghuis, M., Wevers, & Colier, 2001). The depth of the NIRS signal is half the length of the distance between the light source and the detector. If ATT is too thick underneath the NIRS site, this will influence the signal, as the near infrared absorption differs between fat and muscle tissue. Muscle tissue absorbs more light than fat tissue, resulting in a stronger NIRS signal, if ATT under the NIRS is thicker. However, the metabolic activity of fat tissue is lower, which would result in inaccurate muscle oxygenation measurements if ATT values are high (Barstow, 2019; Bhambhani, 2004; Jones, Chiesa, Chaturvedi, & Hughes, 2016). Therefore, Barstow (2019) suggests to measure ATT in every subject to ensure sufficient light penetration. In addition, he points out that, if possible, it would be desirable to include participants with ATT values which are lower than ¼ of the source-detector distance.

This demonstrates a limitation of the present study, since ATT was not obtained. However, ATT at the m. gastrocnemius lateralis is less than at the m. vastus lateralis (Leahy, Toomey, McCreesh, O'Neill, & Jakeman, 2012), making this an optimal site to place the NIRS (Jones et

al., 2016), which was done in the present study. Nevertheless, ATT data was not recorded, thus the results of the current study need to be analyzed with caution. However, since the cohort consisted of sport students, ATT can be suggested to be low.

12.4.   NIRS Data Analysis

Some difficulties occurred while analyzing the NIRS data. The first issue, which needs to be addressed, is the fact that the NIRS data had a very large variety in outcome. The reason for this is that relative concentration changes are measured versus absolute concentrations. Since relative concentration changes are measured, the application plays a role in the data recording. This means that the pressure with which NIRS is applied and the location of the device on the muscle will have an effect on the obtained data and results can vary. The consequence of this seems to be that in some participants the data produced interpretable curves, when plotted against time, whereas in others, no analysis could be performed. This led to difficulties when trying to determine thresholds.

The second problem was that NIRS data was analyzed in relation to the baseline. This means that the data as a function of time varied depending on whether or not the values increased or decreased after the baseline. This produced even more variability.

12.5.   NdAT Determination

Another problem faced was within the method of determining the NdAT. According to the literature, a constant $\Delta[tHb]$ is the requirement for being able to determine the threshold (Grassi et al., 1999). However, no study revealed how constancy was obtained or analyzed, leaving the assumption that $\Delta[tHb]$ constancy was visually determined.

The method chosen to $\Delta[tHb]$ constancy for the present study was not found in the literature but seemed to be a reliable and well calculated method. However, this needs to be taken into consideration when looking at the data.

12.6.   Heterogeneous Group of Participants

A point, which can be criticized in this study is the heterogeneity of the participants. Even though subjects were all physically active sport students, $\dot{V}O_{2peak}$ ranged from 2,198 mL·min^{-1}

to 5,667 mL·min^{-1}, and from 2,146 mL·min^{-1} to 5,229 mL·min^{-1} for ML 3B and MS CPX, respectively. When looking at $\dot{V}O_{2peak}$ relative to body mass, it ranged from 37.9 mL·min^{-1}·kg^{-1} to 62.6 mL·min^{-1}·kg^{-1} for ML 3B and from 39 mL·min^{-1}·kg^{-1} to 57.8 mL·min^{-1}·kg^{-1} for MS CPX. Other studies included much more homogenous groups, both for studies comparing metabolic gas analyzers as well as in studies analyzing NIRS thresholds (e.g. Brandes, Klein, Ginsel, & Heitmann, 2015; Wang, Tian, & Zhang, 2012). In addition, many of the subjects had never run on a treadmill before. This could have influenced the results in the way that the participants' second GXT was more economical since they already had the experience of running on the treadmill. However, GXTs were performed in a randomized order to account for this error.

13.Conclusion

In conclusion, ML 3B and MS CPX produce significantly different $\dot{V}O_2$ values in a cohort of sports students with ML 3B measuring higher, however, this difference can be corrected with an equation. While $\dot{V}CO_2$ is significantly different at certain stages, $\dot{V}E$ is not significantly different between the two systems.

According to this study, it is not possible to determine the NdAT and the NdRCP during running with NIRS validly. There is definitely need for further research on determining thresholds with NIRS in running GXTs. In addition, a study comparing different NIRS parameters to determine NdAT would be necessary.

References

Atkinson, G., Davidson, R. C. R., & Nevill, A. M. (2005). Performance Characteristics of Gas Analysis Systems : What We Know and What We Need to Know. *International Journal of Sports Medicine*, *26*(1), 2–10. https://doi.org/10.1055/s-2004-830505

Austin, K. G., Daigle, K. A., Patterson, P., Cowman, J., Chelland, S., Emily, M., & Haymes, E. M. (2005). Reliability of Near-Infrared Spectroscopy for Determining Muscle Oxygen Saturation During Exercise. *Research Quarterly for Exercise and Sport*, *76*(4), 440–449.

Barstow, T. J. (2019). Cores of Reproducibility in Physiology Understanding near infrared spectroscopy and its application to skeletal muscle research. *Journal of Applied Physiology*, *126*, 1360–1376. https://doi.org/10.1152/japplphysiol.00166.2018

Beaver, W. L., Wasserman, K., & Whipp, B. J. (1986). A new method for detecting anaerobic threshold by gas exchange. *Journal of Applied Physiology*, *60*(6), 2020–2027. https://doi.org/10.1152/jappl.1986.60.6.2020

Beijst, C., Schep, G., Breda, E. Van, Wijn, P. F. F., & Pul, C. Van. (2013). Accuracy and precision of CPET equipment: A comparison of breath-by-breath and mixing chamber systems. *Journal of Medical Engineering and Technology*, *37*(1), 35–42. https://doi.org/10.3109/03091902.2012.733057

Bellotti, C., Calabria, E., Capelli, C., & Pogliaghi, S. (2013). Determination of maximal lactate steady state in healthy adults: Can NIRS help? *Medicine and Science in Sports and Exercise*, *45*(6), 1208–1216. https://doi.org/10.1249/MSS.0b013e3182828ab2

Beltrami, F. G., Froyd, C., Mamen, A., & Noakes, T. D. (2014). The validity of the Moxus Modular metabolic system during incremental exercise tests: Impacts on detection of small changes in oxygen consumption. *European Journal of Applied Physiology*, *114*(5), 941–950. https://doi.org/10.1007/s00421-014-2825-x

Bhambhani, Y. N. (2004). Muscle Oxygenation Trends During Dynamic Exercise Measured by Near Infrared Spectroscopy. *Canadian Journal of Applied Physiology*, *29*(4), 504–523. https://doi.org/10.1139/h04-033

Bland, J. M., & Altman, D. G. (1968). Statistical methods for assessing agreement between two methods of clinical measurement. *The Lancet*, 307–310.

Brandes, M., Klein, H., Ginsel, S., & Heitmann, A. (2015). Comparability of three mobile respiratory gas analyzers. *Sportwissenschaft*, *45*(4), 168–172.

https://doi.org/10.1007/s12662-015-0368-x

Brooks, G. A. (1985). Anaerobic threshold : Review of the concept and directions for future research. *Medicine and Science in Sports and Exercise*, *17*(1), 22–31.

Carter, J., & Jeukendrup, A. E. (2002). Validity and reliability of three commercially available breath-by-breath respiratory systems. *European Journal of Applied Physiology*, *86*(5), 435–441. https://doi.org/10.1007/s00421-001-0572-2

Coquart, J. B., Mucci, P., L'Hermette, M., Chamari, K., Tourny, C., & Garcin, M. (2017). muscle oxyhemoglobin inflection point with time-to-exhaustion during heavy-intensity exercise. *The Journal of Sports Medicine and Fitness*, *57*(3), 171–178. https://doi.org/10.23736/S0022-4707.16.06053-9

Crouter, S. E., Antczak, A., Hudak, J. R., DellaValle, D. M., & Haas, J. D. (2006). Accuracy and reliability of the ParvoMedics TrueOne 2400 and MedGraphics VO2000 metabolic systems. *European Journal of Applied Physiology*, *98*(2), 139–151. https://doi.org/10.1007/s00421-006-0255-0

Díaz, V., Benito, P. J., Peinado, A. B., Álvarez, M., Martín, C., Di Salvo, V., … Calderón, F. J. (2008). Validation of a new portable metabolic system during an incremental running test. *Journal of Sports Science and Medicine*, *7*(4), 532–536.

Douglas, C. G. (1911). A method for determining the total respiratory exchange in man. *Journal of Physiology*, (42), 17–18.

Draper, N., & Hodgson, C. (2008). *Adventure Sport Physiology*. The Atrium, Southern Gate, Chichester, West Sussex, UK: John Wiley & Sons, Ltd.

Duffield, R., Dawson, B., Pinnington, H. C., & Wong, P. (2004). Accuracy and reliabillity of a Cosmed K4b 2 portable gas analysis system. *Journal of Science and Medicine in Sport*, *7*(1), 11–22.

Eston, R., & Reilly, T. (1996). *Kinanthropometry and Exercise Physiology Laboratory Manual*. London: E & FN Spon.

Fontana, F. Y., Keir, D. A., Bellotti, C., De Roia, G. F., Murias, J. M., & Pogliaghi, S. (2015). Determination of respiratory point compensation in healthy adults: Can non-invasive near-infrared spectroscopy help? *Journal of Science and Medicine in Sport*, *18*(5), 590–595. https://doi.org/10.1016/j.jsams.2014.07.016

Foss, & Hallén, J. (2005). Validity and stability of a computerized metabolic system with

mixing chamber. *International Journal of Sports Medicine*, *26*(7), 569–575. https://doi.org/10.1055/s-2004-821317

Grassi, B., Quaresima, V., Marconi, C., Ferrari, M., Cerretelli, P., Quaresima, V., … Cerretelli, P. (1999). Blood lactate accumulation and muscle deoxygenation during incremental exercise. *Journal of Applied Physiology*.

Hill, A. V., Long, C. H. N., & Lupton, H. (1924). Muscular Exercise, Lactic Acid, and the Supply and Utilisation of Oxygen. - Parts IV-VI. *Proc Roy Soc*, *97*, 84–138.

Hill, A. V., & Lupton, H. (1923). Muscular Exercise, Lactic Acid, and the Supply and Utilization of Oxygen. *Quarterly Journal of Medicine*, 135–171.

Hodges, L. D., Brodie, D. A., & Bromley, P. D. (2005). Validity and reliability of selected commercially available metabolic analyzer systems. *Scandinavian Journal of Medicine and Science in Sports*, *15*(5), 271–279. https://doi.org/10.1111/j.1600-0838.2005.00477.x

Hohmann, A., Lames, M., & Letzelter, M. (2007). *Einführung in die Trainingswissenschaft* (4th ed.). Wiebelsheim: Limpert Verlag GmbH.

Hollmann, W., Strüder, H. K., Predel, H.-G., & Tagarakis, C. V. M. (2006). *Spiroergometrie. Kardiopulmonale Leistungsdiagnostik des Gesunden und Kranken*. Stuttgart: Schattauer.

Hopkins, W. G., Schabort, E. J., & Hawley, J. A. (2001). Reliability of power in physical performance tests. *Sports Medicine*, *31*(3), 211–234. https://doi.org/10.2165/00007256-200131030-00005

Hottenrott, K., & Neumann, G. (2010). *Trainingswissenschaft*. Aachen: Meyer & Meyer Verlag.

Jensen, K., Jørgensen, S., & Johansen, L. (2002). A metabolic cart for measurement of oxygen uptake during human exercise using inspiratory flow rate. *European Journal of Applied Physiology*, *87*(3), 202–206. https://doi.org/10.1007/s00421-002-0616-2

Jones, S., Chiesa, S. T., Chaturvedi, N., & Hughes, A. D. (2016). Recent developments in near-infrared spectroscopy (NIRS) for the assessment of local skeletal muscle microvascular function and capacity to utilise oxygen. *Artery Research*, *16*, 25–33. https://doi.org/10.1016/j.artres.2016.09.001

Karatzanos, E., Paradisis, G., Zacharogiannis, E., Tziortzis, S., & Nanas, S. (2010). Assessment of ventilatory threshold using near-infrared spectroscopy on the

gastrocnemius muscle during treadmill running. *International Journal of Industrial Ergonomics*, *40*(2), 206–211. https://doi.org/10.1016/j.ergon.2009.02.003

Keir, D. A., Fontana, F. Y., Robertson, T. C., Murias, J. M., Paterson, D. H., Kowalchuk, J. M., & Pogliaghi, S. (2015). Exercise Intensity Thresholds: Identifying the Boundaries of Sustainable Performance. *Medicine and Science in Sports and Exercise*, *47*(9), 1932–1940. https://doi.org/10.1249/MSS.0000000000000613

Kroidl, R. F., Schwarz, S., & Lehnigk, B. (2010). *Kursbuch Spiroergometrie. Technik und Befundung verständlich gemacht*. Stuttgart: Georg Thieme Verlag.

Kuipers, H., Rietjens, G., Verstappen, F., Schoenmakers, H., & Hofman, G. (2003). Effects of Stage Duration in Incremental Running Tests on Physiological Variables. *International Journal of Sports Medicine*, *24*, 486–491.

Larsson, P. U., Wadell, K. M. E., Jakobsson, E. J. I., Burlin, L. U., & Henriksson-Larsén, K. B. (2004). Validation of the MetaMax II Portable Metabolic Measurement System. *International Journal of Sports Medicine*, *25*(2), 115–123. https://doi.org/10.1055/s-2004-819953

Leahy, S., Toomey, C., McCreesh, K., O'Neill, C., & Jakeman, P. (2012). Ultrasound Measurement of Subcutaneous Adipose Tissue Thickness Accurately Predicts Total and Segmental Body Fat of Young Adults. *Ultrasound in Med. & Biol.*, *38*(1), 28–34. https://doi.org/10.1016/j.ultrasmedbio.2011.10.011

Leprêtre, P. M., Weissland, T., Paton, C., Jeanne, M., Delannaud, S., & Ahmaidi, S. (2012). Comparison of 2 portable respiratory gas analysers. *International Journal of Sports Medicine*, *33*(9), 728–733. https://doi.org/10.1055/s-0031-1301316

Macfarlane, D. J. (2001). Automated Metabolic Gas Analysis Systems: A Review. *Sports Med*, *31*(12), 841–861. https://doi.org/10.3390/polym9120651

Macfarlane, D. J., & Wong, P. (2012). Validity, reliability and stability of the portable Cortex Metamax 3B gas analysis system. *European Journal of Applied Physiology*, *112*(7), 2539–2547. https://doi.org/10.1007/s00421-011-2230-7

McArdle, W. D., Katch, F. I., & Katch, V. L. (2001). *Exercise physiology: Energy, Nutrition, and Human Performance* (5th ed.). Lippincott Williams & Wilkins.

McLaughlin, J. E., King, G. A., Howley, E. T., Bassett, D. R., & Ainsworth, B. E. (2001). Validation of the COSMED K4 b2 portable metabolic system. *International Journal of*

Sports Medicine, *22*(4), 280–284. https://doi.org/10.1055/s-2001-13816

Meyer, T., Davidson, R. C. R., & Kindermann, W. (2005). Ambulatory Gas Exchange Measurements - Current Status and Future Options. *International Journal of Sports Medicine*, *26*(1), 19–27. https://doi.org/10.1055/s-2004-830507

Meyer, T., Georg, T., Becker, C., & Kindermann, W. (2001). Reliability of gas exchange measurements from two different spiroergometry systems. *International Journal of Sports Medicine*, *22*(8), 593–597. https://doi.org/10.1055/s-2001-18523

Midgley, A. W., Bentley, D. J., Luttikholt, H., Mcnaughton, L. R., & Millet, G. P. (2008). Challenging a Dogma of Exercise Physiology. Does an Incremental Exercise Test for Valid ˙VO2max Determination Really Need to Last Between 8 and 12 Minutes ? *Sports Med*, *38*(6), 441–447.

Miura, T., Takeuchi, T., Sato, H., Nishioka, N., Terakado, S., Fujieda, Y., & Ibukiyama, C. (1998). Skeletal muscle deoxygenation during exercise assessed by near-infrared spectroscopy and its relation to expired gas analysis parameters. *Japanese Circulation Journal*, *62*(9), 649–657. https://doi.org/10.1253/jcj.62.649

Perret, C., & Mueller, C. (2006). Validation of a new portable ergospirometric device (Oxycon Mobile) during exercise. *International Journal of Sports Medicine*, *27*(5), 363–367. https://doi.org/10.1055/s-2005-865666

Poole, D. C., Burnley, M., Vanhatalo, A., Rossiter, H. B., & Jones, A. M. (2017). Critical Power: An Important Fatigue Threshold in Exercise Physiology. *Medicine and Science in Sports and Exercise*, *48*(11), 2320–2334. https://doi.org/10.1249/MSS.0000000000000939.Critical

Racinais, S., Buchheit, M., & Girard, O. (2014). Breakpoints in ventilation, cerebral and muscle oxygenation, and muscle activity during an incremental cycling exercise. *Frontiers in Physiology*, *5 APR*(April), 1–6. https://doi.org/10.3389/fphys.2014.00142

Raleigh, C., Donne, B., & Fleming, N. (2018). Association between different Non-Invasively Derived Thresholds with Lactate Threshold during graded incremental exercise. *International Journal of Exercise Science*, *11*(4), 391–403. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/29541332%0Ahttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC5841671

Rietjens, G. J. W. M., Kuipers, H., Kester, A. D. M., & Keizer, H. A. (2001). Validation of a

computerized metabolic measurement system (Oxycon-Pro) during low and high intensity exercise. *International Journal of Sports Medicine*, *22*(4), 291–294. https://doi.org/10.1055/s-2001-14342

Schnabel, G., Harre, D., & Borde, A. (1997). *Trainingswissenschaft. Leistung - Training - Wettkampf*. Berlin: Sportverlag.

Snyder, A. C., & Parmenter, M. A. (2009). Using Near-Infrared Spectroscopy to determine Maximal Steady State Exercise Intensity. *Strength And Conditioning*, *23*(6), 1833–1840.

Van Beekvelt, M. C. P., Borghuis, M. S., M., van E. B. G., Wevers, R. A., & Colier, W. N. J. M. (2001). Adipose Tissue Thickness affects in Vivo Quantitative Near-IR Spectroscopy in Human Skeletal Muscle. *The Biomedical Society and the Medical Research Society*, *101*, 21–28.

Van Der Zwaard, S., Jaspers, R. T., Blokland, I. J., Achterberg, C., Visser, J. M., Den Uil, A. R., … De Ruiter, C. J. (2016). Oxygenation threshold derived from near- Infrared spectroscopy: Reliability and its relationship with the first ventilatory threshold. *PLoS ONE*, *11*(9), 1–16. https://doi.org/10.1371/journal.pone.0162914

Vogler, A. J., Rice, A. J., & Gore, C. J. (2010). Validity and reliability of the Cortex MetaMax3B portable metabolic system. *Journal of Sports Sciences*, *28*(7), 733–742. https://doi.org/10.1080/02640410903582776

Wang, B., Tian, Q., & Zhang, Z. (2012). Comparisons of local and systemic aerobic fitness parameters between finswimmers with different athlete grade levels. *European Journal of Applied Physiology*, *112*, 567–578. https://doi.org/10.1007/s00421-011-2007-z

Wang, B., Tian, Q., Zhang, Z., & Gong, H. (2009). Comparison between the changes in muscle oxygenation and blood lactate concentration in finswimmers during incremental exercise. *Proc. of SPIE*, *7519*. https://doi.org/10.1117/12.845434

Wang, B., Xu, G., Tian, Q., Sun, J., Sun, B., Zhang, L., … Gong, H. (2012). Differences between the vastus lateralis and gastrocnemius lateralis in the assessment ability of breakpoints of muscle oxygenation for aerobic capacity indices during an incremental cycling exercise. *Journal of Sports Science and Medicine*, *11*(4), 606–613.

Wang, L., Yoshikawa, T., Hara, T., Nakao, H., Suzuki, T., & Fujimoto, S. (2006). Which common NIRS variable reflects muscle estimated lactate threshold most closely? *Applied Physiology, Nutrition, and Metabolism*, *31*(5), 612–620. https://doi.org/10.1139/h06-069

Wasserman, K. (1984). The Anaerobic Threshold Measurement to Evaluate Exercise Performance. *The American Review of Respiratory Diseases*, *2*, 35–40.

Wasserman, K., Whipp, B. J., Koyal, S. N., & Beaver, W. L. (1973). Anaerobic threshold and respiratory gas exchange during exercise. *Journal of Applied Physiology*, *35*(2), 236–243.

Weineck, J. (2010). *Optimales Training*. Balingen: Spitta Verlag.

Wilmore, J. H., Costill, D. L., & Kenney, W. L. (2008). *Physiology of Sport and Exercise* (4th ed.). Champaign, IL, USA: Human Kinetics.

List of Tables

List of Figures

Abbreviations

ATP	Adenosine triphosphate
AT	Anaerobic threshold
ATT	Adipose tissue thickness
CoV	Coefficient of Variation
$\Delta[HHb]$	Deoxyhemoglobin
$\Delta[O_2Hb]$	Oxyhemoglobin
$\Delta[tHb]$	Total hemoglobin
GXT	Graded exercise test
HR	Heart Rate
ICC	Intraclass correlation coefficient
LoA	Limits of Agreement
LT	Lactate threshold
ML 3B	MetaLyzer 3B
MLSS	Maximal lactate steady state
MS CPX	MasterScreen CPX
NdAT	NIRS derived anaerobic threshold
NdRCP	NIRS derived respiratory compensation point
NIRS	Near infrared spectroscopy
OI	Oxygenation index ($\Delta[O_2Hb] - \Delta[HHb]$)
RCP	Respiratory compensation point
TSI	Tissue saturation index (ratio of $\Delta[O_2Hb]$ to $\Delta[tHb]$)
V_{max}	Maximal running speed
$\dot{V}CO_2$	Carbon dioxide output
$\dot{V}E$	Minute ventilation
$\dot{V}O_2$	Oxygen uptake
$\dot{V}O_{2peak}$	Maximal oxygen uptake

Declaration

I, Keltey Kathleen McGirr, hereby confirm that this manuscript is my own work, is not copied from any other persons work (published or unpublished), and has not been previously submitted at the University of Vienna or elsewhere. All foreign sources of information used in the text are listed under references.

_____   _____

Name   Location, Date