ANALYSIS OF RELIABILTY AND VALIDITY OF CRITICAL POWER TESTING IN THE FIELD

Bettina Karsten

A thesis submitted in partial fulfilment of the requirements of the University of Greenwich for the Degree of Doctor of Philosophy

May, 2014

DECLARATION

"I certify that this work has not been accepted in substance for any degree, and is not concurrently being submitted for any degree other than that of Doctor of Philosophy being studied at the University of Greenwich. I also declare that this work is the result of my own investigations except where otherwise identified by references and that I have not plagiarised the work of others".

Ms Bettina Karsten: _____

……

PhD Supervisors

Dr Chris Beedie: _____

Dr Simon Jobson: _____

Dr James Hopker: _____

Date:

ACKNOWLEDGEMENTS

I only started to really believe in the completion of this thesis, once I was able to see the finish line without the use of a telescope. Reaching it was the result of a great team, who supported this effort in such a variety of ways.

Firstly I would like to thank Dr Chris Beedie for his outstanding input in this thesis. Being continuously challenging, believing in me and being never far away from a positive comment has helped me through many thin moments. I could not have wished for a better supervisor. Chris made sure my PhD did not take one day longer than necessary. I can only hope that this thesis will not just do some justice to his excellent support and inspiration but maybe even put a smile on his ever-challenging face. Furthermore I would like to thank both my second supervisors, Dr Simon Jobson and Dr James Hopker, the physiological masterminds behind this thesis. Simon with his wonderfully applied, hands on approach and James with his very careful, considered approach have provided a fantastic balance for the million questions, which came up along this journey. I hope to continuously work with this amazing supervisory team in the future.

This thesis would have never been possible without the cyclists, who participated in all the research. Their epic physical efforts, their sweat, their blood and their wonderful dedication provided me with the much needed results. I really enjoyed working with each and every one of them, sharing their passion for sport and cycling. Performing research can be demanding but these cyclists made it not just worthwhile but also very enjoyable.

There are also a few members of staff at Greenwich, who deserve a huge thank you. Prof Paul Ward not just let me get on with my thesis but also provided me with the necessary space to do so. I did not take this for granted. Kelly Cooper, who endlessly booked the

laboratory for me, maintained the required equipment and finally jumped out of a plane with me - thank you. Another colleague, who literally saved my life on a few occasions, is Amanda Lewis, our finance manager. Within reason, and if I really needed something - I got it. Amanda, the posted chocolate will never do justice to your support. Finally I would like to express my gratitude to Dr 'Harry' Colpus, whose philosophical yet pragmatically approach to life and its' research questions has been very supportive.

Another magnificent person to thank is Geoff Wiles. What can I say? Geoff, an absolute cycling legend and now track champion of the world, literally recruited me into his sport during his undergraduate degree at Greenwich. Cheers Geoff, you really got me into trouble here. Having met a number of his fellow cyclists, I now know that nothing less than a life transformation occurs when spending time with Geoff Wiles. I would also like to recognise his endless support on technical issues, which were numerous. Geoff was always there and without him and his magic bike tools, I would have crashed and burned badly.

Finally I would like to express my absolute gratitude to my partner Anne (vive la France), my family (Eltern, es ist endlich vollbracht) and my friends for their love, support, patience and absolute faith in my ability to master this PhD thesis. Please accept my apologies for not having been there much, for being grumpy, for being tired, for being totally boring and for having a little racing bike permanently going around in my brain. I am really looking forward to fully being there again.

Finally I can't wait to regularly race my bike through the beautiful English countryside.

Bettina Karsten

May 2014

In loving memory of Margaret King,

An inspirational teacher

and

Penny Humphrey

A simply wonderful friend

“All my life I've been surrounded by people who are smarter than I am, but I found I could always keep up by working hard.”

Glenn Seaborg

ABSTRACT

Critical Power (CP) provides a useful indicator of training status in continuous activities lasting between approximately 2 and 30 minutes. To date, determination of CP has been mostly constrained to the laboratory. The conventional laboratory protocol commonly requires multi-day time-to-exhaustion tests. The thesis below addresses whether CP can a) be determined without multi-day exhaustive trials and b) be determined in the field. Studies compared the findings of conventional laboratory determination methods with novel protocols in which either the testing mode, the recovery period between exhaustive trials, or the environment were manipulated. Study 1 demonstrates that the recently developed 3-min all-out protocol does not result in valid CP values, when using the isokinetic ergometer mode. Results indicated low levels of agreement (mean of 23-45 W) between conventionally determined CP and values derived through the 3-min all-out protocol. The average prediction error associated with the relationship between CP and the 3-min all-out End Power was 7%. In Study 2, values of CP derived through a conventional laboratory CP protocol were compared with those determined outdoors on a cycling track. High levels of agreement (mean of 2 - 14 W) were observed between the laboratory and field values of CP. The average prediction error associated with the relationship between laboratory and field CP was 2.2%. Based in the laboratory, Study 3 compares a 24 h recovery protocol with a 3 h and a 30 min recovery protocols. High levels of agreement (mean of -2 - 11 W and -2 - 8 W respectively) were observed across protocols. The average prediction error associated with the relationship between the 24 h and 3 h and the 24 h and 30 min protocols was 2.4% and 3.3% respectively, suggesting that determination of CP could be made

more ‘athlete-friendly’ by shortening the conventional 3-day protocol to one day. Study 4 uses three protocols to evaluate the shortened 30 min protocol in ecological valid open road conditions. Values of CP derived from laboratory protocols were compared to a) those derived from pre-planned and ‘grouped’ maximal efforts of 3 min, 7 min and 12 min with a 30 min recovery period between efforts (protocol 1), b) those of discrete and randomly performed, yet still pre-planned maximal efforts of the same durations (protocol 2), and c) to those extracted from self-directed training and racing of these same durations (protocol 3). The average prediction error associated with the relationship between the laboratory and the field values of CP was 3.1% (protocol 1), 4.9% (protocol 2) and 4.1% (protocol 3). Results, whilst providing high levels of agreement, also suggested that in particular protocols 1 and 3 potentially provide a practical and arguably ecologically valid alternative to the conventional laboratory protocol. Study 5 further investigates the overall CP determination procedure by comparing collected values of CP derived through 3 data points with both, CP laboratory and field values derived through 2 data points. High levels of agreement and low prediction errors (average 3.2%) associated with the relationship between 3 data points and 2 data points-derived CP were observed. Studies collectively provide support for the acceptance of field performance testing using CP, with either a 30 min inter-maximal effort recovery period or alternatively the extraction of non-planned specified maximal efforts from training and racing data. Overall the investigations described in the thesis suggest that CP determination is feasible beyond the laboratory and that consumer-level technology provides satisfactory ease and reliability of measurement in this context. Moreover, these novel CP determination methods allow coaches to continuously monitor their athletes.

TABLE OF CONTENTS

TABLES

FIGURES

ABBREVIATIONS

[ ]	**Concentration**
Δ	**Delta; a difference or a change in value**
ADC	**Anaerobic Distance Capacity**
ARD	**Anaerobic Running Distance**
ATP	**Adenosine 5' -triphosphate**
AWC	**Anaerobic Work Capacity**
Ca^{2+}	**Calcium**
CE	**Cycling Efficiency**
CMD	**Central Motor Drive**
CNS	**Central Nervous System**
CO_2	**Carbon Dioxide**
CoV	**Coefficient of Variation**
CP	**Critical Power (W)** asymptote of the hyperbolic relationship between power output and time to exhaustion
CS	**Critical Speed ($m \cdot min^{-1}$)**
CV	**Critical Velocity ($m \cdot min^{-1}$)** determined from the hyperbolic between velocity and time to exhaustion
CWR	**Constant Work Rate**
EP	**End Power (W)**
F	**Force (Nm)**
GE	**Gross Economy**
GET	**Gas Exchange Threshold**
H^+	**Hydrogen ion**
HR	**Heart Rate ($b \cdot min^{-1}$)**
HR_{max}	**Maximum Heart Rate ($b \cdot min^{-1}$)**
IAT	**Individual Anaerobic Threshold**
ICT	**Individual Critical Threshold**
J	**Joule**
K^+	**Potassium**
kJ	**Kilo Joules**
La	**Lactate**
LBF	**Leg Blood Flow**
LoA	**Limits of Agreement**
LT	**Lactate Threshold** exercise intensity which elicits a sustained increase in blood [lactate] above resting values
MAOD	**Maximal Accumulated Oxygen Deficit (L)**
MAP	**Maximal Aerobic Power**
MHC	**Myosin Heavy Chain**
MLSS	**Maximal Lactate Steady State** the highest $\dot{V}O_2$ at which blood lactate concentration can be stabilised
MLSSV	**Maximal Lactate Steady State Velocity**
MRS	**Magnetic Resonance Spectroscopy**

n	**Number (i.e. Participants)**
$NaHCO_3$	**Sodium Bicarbonate**
NIRS	**Near Infra-Red Spectroscopy**
NO	**Nitric Oxide**
nNO	**neural Nitric Oxide**
O_2 deficit	**Oxygen deficit**
P-1/t	**Power-Inverse Time Relationship**
PCr	**Phosphocreatine**
pH	**logarithmic scale which expresses the acidity or alkalinity of a solution**
P_i	**inorganic Phosphate**
P_{max}	**Instantaneous Maximal Power**
PO	**Power Output**
PP	**Peak Power**
r	**Correlation Coefficient**
RPM	**Revolution per Minute** used to define pedal cadence during cycle ergometry
SA	**Surface Area**
SD	**Standard Deviation**
SEE	**Standard Error of the Estimate**
SR	**Sarcoplasmatic Reticulum**
Tau (τ)	**Tau**
TT	**Time Trial**
TTE	**Time-to-Exhaustion**
VT	**Ventilatory Threshold**
W	**Watt**
W´	**Anaerobic Work Capacity** curvature constant of the power-duration relationship (J; kJ)
WEP	**Work done above End Power**
WR_{peak}	**Work Rate Peak (W)**
W-t	**Work-Time Relationship**
$\dot{V}O_2$	**Oxygen Uptake ($l \cdot min^{-1}$; $ml \cdot kg \cdot min^{-1}$)** volume of oxygen extracted and utilised from the inspired gas per unit of time
$\dot{V}O_{2max}$	**Maximal Oxygen Uptake ($l \cdot min^{-1}$; $ml \cdot kg \cdot min^{-1}$)** maximal oxygen uptake per unit of time
$\dot{V}O_{2peak}$	**Peak Oxygen Uptake ($l \cdot min^{-1}$; $ml \cdot kg \cdot min^{-1}$)** peak oxygen uptake per unit of time

CONFERENCES/PUBLICATIONS

Conference proceedings

Karsten, B., Jobson, S. A., Hopker, J. G., Petrigna, L., & Beedie, C. No differences between 24h, 3h and 30min recovery in the estimation of critical power and W′ in cycling. European Congress of Sports Science, 18th ECSS Congress 26-29 June 2013, Barcelona/ESP

Karsten, B., Jobson, S. A., Hopker, J. G. & Beedie, C. Critical power determination in the field. European Congress of Sports Science, 19th ECSS Congress 2nd – 5th of July 2014, Amsterdam/NL

Karsten, B., Jobson, S. A., Hopker, J. G. & Beedie, C. Determination of Critical Power from road data, World Conference of Cycling Science, 2nd – 3rd of July 2014, Leeds/UK

Peer reviewed publications

Karsten, B., Jobson, S.A., Hopker, J., Passfield, L., Beedie, C. (2014) An all-out 3-min test using the SRM isokinetic mode does not provide a valid estimate of critical power in cycling. *International Journal of Sports Medicine.* 35(4):304-309

Karsten, B., Jobson, S.A., Hopker, J., Jimenez, A., Beedie, C. (2014) High agreement between laboratory and field estimates of critical power in cycling. *International Journal of Sports Medicine.* 35(4):298-303

Karsten, B., Jobson, S.A., Hopker, J., Stevens, L., Beedie, C. Validity and reliability of critical power field testing. *European Journal of Applied Physiology*.
DOI: 10.1007/s00421-014-3001-z

CHAPTER 1: GENERAL INTRODUCTION

During the 2008 Olympic Games Great Britain won 14 medals in cycling events and with it the top place in the cycling medal league. At the 2012 Olympic Games Great Britain won a total of 12 medals and again topped the medal league in cycling [1]. The UK also produced the Tour de France winners in two consecutive years, 2012 and 2013. Since 2008 cycling in the UK has consequently gained significantly in popularity with a market growth of 28% and bike sales of 3.7 million in 2010 [2]. Reasons for taking up cycling are various. From rising fuel costs to carbon neutral commuting, recreational purposes, health benefits to sporting competitions, cycling has become a major physical activity in the UK [2]. This however is in contrast to a report titled 'sport and exercise science and medicine; building on the Olympic legacy to improve the nation's health', released by the House of Lords [3] in July 2012, which voiced concerns about the general quality and robustness of sports science research and its application into real-world sport and exercise. According to the report there is little evidence that findings in sports science are causing an impact on elite performance. Whilst the difficulties associated with research on elite athletes were acknowledged, a general recommendation was made to combine observations from elite athletes with rigorous research conducted on recreational athletes. Making science relevant, applied and transferable should therefore be of importance in the design of any sports science research. In addition, the report acknowledged a scientifically demonstrated relationship between physical activity and health benefits but highlighted an apparent lack in understandings of the underpinning mechanisms of that relationship. Physical activities, according to the Department of Health [4], are defined as every day activities, active recreational activities and sporting activities. Cycling satisfies all three definitions and it is used by the Department of Health [4], as a suitable example of physical activity. The ongoing success of competitive cycling in the UK partially reflects the widespread specific application of sports science methods and research. Arguably, however, these need to be applied further to fulfil the recommendations of the House of Lords [4] by being relevant, robust and transferable.

Cycling has been well researched in a laboratory setting and a link between cycling laboratory research and real-world cycling has been demonstrated [5]. However there is a

need for relevant field performance tests in cycling, as many of the field tests that coaches use are not sufficiently sensitive or reliable to provide a valid estimate of training effects [6] and have not been validated [7]. Whilst performance tests are an integral part of evaluation for competitive cyclists, relevant tests are generally performed in a controlled laboratory environment using a stationary cycle ergometer [8]. This was criticised by Peveler [9] and by Bertucci and Taiar [10] who stated that ergometer cycling does not replicate real-world cycling well enough to gain meaningful comparisons of performance. Despite many of the technological and physiological advances in recent years, surprisingly limited research has been performed addressing the translation of standardised laboratory tests into the field and consequently into "real-world" cycling [11,12]. Technological developments, such as mobile power meters, can potentially bridge this gap between the research laboratory and the real world. The major advantage of power meters is the provision of real time training and competition feedback such as power output (PO), cycling velocity or distance covered. This can be useful as it allows power-based training targeting specific adaptive processes, such as aerobic or anaerobic power, without having to rely on physiological feedback, such as heart rate (HR) or blood lactate concentration [lactate] [13]. HR has the disadvantage of a delayed response, for example during repeated short high-intensity exercise bouts [14], whilst the measurement of lactate often requires the athlete to stop his/her performance to sample blood. Consequently, the use of power meters has increased in popularity as evidenced by the increasing number of manufacturers developing their own power meter brand [15]. Previously only used by professional and elite cyclists, these devices are now commonly seen in amateur road, track and off-road cycling [16]. Whilst interpretation of the power data is still challenging [17], the measurement of PO contains such low error that it has been deemed suitable for tracking the small performance changes typically seen in elite cyclists [8]. Therefore field testing applications which use PO as testing variable could detect such small performance changes and consequently should be considered in order to further advance cycling research.

One index of performance in cycling is Critical Power (CP). CP is defined as a training and performance intensity sustainable over prolonged periods of time without a

continuous loss in homeostasis [18]. The determination[a] of CP requires accurate measurements of PO values which have, to date, mostly resulted in CP being largely constrained to the laboratory. Furthermore CP determination commonly requires multi-day time to exhaustion (TTE) trials. The combination of a resource-intensive testing protocol constrained to a laboratory setting has arguably not allowed CP to become a routinely assessed performance parameter in the "real world" of cycling [19].

Using cycling as a vehicle of physical activity, this thesis attempts to achieve some real-world impact at both, recreational and elite level. The over-arching aims of this research thesis are to question whether it is possible to accurately determine CP using a less cumbersome testing method, and whether CP, with a specific focus on road testing, can be determined in the field. The following literature review provides an overview on the meaning and significance of CP and its relevance for exercise tolerance and cycling. In order to contextualise CP, cycling PO and the importance of the performance monitoring process, current confinements of such process to the laboratory and the relevance of power meters are firstly introduced. A brief overview on the history of CP, its significance in human and cycling performance is subsequently provided which extends to the underpinnings of the power-duration concept of CP. Next, the physiological meaning of CP is put into the context of the maximal lactate steady state (MLSS) and an argument is presented on why CP cannot replace the MLSS and how relevant research has led to the construct of a new training zone/intensity domain. The physiological characteristics of each domain and their relation to training adaptations are described. This is followed by a discussion on performances around the CP intensity which explores the underpinning physiology of exercise tolerance further. This debate leads to the physiological meaning of the second parameter of the power-duration relation, W' and its particular role in exercise tolerance. A presentation on various mathematical models of CP, their inherent assumptions and their meaning are provided next which

[a] Various terms are used in the literature (i.e. estimate, derive, determine, test) which identify the process of gaining CP results. Whilst recognising distinct differences of each of these terms, 'determination of CP' will be used consistently within the current thesis to describe the process of obtaining results. In mathematical terms 'determination' is commonly used in the sense of fixing or defining a position, which is suitable in the context of this work [381].

leads to an in-depth discussion on factors affecting the determination of CP. Research comparing laboratory with field cycling performances and how mobile power meters has progressed such research is followed by a statement regarding the need of refined CP determination methods. Furthermore a debate will be presented which outline causes why CP to-date has not made any impact on real-world cycling. Specifically stating the research challenge, individual research questions are presented which provide a logical sequence to the PhD research process. This is followed by a general method section and leads to the individual studies. The final chapter summarizes this PhD, highlights its limitations, provides examples of future studies and presents the final argument that demonstrates how the study findings can be applied to cycling at all levels and how the findings can be used as a template for other sports.

CHAPTER 2: LITERATURE REVIEW

2.1 Power output in cycling; relevance of the monitoring process and ecological validity of laboratory testing

During cycling the leg muscles need to repeatedly produce high levels of mechanical force (F) on the pedals, usually over extended periods of time. In angular motion, multiplying F by the moment arm (d) (i.e. the perpendicular distance of the line of the force to the axis) gives the Torque (T = Fd), which when multiplied by the angular velocity ($\boldsymbol{\omega}$), gives the PO (PO = T$\boldsymbol{\omega}$), where $\boldsymbol{\omega}$ is expressed in $rads^{-1}$, Torque in Nm and PO in watts (W). Power is generated at the cost of the systematic and repeated imposition of physiological and mechanical stresses on the muscles in question [20]. Over time and with appropriate training, these stresses result in physiological and anatomical adaptations in the muscles, as well as the enhanced functioning of more central systems of the body supporting these muscles (e.g., cardio-vascular, respiratory and neuromuscular). The overall goal of training is to maximise these adaptive processes which allow the body to produce greater maximal levels of PO and/or maintain existing PO over longer durations. Effective sports training aims to maximise these adaptive processes that underlie the majority of strategies targeted at improving performance. Effective testing consequently requires protocols which are sensitive to detecting the small changes in performance capacity often seen in well-trained athletes [21,22].

Performing at a high level of competitive cycling necessitates the right balance between training load and recovery allowing the body to adapt maximally and to avoid injury [23]. High training loads, in particular high intensity training over extended periods of time can cause high levels of physiological and biomechanical stress and critical levels of fatigue. Positive adaptations and possibly negative responses to training can be monitored by on-going performance tests but to date, these tests still can present a significant challenge. Often requiring a maximal or near maximal effort or multi-day testing, performance tests can consequently compromise training whist contributing to fatigue accumulation.

Improvements in competitive cycling performance can be predicted by the observation of enhanced performance markers in training. Unfortunately the measurement of these predictors commonly not only requires sophisticated technologies, such as gas analysis and/or an electronically controlled ergometer, but also the expertise of one or more sports scientists. In short, the measurement of training-induced physiological and mechanical improvements in PO in cycling requires substantial time and financial resources.

Over and above issues of resources, laboratory based performance indices are also compromised by relatively low ecological validity. In short, an indoor or 'fixed' bicycle ergometer - fixed' implying that it is fixed in a stationary position and cannot move in any of the three planes in which bicycles move in the real world - is a relatively crude measure of all of the mechanical forces and physiological processes involved in cycling, and therefore provides only a crude approximation of real cycling performance [10,24]. Using bicycles equipped with power meters in the laboratory overcomes some of these related issues, such as the exact replication of the participant's usual riding position. However, the above holds true even for cycle-specific laboratory treadmills, which, despite being more ecologically valid than a fixed ergometer in allowing the bike to move in space in a more realistic manner, controls the speed of the rider in a way that would never happen in the real world.

Table 1. Summary of main forces and process relevant in cycling and degree to which these can be realistically reproduced in the laboratory

Force/Process	Sports Science Laboratory
Rolling resistance	Ergometer dependant. Some models can be calibrated for realistic rolling resistance.
Gravitational resistance	Not reproducible: exception uphill treadmill cycling.
Drag force	Limited reproducibility and ergometer dependant.
Head wind	Limited reproducibility and ergometer dependant.
Tail wind	Limited reproducibility and ergometer dependant.
Bicycle oscillation	Ergometer dependant Exception: Bicycle treadmill.
Weather conditions, such as humidity and temperature	Not reproducible. Exception: environmental laboratories.
Specific bicycle settings	Only on ergometers, which allow use of personal bicycles.
Sense of motion	Not reproducible. Exception: bicycles treadmill.
Self-control of speed or power	Protocol dependant.
Competition conditions, i.e. racing against other riders	Not reproducible.

To reliably evaluate the effectiveness of training interventions, more relevant data are required. A substantial step in this direction has been the development and recently increased refinement of mobile power meters, which can be mounted on real road bicycles. With the addition of telemetric technology, such power meters provide information in real time, alongside other real time information such as HR, cadence, speed and distance covered. These are useful when quantifying training loads and training adaptions, i.e. by a decrease in HR over set submaximal intensities [25] or by optimising pacing strategies in endurance events [26]. Such technologies could potentially provide coaches and scientists with more ecologically valid data, as it also allows power to be traced over time as a function of training or competition. Furthermore Atkinson and Brunskill [27] recommended the use of a power meter when applying variable racing strategies as it provides feedback on intensities in real time without reliance on HR or perception as the sensitivity of these variables is too low to monitor the meaningful changes in power during a race.

2.1.1 Development and function of mobile power meters

Power meter prototypes were first tested by the professional Team Strawberry during the race across America using the "Power Pacer" and by Greg LeMond using an "SRM" (Schober Rad Messtechnik, Juelich, Germany) in 1980. In 1989 mobile power meters became commercially available. The principle units in power meters are strain gauges. Strain gauges measure the applied torque, created by the rider and combined with angular velocity calculates power (PO = T**ω**). Strain gauges come in various types and forms but any metal in principal constitutes a strain gauge as it changes its resistance due to strain. Measuring strain related to e.g. stress, torque or force, the main concept of all strain gauges is a change in resistance of materials caused by a mechanical change in length and cross sectional area as a function of strain. In cycling, applied torque creates that strain on an object bonded to one or more strain gauges. As the object is being deformed, the strain gauge is deformed in tandem causing its resistance to change. The resistance change is related to the strain by a known factor, which is termed Gauge Factor. Manufactures of power meters generally provide the Gauge Factor, giving the sensitivity of a specific strain gauge as part of the calibration information. With a change in strain, i.e. a change in the size of the object, the resistance of the strain gauge varies. Strain is the ratio of that change in dimension of the object to its original dimension and the resultant resistance is proportional to the strain applied.

The power meters most commonly used in research are arguably the SRM power measuring crank system and PowerTap (CycleOps, Madison, USA). Both systems have been used in research studies individually or jointly as a validity and reliability reference value against an alternative power meter measuring device (SRM [28–33]; PowerTap [34,35]; jointly [36,37]). Both systems have also been subject to validity and reliability investigations but it is the SRM system which since the late 1990s has been accepted as a valid and reliable power measuring device and since has become one of the gold standards in cycling ergometry. For example Jones and Passfield [38] dynamically assessed the agreement between the PO read by the SRM technology in comparison with a standardised and, for the purpose of the study modified Monark cycing ergometer. The researchers concluded that the SRM system provides a valid method of assessing PO in the laboratory during scientific research. Martin et al. [39] in the same year also suggested that the SRM system provides a valid measure of cycling PO. During the following year

Lawton et al. [40] provided further evidence of the reliability of the SRM system. However accurate, PO values in this study were dependent on the specifics of the particular crank with measurement error ranging between 0 to 10%. More support was produced by Abbiss et al [41] who also suggested that the SRM power meter provides valid and reliable PO values [42–44].

Depending on the specific model, the SRM power meter is equipped with a set number of strain gauges. The SRM power meter is located within the crank system of the bicycle (Figure 1). The containing strain gauges which measure the applied torque have variable resistive values which change with small deformations of the detection unit. The SRM strain gauge devices are mechanically mounted inside the spider of the crank (i.e. the object) which is set between the crankshaft and the chainwheels. Force applied to either pedal distorts the strain gauges as they transfer the torque to the chain rings via the cranks. The multiplication of the rotational speed (i.e. angular velocity of the cranks) with the torque produced by the rider results in the measured PO (W). In older models, rotational velocity was directly detected via a magnet attached to the pedal arm, whilst later models use a sensor built into the crank unit, which uses an algorithm to calculate the value. The handlebar mounted power control unit of the SRM power meter provides the rider with real time feedback of the produced PO every second.

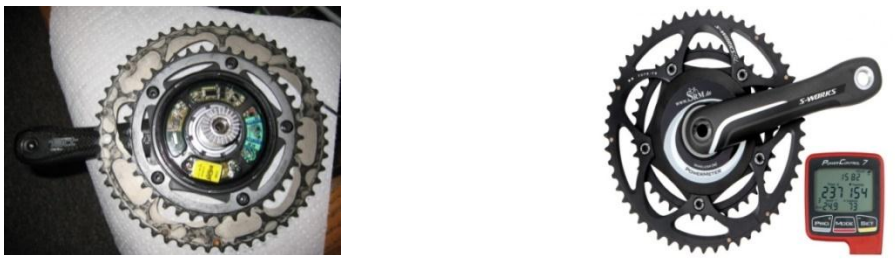

Figure 1. Example of a SRM power meter - internal and external [45].

More recently the PowerTap system has also been accepted as a valid and reliable PO measurement device, when compared to the SRM system. Bertucci et al. [44] used a road racing bicycle equipped with a PowerTap and SRM crank and tests were performed either on the road (3 h road cycling) or on a motorised treadmill at different slopes and different cycling cadences. Results demonstrated a non-significant mean -1.2 % difference in PO during submaximal constant work-rate tests between 100 W and 400 W. During submaximal incremental tests the mean error was 2.9 ± 3.3 W. Non-significant differences were also demonstrated for the 3 h road cycling tests between PO measurement devices. The PowerTap provided coefficient of variation (CoV) values for repeated submaximal incremental tests of 1.8 ± 0.6% which is highly comparable with the SRM power meter performance (CoV: 1.5 ± 0.4%). An 8% significant difference was established for 8 s sprint cycling efforts when using a small gear ratio (39/23). This difference however was not evident using a middle or higher gear ratio. For submaximal intensities between 100 W and 450 W Bertucci et al. [44] consequently deemed the PowerTap system as valid, reliable and suitable to measure PO during road cycling. In their review on cycling ergometry and mobile power meters, Paton and Hopkins [46] reported the contributions of ergometer error and biological variation to the error of measurement in a performance test. The study used a road bicycle equipped with both systems (SRM and PowerTap) which was mounted onto a Kingcycle. Cyclists had to perform three 5-min Time Trial (TT) efforts and results indicated measurement errors for the PowerTap of 1.5% and for the SRM of 1.6%. Deducting the components of cyclist error, these values were further reduced to 0.9% and 1.1% respectively.

The PowerTap uses its technology in the rear hub of the wheel (Figure 2). The hub contains a torque sensor that monitors torque 60 times per second. Equipped with strain gauges the torque sensor measures forces within the hub. Forces created by the cyclists are transmitted from the pedals via the cranks onto the chainring, which consequently creates tension on the chain. This in turn transfers the produced torque to the rear cassette and the strain gauges contained within the rear hub are deformed. A known predictable relationship between the deformation of the strain gauges and how much force is applied to achieve a quantified deformation is used to calculate resultant torque values. Like the SRM power meter, older PowerTap models measured rotational velocity directly via a magnet attached to the pedal arm whilst more recent models are

now equipped with an algorithm using rotational speed sensor built into the rear hub. The PowerTap averages force sampled and by multiplication with the rotational speed provides the rider with a second by second PO reading.

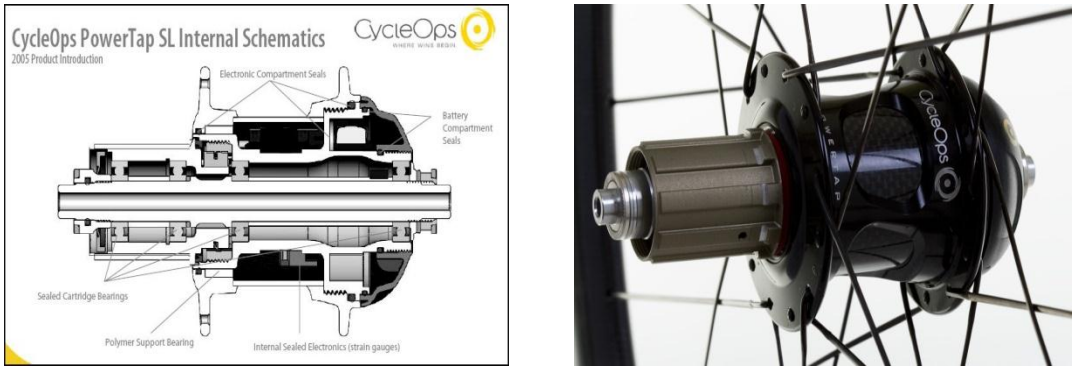

Figure 2. Example of a Powertap power meter. Internal and external [47] .

Software programmes such as TrainingPeaks (Peaksware LLC, Boulder, USA) facilitate the tracking of variables (i.e. PO or HR) over time. Such software packages also facilitate customised data analysis, such as calculation of the mean maximal PO over set time periods. These can span individual performances or use periodized training plans, allowing coaches to systematically monitor training adaptations. The development of power meters and software programmes therefore presents several new opportunities for researchers and practitioners to enhance the current understanding of real-world cycling, and to investigate novel testing methods, which are concerned with useful and associative information about current and future performance. Moreover, the development of power meters raises the relation between research laboratory and real-world by creating a theoretical possibility of ecologically valid but also reliable field-testing.

2.2 History of critical power

The seminal work of A.V. Hill in 1927 [48] plotted world record velocities over time (Figure 3), and led to the original idea about varying causes of muscle fatigue for exercise intensities of different durations. The resultant relationship formulates the principle model of performance intensity (i.e. power) and its tolerable exercise duration (Figure 4; panel A) which is still valid today [19].

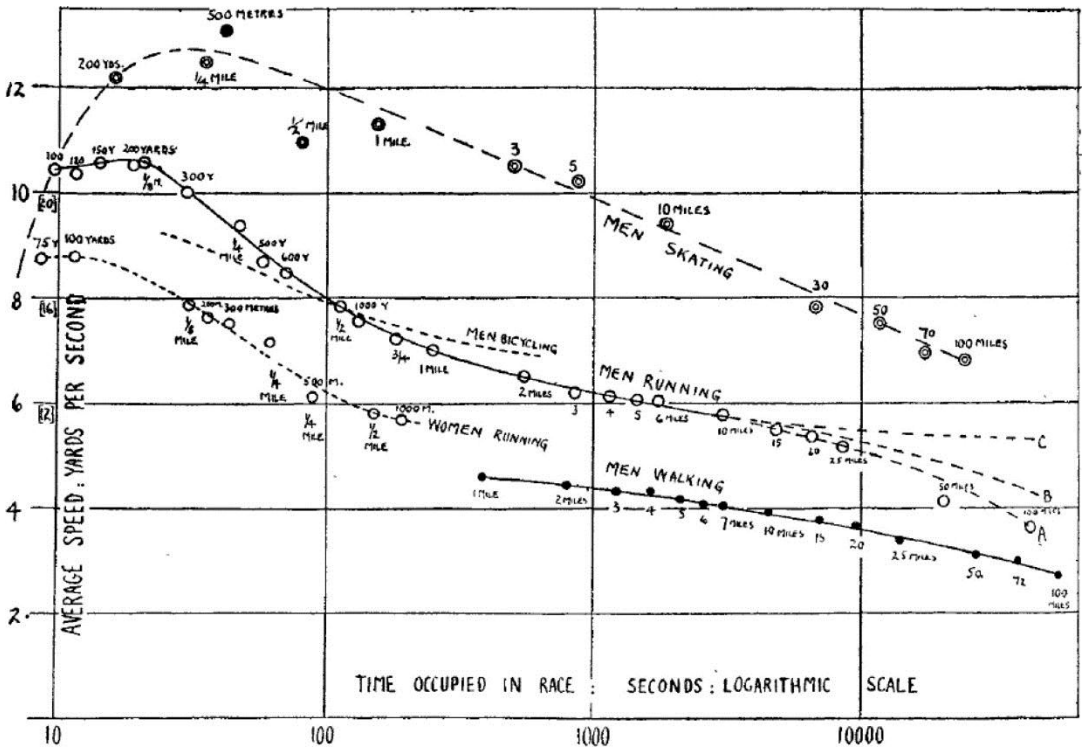

Figure 3. Plots illustration of the relationship between work-rate and exercise durations. Panel A. A. V. Hill's original plot of world record performance times versus performance speed for various sports. Taken with permission taken from Joyner and Cycle. Endurance exercise performance: the physiology of champions. *J. Physiol.* 586, 35–44 (2008) [49].

The construction of a performance velocity-time curve by Hill in 1927 provided the basis of future works by Monod and Scherrer [51], who in 1965, used isolated muscles to describe the power-duration relationship mathematically. Monod and Scherrer [51], accepted as the seminal researchers of CP, formulated that the total work performed by either one muscle or one synergistic muscle group is linearly related to its tolerable exercise duration (Figure 4; panel B). Monod and Scherrer termed the slope of this relationship CP and defined it as a 'threshold of local fatigue'. CP was originally believed to represent an exercise intensity where fatigue does not occur and, in theory, exercise is indefinitely sustainable. By definition, CP must therefore be solely aerobic in nature and unlimited in its capacity. For sports which do not allow a direct measurement of power, the analogous terms of CP are used. Generally these are either Critical Velocity (CV) in running or Critical Speed (CS) in swimming, the difference being whether the athlete performs uni- or bidirectional.

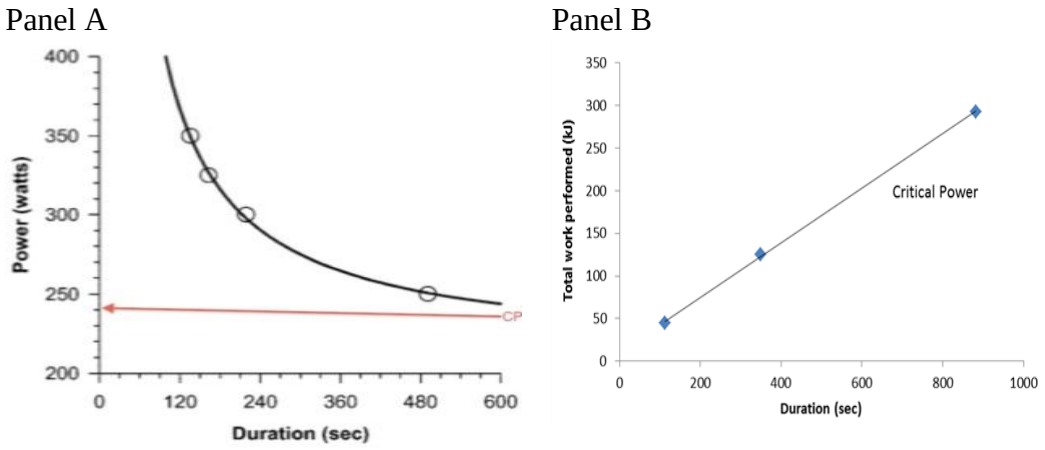

Figure 4. Exemplary illustration of Monod and Scherrer's model of critical power. Panel A illustration the relationship between power and exercise duration (Panel A) and panel B between total work performed and exercise duration CP model. Panel A illustrate a short maintainable high power as a function of tolerable time or event duration [50].

Monod and Scherrer [51] in defining the linear relationship between imposed work rate and tolerable duration also defined what they termed 'anaerobic work capacity' (AWC) which is utilised when performing at intensities above CP. The work capacity above CP is fixed and remains constant regardless of the rate of discharge. Originally, it had been

thought that this ‘anaerobic work capacity’ comprised the energy produced through phosphocreatine hydrolysis, anaerobic glycolysis, and a small aerobic contribution from O_2 stores [52–54] but more recent research has found this assertion to be incorrect. Pertinent research consequently uses the term W′ [18,55–57], which will be used consistently within the current thesis to describe.

Figure 5 schematically illustrates key physiological parameters which contribute to the determination of performance power as identified by Joyner and Coyle [49]. Whilst calculated mathematically, CP is related to a number of ‘traditional’ physiological parameters and the supporting role of CP in the determination of endurance performance was recognised by Jones and Carter [58]. The precise relationship between CP and some of these physiological parameters has been subject to a number of investigations. Smith and Jones [59] in a running study did not find a significant difference between CV and the Lactate Turn Point (LTP). Similarly Dekerle et al. [60] when comparing CP and the respiratory compensation threshold did not identify a significant difference. Furthermore McLellan et al. [61] identified a strong correlation between the individual anaerobic threshold, which was defined as the highest metabolic rate where blood [lactate] attains a steady state, and CP. Additionally Jenkins and Quigley [62] demonstrated a significant correlation between maximal oxygen uptake ($\dot{V}O_{2max}$) and CP, which was analysed before and after an 8 week endurance training intervention. Likewise, a strong relationship between CP and $\dot{V}O_{2max}$ and between CP and the ventilatory threshold (VT) was found by Smith et al. [63] when using only trained cyclists. Moreover the study demonstrated a strong correlation between CP and TT performance power. Therefore CP as a mathematical model encompasses an assessment of all physiological processes apparent in performance power.

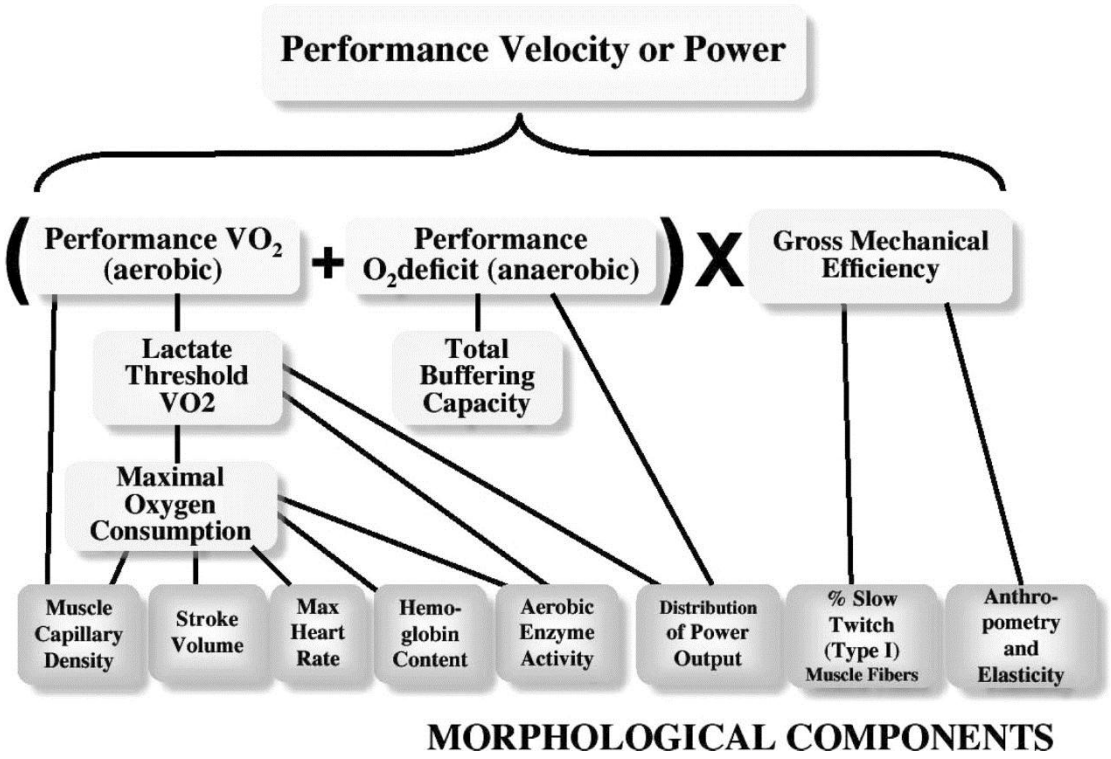

Figure 5. Schematic of the multiple physiological factors that interact as determinants of performance velocity or power output. With permission taken from Joyner and Coyle Endurance exercise performance: the physiology of champions. *J. Physiol.* 586, 35–44 (2008) [49].

The significance of CP is the variety of conditions for which it applies. It can be used:

i) As a training intensity marker [64,65],

ii) As a performance predictor [55,63,66],

iii) As a monitor for changes in endurance fitness [62,67–69],

iv) To assess the effectiveness of particular training periods [70],

v) To determine the strength and weakness of athletes [71].

CP as an endurance fitness marker has shown good test-retest reproducibility and produced high correlation coefficients of > 0.9 [72,73]. CP has also shown a positive correlation with endurance performances which last longer than the durations used in the modelling process [59,74,75].

It has to be noted however, that there are competing models which also describe endurance performance based on PO. For example Coggan and Hunter [16] developed the concept of 'functional threshold power' (FTP), which is an exercise intensity that can be maintained for 60 minutes. FTP is generally assessed either by a 60 min TT or alternatively by a 20 min TT (minus 3-5% equals FTP). Coggan and Hunter [16] further developed training zones based on FTP and whilst being popular with cyclists, very little research has been undertaken which investigates the reliability and validity of FTP or which used FTP as a variable [76].

Another competing model is that of the 'power-law', originally stipulated by Garcia-Manson et al. [77]. The power-law describes the relationship between time (or speed) and distance based on record times. Passfield et al. [78] suggested that endurance performance is better described by the power-law model. In a running study, the power-law model determined that performance prediction times for efforts of less than 2 and more than 20 min, were more accurate, compared to those predicted from the CP model. However, to date the model appears not to have attracted much research attention.

Indeed, having been the subject of a large number of research papers there are several criticisms that could be levelled at the CP model, which require careful consideration when applying the model to athletes. Two main limitations are that a) the model implies CP to be sustainable for an indefinite period of time and b) that at the onset of exercise when exercising at CP intensity the model describes energy provision to be solely supplied by aerobic metabolism [79,80]. In spite of the apparent limitations, which are discussed in more detail under heading 2.13, the justification of CP as a meaningful research topic is based on the following scientific principles:

i. CP provides good test re-test repeatability values [72,73,81]
ii. CP is a valid performance measurement [59,63,74,75]
iii. CP is more ecologically valid than a test of an isolated variable e.g. $\dot{V}O_{2max}$ [11]
iv. CP encompasses an assessment of the integrated physiology of a cyclist in a performance setting [19]

2.2.1 The power-duration concept of critical power

The CP concept and its inherent hyperbolic and linear relationship between power and tolerable exercise duration follows fundamental principles of integrative physiology and human performance [82]. It is however derived from a mathematical model and therefore not reflective of a direct physiological response. Theoretically, performing at CP results in the highest level of physiological aerobic steady state [52], where 'steady state' indicates that energy demands are met by energy supply. CP therefore represents a unique metabolic rate above which a progressive loss of homeostasis is manifested, identifying the upper limit of sustained exercise [65,67,82]. Monod and Scherrer [51], in their original work on the subject, reported CP to represent the fatigue threshold marker that cannot be determined by a single performance test. Monod and Scherrer [51] thought that CP could also be used as a marker of exclusive aerobic exercise provision and as an exercise intensity which is maintainable for "a very long time without fatigue". Other eminent researchers such as Poole et al. [65] confirmed the definition by Monod and Scherrer, as their research investigations also indicated CP to be the highest constant work rate at which steady state for which values for ventilation, gas exchange ($\dot{V}O_2$) and blood acid base status could be achieved. Along this original definition, performing exercise above CP intensity consequently causes fatigue accumulation, loss of power and eventual attainment of $\dot{V}O_{2max}$ [83].

The aerobic nature of CP has been demonstrated by manipulation of oxygen transport [84,85] and via endurance training [62]. Vanhatalo et al. [85] found the CP parameter of the power-duration relationship to be sensitive to the inspiration of hyperoxic air, as CP values were significantly higher under hyperoxic (i.e. 40% O_2) than normoxic conditions. These findings are complemented by the results of Dekerle et al. [84] who demonstrated the effects of hypoxic air (15% O_2) inhalation, as CP values were systematically reduced when compared to CP values determined under normoxic conditions. Moreover, aerobically fit participants were less affected by the reduced oxygen content. With a focus on the effects of continuous endurance training Jenkins and Quigley [62] reported a mean 31% increase in CP after an 8-week intervention in untrained participants where the exercise intensity was equal to CP.

The second parameter of the power-duration relationship, W′ (pronounced W prime) is utilised when exercising above CP intensity. The anaerobic nature of W′ has also been subject to extensive research [18,86–89]. For example, Vanhatalo et al. [85] reported a mean 18.7% reduction in W′ under hyperoxic conditions and Dekerle et al. [84] demonstrated the independency of W′ under hypoxic conditions. The latter study resulted in a non-significant difference for values of W′ between normoxic and hypoxic conditions. It is clear that utilisation of W′ causes a physiological non-steady state intensity, where the energy demand exceeds the energy supply. A depletion of W′ consequently results in either physical exhaustion (i.e. when W′ is zero), or if exercise is to be continued, the need to reduce PO to a level below CP [18,89,90]. Depending on the duration and magnitude of the decreased PO this reduction can allow for a recovery of the parameter and a return to a metabolic steady state. However, considerably less is known about the physiological underpinnings of W′ [91] and more recent findings suggest that W′ appears to reflect an athlete's ability to exercise under increasing levels of fatigue caused by its own utilisation [87]. In short, at a continuous exercise intensity above CP, the utilisation of W′ results in the accumulation of fatigue related metabolites. Furthermore, at a continuous exercise intensity above CP, the utilisation of W′ also results in the athlete having to perform under non-steady state, i.e. fatiguing conditions with decreasing power levels [87].

The ability to sustain a high PO for a prolonged period of time is one of the decisive factors in cycling endurance performance success [49,83,92,93]. Maintaining a high but tolerable PO for a prolonged period of time (i.e. at steady state) without duress is a common experience for endurance athletes. However when only marginally increasing the PO (i.e. above steady state), the tolerable duration at that PO is dramatically decreased, with fatigue accumulation consequently occurring. The PO transition point between tolerable and intolerable exercise intensity is enshrined within the power-duration relationship and based on its physiological and mathematical meaning, corresponds with the upper limit of sustainable exercise, i.e. CP. The magnitude of the PO above CP dictates the level of accumulated fatigue. In short the higher the PO above CP intensity the higher the levels of accumulated fatigue. Knowledge of the highest tolerable PO is therefore of significant value to a cyclist as it allows him/her to apply an

appropriate racing strategy, with surges of accelerations and changes of terrain or conditions (i.e. hilly, headwind, tailwind, etc) [27,89,94].

The robust concept of the power-duration relationship (i.e. it's sensitivity to manipulation of oxygen transport and endurance training) has been demonstrated in other exercise modes such as running [95], swimming [84], rowing [96], kayaking [97] football [74], canoeing [98], table tennis [99], hockey and rugby [100]. Furthermore, it has been successfully implemented in research on wheelchair athletes [101], adolescents [102–104], the elderly [105], clinical populations [106] and various animal species, such as mammals, rodents, crustacean, fish and amphibians [82,107–112].

2.2.2 Maximal lactate steady state and critical power

Previously perceived as being equal, the more recent literature has demonstrated that CP is located only approximately, i.e. above at the MLSS intensity [113]. The MLSS reflects the highest sustainable intensity without a drift in blood lactate, which is associated with the accumulation of fatigue by more than 1mM between minutes 10 and 30 of a constant load test [114]. The MLSS represents an equilibrium between blood lactate appearance and disappearance and it has a close relationship with endurance performances [115]. For example, the average velocity over a marathon is slightly below the MLSS [115]. Thus the MLSS is important as the corresponding intensity demarcates the boundary between the heavy and the very heavy (alternatively: heavy and severe) exercise domain (i.e. drift of physiological variables such as $\dot{V}O_2$ and blood lactate towards a maximal tolerable limit) [65] (see 2.3 for further discussion on exercise domains). Anaerobic glycolysis, as a provider of anaerobic energy, allows individuals to engender PO values which cannot be attained nor sustained through aerobic metabolism alone. A dynamic balance between glycolysis related lactate production and utilisation/removal has been suggested by Brooks [116]. This balance contributes to performance intensities which are at the highest physiological steady-state level, i.e. not leading to a continuous loss of homeostasis and are consequently reflected by a high, sustainable but tolerable PO value. According to Antonutto and di Prampero [117] the physiological importance of the MLSS is that it defines the exercise intensity above which the anaerobic metabolism (i.e. W′) significantly contributes to the energy supply, i.e. beyond a physiological steady-state.

Billat et al. [115] stated that the rate of metabolic adenosine triphosphate (ATP) turnover increases as a direct function of metabolic power output which is indicated by a high blood [lactate]. This suggests that individuals with high MLSS values are more likely to translate this metabolic power into high PO performance values.

Whilst the MLSS value is a good indicator of intensity [118], its estimation requires an elaborate, tedious testing methodology, in turn requiring athletes to perform an incremental maximal exercise test in addition to three to six subsequent constant work-rate tests [114]. Tests are performed on different days and blood lactate during the constant work-rate tests is sampled every 5 minutes. MLSS tests are consequently strenuous and invasive for athletes. Being an important physiological marker, a number of researchers as a result, investigated alternative MLSS testing protocols [119–124]. Despite these efforts, none seem to have fully replaced the original method and no field testing method in cycling to date has been researched. Besides the invasive and strenuous nature of MLSS testing as a function of training status or measure of performance index, the corresponding intensity (i.e. PO) related to the MLSS frequently changes in response to training [125], making this test impractical for scientists, coaches and athletes.

CP highly correlates with the MLSS [126] but does not require lactate analysis for its determination. CP is an important variable in sports science research as the determination of a sustainable PO is important for aerobic capacity diagnostic and training programme design purposes. CP as being approximately at MLSS intensity could theoretically replace the MLSS.

However, there is now compelling evidence which demonstrates that CP overestimates the PO that corresponds with the MLSS (MLSSP). For example in cycling, Jenkins and Quigley [127] demonstrated that CP is actually located within a few percentages of the MLSS but not at the MLSS intensity. Overend et al. [105] investigated the differences between elderly and young athletes when cycling at CP. Whilst the study was able to determine CP in the elderly, it concluded that CP may not represent a true non-fatiguing work rate in either population. Poole et al. [65] reported a 24 min sustainable duration when performing at CP. However, due to technical limitations, individual tests were terminated at min 24. The researchers consequently defined CP as a performance intensity of prolonged duration.

Hill and Ferguson [128] in support of Poole's et al. [65] work defined CV in running as the highest velocity at which physiological and metabolic variables can achieve a steady state. On the contrary the work by Pepper et al. [129] revealed a significant difference between corresponding intensities of CV and MLSS in running. In swimming, Wakayoshi et al. [130] stated a possible correspondence between CS and the exercise intensity at MLSS. Smith and Jones [59] whilst not finding a statistical difference between CV and the maximal lactate steady state velocity (MLSSV) cautioned that the extent of disagreement between the variables was too great to be used interchangeably. In football Denadai et al. [74] also found a difference between MLSSV and CV. CV intensities were consistently higher than those corresponding to the MLSS. Utilising an alternative MLSS testing protocol, Sid-Ali [131], on the contrary demonstrated almost identical values for intensities corresponding to CV and MLSS in running. The alternative MLSS test was based on a two step-protocol originally developed for cycling ergometry, and the corresponding values were estimated. In cycling, McLellan and Cheung [61] using the individual anaerobic threshold (IAT) as an indirect measure of the MLSSV, also found a significantly lower occurring IAT when compared to CP. Housh et al. [66] investigated the intensity associated with 60 min sustainable exercise and found an approximate 17% overestimation of CP relative to that required to complete the exercise task. The first researchers who directly and independently investigated the agreement between CV and MLSSV were Smith and Jones [59]. Five out of eight participants demonstrated CV to be higher than the MLSSV and the researchers consequently stated a tendency of CV to overestimate the MLSSV. Using only trained athletes Brickley et al. [132] also concluded CP not to represent a sustainable steady-state intensity. This population demonstrated a range of time to task failure between 20 and ~ 40 min. Using only trained cyclists Carita et al. [126] provided further evidence of CP being significantly higher than the MLSS intensity (313 W vs 287 W respectively). Based on the increasing evidence of a non-steady state intensity, Dekerle et al. [60] suggested that CP and MLSS are actually different physiological parameters and that "an accurate physiological meaning of CP is still unknown". Pringle and Jones [113] in cycling also found a significant overestimation of the MLSSP by CP and the researchers consequently suggested that these variables could not be used interchangeably. In swimming Dekerle et al. [133] also found an overestimation of the MLSS by CS and the researchers similarly stated that the disagreement was too great to use the two variables interchangeably.

Importantly Smith and Jones [59], Dekerle et al. [60,133], Pringle and Jones [113], Denadai et al. [74] and Carita et al. [126] are the only researchers who compared direct measures of both variables. All other studies used CP values as an indirect intensity measurement for the MLSS or used an alternative MLSS testing protocol. Based on the discussion presented above CP and its equivalents are located above the intensity associated with the MLSS. Furthermore CP and its equivalents represent a work-rate which does accumulate fatigue and a direct determination of the MLSS is necessary when precision is required in studies which investigate MLSS exercise.

However due the non-invasive nature, non-reliance on continuous blood sampling or expensive analytical equipment, the determination of CP might still be the more attractive option for coaches to evaluate the fitness level of their athletes. Furthermore a less elaborate, ecologically valid method, which provides an intensity marker close to the MLSS intensity, i.e. CP would be exceptionally useful for scientists, coaches and athletes. Future research could be directed towards investigating the exact magnitude of that overestimation in different sports.

2.3 Exercise intensity domains

Athletes commonly train according to intensity zones, which target certain physiological and anatomical adaptations. Physiologists generally refer to these zones as exercise intensity domains which in training terms are often sub-divided according to their specific adaptation or training characteristics (Table 2). The boundaries between intensity domains are commonly based on physiological landmarks. In sports science HR, $\dot{V}O_2$ and blood lactate responses provide three of the main physiological measurements for the description of behaviour in a particular domain with breakpoints, such as the LT or the gas exchange threshold (GET), being used as domain demarcation points. Athletes either use the direct measurement (i.e. HR) or the intensity associated with a particular boundary (i.e. LT) as an indication of training intensity. To determine individual training zones and to assign i.e. intensity values (i.e. HR or PO) for each of these, athletes are generally tested in the laboratory. As exercise is a continuum, the absolute ‘strictness’ of these demarcation markers has not been fully evidenced within the research literature to date [134].

Distinct physiological events during sub-maximal constant work-rate exercises at different intensities are well documented in the literature [65,88,135–141]. Mainly the adjunct blood lactate and $\dot{V}O_2$ response result in distinct individual metabolic and physiological landmarks but also in perceptual difference. Published research has used either three [142,143] (moderate, heavy and severe; Figure 6) or four [86,128] (moderate, heavy, severe and extreme with an alternative terminology of moderate, heavy, very heavy and severe) intensity domains. These are depicted in table 2. The terms of moderate, heavy, very heavy and severe are used in table 3 to describe specific sub-maximal domains. It is not unreasonable to argue that research has led to the creation of an additional domain, given the more recent evidence which identified CP to exceed the intensity associated with the MLSS [59,60,113]. Although narrow, this more accurately describes the boundaries between domains and does not use the MLSS interchangeably with CP (see 2.2.2).

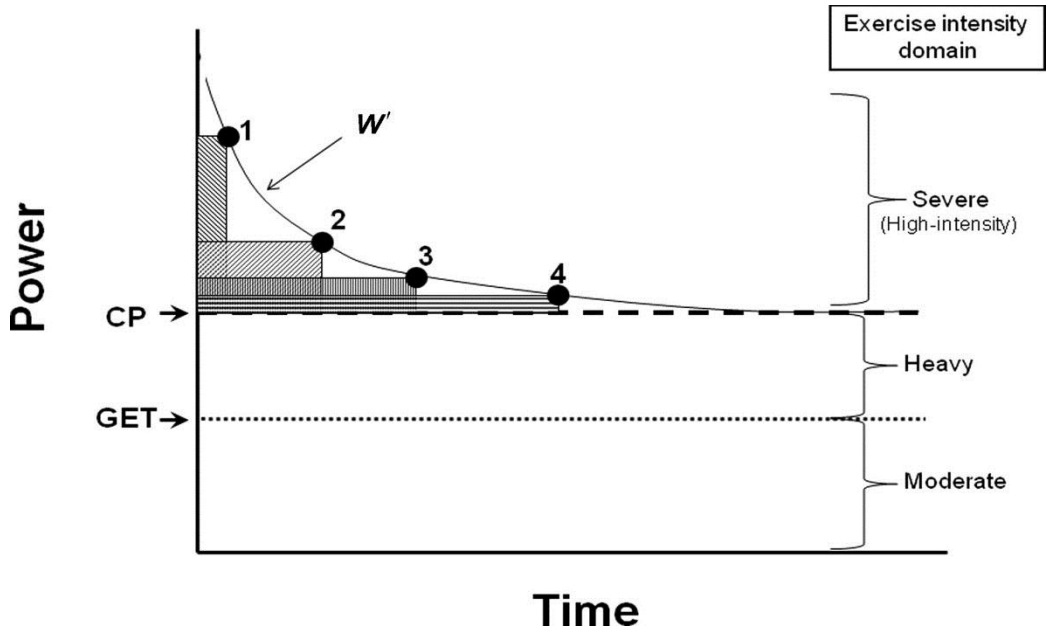

Figure 6. Schematic diagram of power versus time-to-exhaustion relationship for high-intensity exercise. With permission taken from Poole D C Exp Physiol 2009;94:197-198

Table 2. Illustration of four exercise domains

Intensities	Boundary	Predominant Energy System
Extreme	Above CP	
Severe	Upper: CP	
	Lower: MLSS	Aerobic and anaerobic glycolytic
Heavy	Upper: MLSS	Aerobic oxidative and anaerobic glycolytic
	Lower: LT/VT	
Moderate	Upper: LT/VT	Aerobic Oxidative Systems

Table 3. Training zones and exercise intensity domains

Zone	**Domain**	**Lower Boundary**	**Upper Boundary**	**Adaptations**	**Training Zone/ Session Type**
Recovery	Moderate	Rest	LT or GET	**Stress/Adaptations** Cardio-vascular stress. Capillary density, oxidative metabolism enzymes	Short rides (non-training) for recovery
Zone 1					Long rides of up to 6 hours. Economy/ efficiency and fat utilisation development
Zone 2					Long rides of up to 4 hours. Aerobic base development
Zone 3	Heavy	LT/GET	MLSS	**Stress/Adaptations** Increased cardio vascular stress Capillary density, oxidative metabolism enzymes	2-3 hours maximal. Aerobic capacity development

Zone 4	Very Heavy	MLSS	CP	**Stress/Adaptations** CV stress, increased cellular stress. Capillary density, oxidative metabolism enzymes (muscular buffering capacity)	Up to 1 hour. Race pace preparation
Zone 5	Severe	CP	$\dot{V}O_{2max}$	**Stress/Adaptations** Increased cellular stress. Improved muscular buffering capacities. Increased cardiac Output, smaller effects on capillary density and oxidative metabolism enzymes	Up to 40 min. Lactate clearance and adaptation to race speed
Zone 6	Extreme	$\dot{V}O_{2max}$	Peak Power	**Stress/Adaptations** High stress and adaptation on muscular buffering capacities. Cellular stress with less stress/ effects on capillary density and oxidative metabolism enzymes	Accumulation of 20 min. Use of high intensity interval training. Increase in maximum power and improvement of lactate production/clearance

The characteristics of the sub-maximal moderate, heavy, very heavy and severe domains are described as follows:

2.3.1 Moderate-intensity domain

The upper boundary of the moderate intensity domain is defined as both, the GET or LT [144–146]. Both terms can be used inter-changeably as they are determined by the same physiological event [147]. Exercising in this domain causes minor or no alteration in the acid-base status. An increase in glycolysis results in an elevated metabolic rate to meet, for example, the energy demands of the transitional phase (i.e. from rest to exercise). It may also cause some temporal blood lactate overshoot which if exercise is to be continued in this domain will return close to resting levels[148,149]. During this phase the ATP breakdown exceeds oxidative ATP re-synthesis and intramuscular oxygen and phosphorcreatine (PCr) stores are utilised to subsidise ATP provision. During constant load cycling exercise in this domain, $\dot{V}O_2$ increases with a gain of 9-11 $ml \cdot min^{-1} \cdot W^{-1}$ above that of unloaded pedalling [138] and healthy individuals attain a $\dot{V}O_2$ steady-state within ~2-3 min [135,139]. As metabolic variables do not drift in this domain, individuals can sustain exercise for 4 to 6 hours provided that factors such as substrate depletion [150], hyperthermia [151] and central fatigue [152–154] do not occur.

2.3.2 Heavy-intensity domain

The lower boundary of the heavy domain is defined as the GET/LT with the MLSS demarking the upper boundary [146,155]. In this domain $\dot{V}O_2$ increases continuously and reaches a delayed steady state which exceeds that predicted from the sub-LT workload relationship [106,140,156]. $\dot{V}O_2$ increases with a gain of 13 $ml \cdot min^{-1} \cdot W^{-1}$ creating an additional O_2 cost, termed "$\dot{V}O_2$ slow component", which originates predominantly within the working muscles [157,158]. The $\dot{V}O_2$ slow component has been defined as a continued rise in $\dot{V}O_2$ beyond the third minute of exercise [159]. Following a transient overshoot during initial 5 minutes of constant load exercise, blood lactate eventually stabilises at an elevated level around 2-5 mM. Despite an increased metabolic demand, healthy individuals in this domain attain a steady state within ~ 2-3 min [155] but

depending on the magnitude of the slow component this may be delayed by 10-15 min or more, i.e. when the work-rate corresponds with the MLSS intensity [138]. The upper boundary of this domain is defined as the highest $\dot{V}O_2$ at which blood lactate (and $\dot{V}O_2$) can stabilise, i.e. the MLSS [138]. Exercise in this domain is sustainable for less than 3 hours [160]. Fatigue during heavy exercise is likely due to limitations in the rate or capacity for substrate utilisation and/or hyperthermia [84,140,161] and/or neuro-muscular fatigue [154].

2.3.3 The very heavy-intensity domain

Previously accepted as (broadly) coinciding at the same intensity [127], MLSS and CP now form an additional intensity domain [162]. The lower boundary is consequently defined as the intensity corresponding to the MLSS with an upper boundary demarcated by CP [141]. Above the MLSS, the anaerobic system increasingly contributes to energy requirements resulting in a continuous upward drift of blood lactate and $[H^+]$ [117]. Blood lactate increases above the MLSS as a function of time and intensity, not attaining a steady state any longer [146]. Exercise is typically terminated when blood [lactate] reaches 8-12 mM [155]. As with the heavy domain, $\dot{V}O_2$ increases as a function of both, time and work rate. If exercise is to be continued sufficiently long enough, $\dot{V}O_2$ projects towards maximum [65,163]. The $\dot{V}O_2$ slow component can reach a magnitude of 0.5 – 1.0 $L \cdot min^{-1}$. Whilst previous demarcation points are characterised by a physiological event, CP is characterised as an indirect marker of physiological intensity. Moreover CP demarcates the transition point between tolerable and not-tolerable exercise intensity which appears to correspond to the characteristics of the slow component at this work-rate. This justifies CP being used as an additional intensity demarcation point relevant to an athlete's training and performance [64,164]. Exercise in this domain typically is sustainable for up to ~30- 40 min and it terminates at volitional or metabolic fatigue [59,113,132]. Termination might be the result of a progressive recruitment of additional fibres [165,166], neuro-muscular fatigue [167] or it occurs concomitant with the development of progressive inefficiencies within already recruited but fatigued fibres [168,169]. A cumulative fatigue effect between additional motor unit activation and metabolic fatigue was recently been supported by Sih et al. [170].

2.3.4 The severe-intensity domain

Exercise intensities within the severe domain comprise work rates located between CP and the highest work-rate for which $\dot{V}O_{2max}$ is still attainable, i.e. an intensity which is maintainable sufficiently long enough to reach $\dot{V}O_{2max}$. Blood lactate and $\dot{V}O_2$ increases inexorably to exhaustion, which occurs parallel to W' expenditure [65]. These might not reach their maximal levels if the corresponding work-rate is closer to $\dot{V}O_{2max}$ intensity, i.e. only a short tolerable time. If evident, the $\dot{V}O_2$ slow component develops after 2-3 minutes of exercise and rises as a function of time and work-rate. At the lowest work-rate (i.e. close to CP) the slow component can reach a maximal value of 1 – 1.5 $L \cdot min^{-1}$ $\dot{V}O_2$ [155]. Like in the very heavy intensity domain, if exercise is performed closer to CP intensity, blood [lactate] at exercise termination reaches values between 8-12 mM [155]. Consequently exercise is shorter than in the very heavy domain (i.e. less than 30-40 min) but long enough to attain $\dot{V}O_{2max}$ (i.e. 2-3 min). Task failure in this domain is associated with the accumulation of fatigue related metabolites (mainly P_i) and altered Ca^{2+} handling [171].

2.4 Physiology of performances around the critical power intensity

The following review on performances at, below and above CP provides consistent evidence that human and indeed some animal responses to exercise are dictated by the power-duration relationship. Further CP can be deemed as a reliable indicator for continuous activities between approximately 2 and 30 minutes [55] and hence is defined as a physiological measurement of sustainable exercise, making the determination of CP important and relevant.

Studies which investigated CP sustainability revealed not just a large inter-individual, but also inter-study variability [64,132,172–177] with a reported durations of 20 to 40 minutes. This might , as suggest by Hopkins et al. [178] be partially due to small errors in PO which can result in a much larger change in TTE duration. Additionally, CP will vary dependent on the mathematical model used for its determination [179–181] (see 2.6). Furthermore, the chosen TTE duration affects CP values obtained [79,164,182] (see 2.7.2). Finally differences might be due to a practice effect. For example, Hill and Smith [173]

found a 27% increase in time-to-exhaustion during a second trial of CP exercise performance. These limitations in accuracy have to be considered when comparing results between studies and highlight the need to standardise protocols for CP determination.

Investigating the precise physiological behaviour of CP, Poole et al. [65] recorded the metabolic and respiratory responses when exercising at both a work-rate equivalent to CP, and slightly above CP. All participants in this study were able to complete a 24-min exercise task at CP intensity. However, exercise tolerance drastically decreased by ~ 7 min at an intensity equating to 5% above CP. Measured variables such as $\dot{V}O_2$, blood [La]/[pyruvate] ratio were found to inexorably increase towards maximal values during the higher intensity task. For example, blood [La] reached values of 11.3 ± 1.4 mM with no evidence of a $\dot{V}O_2$ slow component at exercise termination. This indicated that individuals were performing at the higher end of the severe domain towards a work-rate closer to $\dot{V}O_{2max}$ intensity. Following on their earlier research, Poole et al. [67] demonstrated the effects of a 7-week intense interval cycling training on the metabolic and respiratory profile when performing at and above CP intensity. CP remained unaffected following the intervention but values for W′ and LT were significantly increased. $\dot{V}O_2$, blood [La] and pH eventually reached stable levels pre and post training intervention when performing at CP intensity. Above CP intensity with no attainment of $\dot{V}O_{2max}$ values, a progressive increase in blood [La] and a decrease in pH was evident. Poole et al. [67] consequently concluded that CP represents an upper limit of exercise intensity at which $\dot{V}O_2$, blood [La] and pH eventually stabilizes, whilst any intensity performed above CP results in imminent fatigue.

Interestingly, using highly trained endurance runners, Billat et al. [183] demonstrated a maximal tolerable duration of ~ 17 ± 4.4 min without the occurrence of a $\dot{V}O_2$ slow component, but a $\dot{V}O_2$ steady-state attainment when performing 5% above CV intensity. Runners reached exhaustion before reaching their $\dot{V}O_{2max}$ values. When comparing their results to those of Poole et al. [67] or Roston et al. [184], Billat et al. [183] speculated that disparity in findings were due to the different exercise modes (i.e. running vs. cycling) and/or to the population groups used (i.e. highly trained runners vs. physically active or

inactive subjects). Similar to Poole et al. [65] but using only trained cyclists, De Lucas et al. [64] investigated the pulmonary, ventilatory and blood [La] responses when cycling at CP and 5% above CP intensity. Individuals were able to sustain durations of 22 ± 7.5 min when performing at CP but only 13.3 ± 5.8 min when performing at the higher intensity. Physiological variables obtained from above CP intensity tests were significantly higher compared to those obtained from the equal CP intensity tests.

Jones et al. [18] used 31P-Magnetic resonance spectroscopy (^{31}P-MRS) to investigate the muscle metabolic response and fatigue mechanisms when performing at 10% below and 10% above CP using single leg knee extension. When performing below CP intensity, all measured variables stabilised within 3 minutes and values remained stable until the completion of the 20 min exercise task. After an initial rapid decrease, [PCr] stabilised at a 75% baseline value, with [Pi] exhibiting an initial temporary rise but stabilised within 1 min. During an initial transient time of the first minute of exercise, values for pH increased before reaching a maximal decrease at ~ 3 min which was followed by a slight recovery value until the end of exercise. This end exercise pH value was similar to that recorded at resting stage. All individuals were able to complete the task without duress at this intensity. Contrarily, a progressive loss in homeostasis was evident for the 10% above CP trials. At exhaustion [PCr] was decreased down to 26 % when compared to baseline value with [Pi] increasing more rapidly during the task. At min 6, values for pH reached levels which are generally observed at the exhaustive stage of high intensity exercise. Individuals were able to sustain this high intensity exercise for a duration of 14.7 ± 7.1 min. Jones et al. [18] consequently defined CP as the highest possible constant work-rate which does not exhibit a progressive depletion of high-energy phosphates and the accumulation of fatigue related metabolites.

Brickley et al. [185] investigated the metabolic responses of 30 min constant load exercise versus a 30 min oscillating protocol in trained athletes. By averaging the same mean PO, the researchers found no significant differences in pH, muscle [La] and muscle [glycogen] between the protocols. Participants in this study had to perform one alternating task pattern of cycling for 30 s at 158% CP and 120 s at 73% of CP in addition to one 90% CP constant work-rate task. The main findings were similar muscle metabolic responses between the protocols and that the alternating intensity protocol did not result in greater metabolic perturbation when matched for total work performed. Concentrations of muscle metabolic variables changed significantly during both patterns

of exercise, which was assessed using vastus lateralis muscle biopsy samples pre, mid-point and post exercise. Brickley et al. [185] indicated three main possible reasons out which either individually or jointly caused the lack of change in muscle metabolic response for the alternating exercise protocol. Reason one suggested that the 2 min low intensity recovery duration might have been sufficiently long enough for a full restoration of i.e. [PCr]. Reason two suggested a possible dampening effect of the metabolic response to a lower intensity exercise bout after a higher intensity bout and reason three suggested a possible dampening effect of the metabolic response to prior alternating bouts of exercises. The authors however highlighted some major methodological limitations and cautioned researchers to carefully consider theses results, which were based solely on statistical differences. These findings imply that the parameters of the power-duration relationship, if matched for total work performed will under intermittent exercise conditions result in an equal metabolic response. If considered correctly this might be highly relevant for pacing strategies, an important factor in successful road cycling races [49,186].

Technical and ethical limitations led to more research being performed on animals. Deeming rats as a valid and reliable model to represent human responses, Armstrong and Laughlin [187–189] during the 1980's started to investigate inter- and intramuscular hind-limb blood flow responses to running exercise. Based on this model, Copp et al. [82] observed the blood flow in rats hind-limbs. Using a treadmill, rats performed a maximal incremental running test in addition to five constant load tests for the determination of CV. Consequent TTE runs were performed at ~15% above and ~15% below CV and hind-limb blood flow was measured using injected radiolabelled microspheres. Runs performed at the below CV intensity were ~ 5 times greater (~ 45 min), than those performed above CV intensity (~ 10 min), confirming CV as a marker of the upper limit of sustainable exercise performance in rats. The main findings of the study during the higher intensity TTE trial were a significantly elevated blood flow to the total hind-limb skeletal muscles with ≥ 69% of a 35% total increase being distributed to glycolytic type IIb/d/x fibres. Type I fibres during the high intensity exercise could not match the energy demand and without a significant increase in blood supply, fibres progressively accumulated fatigue. The relative greater distribution of blood flow to less efficient [190–193], higher threshold glycolytic [194,195] fibres also indicated a disproportional increase in type IIb/d/x fibre recruitment. Copp et al. [82] consequently confirmed CV as a

unique metabolic rate in rats above which an inherent progressive instability of oxidative metabolism is exhibited. The researchers concluded that a concomitant effect between progressive metabolic inefficiencies within recruited fibres and the progressive increase in motor unit recruitment composed of type IIb/d/x fibres as an underlying mechanism of the $\dot{V}O_2$ slow component.

2.5 Physiological meaning of W′

The following review summarizes a to-date unclear understanding of the exact nature of W′. Without such clear understanding, Dekerle [196] advises prudence when interpreting the value of W′ and its changes over training. Furthermore, with researchers such as Gaesser et al. [197] stating an inherent difficulty in determining W′ accurately or reliably, further works which investigate the true dynamics and content of W′ are required. Reported W′ values in individual studies of this thesis did not lead to conclusive outcomes, which justified the decision of using CP alone as the overarching research subject.

A central component of road racing success is the ability to produce high PO values during short periods of time which can produce tactical advantages [49,186]. It is not uncommon to produce somewhere between 20-70 sprint efforts above the Maximal Aerobic Power (MAP) in cycling road races [198]. The energy supply for efforts of such high PO values (i.e. > CP) is accounted for by the parameter of W′. W′ has been subject to a number of recent investigations [56,57,86,87,89,199] and it continues to raise fundamental questions about its underlying physiology [18,199,200]. Increasing as the result of interventions such as creatine supplementation [56,201–203] or high-intensity strength and sprint training [68], W′ is reduced after glycogen depletion [203], prior high intensity exercise with limited recovery [204] and it remains unaffected by interventions such as hypo- or hyperoxia [84,85]. Recently Sawyer et al. [205] reported an increase in W' after an intervention of strength training. Contrary to these Clark et al. [206] when applying high intensity interval training found an improvement in CV but a decrease in ARD.

W′ as a fixed capacity of anaerobic work is equivalent to the total work performed minus the work derived from aerobic metabolism, represented by CP [207,208]. Like CP, W′ is also subject to a key assumption embedded in the modelling of the power-duration relationship and discussed further under heading 2.6. Whilst W′ is represented as a simple mathematical value, expressed in Joules (J) or kilo Joules (kJ), it is reflective of some physiological variable or variables and together with CP defines tolerable exercise durations [87]. However its physiological base to date remains controversial. A number of authors have considered W′ to be synonymous with the maximal O_2 deficit (MAOD) [54] or the anaerobic work capacity [52,79]. According to Moritani et al. [52] and Poole et al. [65], W′ is reflective of a finite anaerobic intra-muscular energy store comprised of oxygen bound myoglobin stores, glycogen and high energy-phosphates. W′ is depleted at a rate somewhat proportional to the magnitude of the power requirements above CP, making W′ capacity, not rate limited. Whether W′ reflects a finite amount of energy store or a more recently proposed build-up of fatigue related metabolites, such as H^+, di-pronated inorganic phosphate ($H_2PO_4^-$) and potassium ions (K^+) to some tolerable level [199,200,209] is a research question which has attracted some considerable attention [86–89]. The following review demonstrates to-date research findings attempting to eliminate the underlying physiology of W′.

Historically W′ was perceived as comprising energy derived through substrate-level phosphorylisation utilising PCr and glycogenesis with a small aerobic contribution from myoglobin and haemoglobin bound O_2 stores [19]. Investigating PCr recovery kinetics using ^{31}P-MRS analysis , Forbes et al. [210] conducted their research on repeated bouts of heavy exercise, separated by either 3 min, 6 min or 15 min. There was no difference between the on-transient time constant of the PCr primary component between repeated exercise bouts. However, the amplitude of the PCr slow component and the total PCr breakdown were reduced in each of the subsequent bouts. In a follow up study, Forbes et al. [211] examined PCr recovery kinetics in humans and rats after low and high intensity exercise bouts. Contrary to rats, humans demonstrated a single-exponential PCr recovery component after repeated bouts of low intensity exercise, indicating predominantly oxidative metabolic recovery kinetics. After high intensity exercise PCr recovery kinetics exhibited a prevalent initial fast component, indicating a greater reliance on glycolytic ATP production towards PCr re-synthesis in both, humans and rats. Results

further suggested that the heterogeneity of the oxidative capacity among skeletal muscle fibres in humans does not contribute to a higher-order PCr recovery pattern and that glycolytic ATP production are part of PCr recovery kinetics. PCr breakdown and recovery kinetics therefore contribute to the behaviour of W′ during high intensity exercise.

Jones et al. [18] demonstrated that W′ is utilised at a predictable rate during sustained exercise above CP. The magnitude above CP of that predictable rate determines the tolerable duration of that intensity and will lead to W′ depletion. Following this, exercise cannot be tolerated at the same intensity and has to be reduced below CP. Consequently, the higher the sustained PO above that of CP, the faster the expenditure of W′ and the greater the rate of fatigue related metabolites accumulating such as P_i, ADP, H^+, and extracellular K^+.

Investigating the physiological meaning of W′, Ferguson et al [87] suggested that a rate determined W′ utilisation is coupled somewhat proportionally with the rate of fatigue related metabolite build up. Based on finding by Rossiter et al. [212] the study used $\dot{V}O_2$ and arterialized capillary blood [La] as proxy for intramuscular PCr kinetics and lactate recovery kinetics respectively. Attempting to elucidate whether there is a linear or some more complex W′ utilisation function and build-up of fatigue related metabolites, Ferguson et al [87] questioned to which degree blood [La] and PCr kinetically correlate with W′. Individuals were required to perform a total of four CP tests. The first test involved four TTE trials for conventional CP and W′ determination. All TTE trials in the remaining 3 CP tests were preceded by a 6 min exhaustive exercise bout followed by recovery baseline cycling phase at 20 W for 2 min, 6 min or 15 min. This was the first study investigating the putative physiology of W′ and results demonstrated, that contrary to the findings of Coats et al. [199], CP was not affected by a prior bout of exhaustive exercise targeted at W′ depletion. However all TTE trial durations which followed the exhaustive exercise bout were shorter in durations and depending on the recovery protocol resulted in systematically and significantly lower W′ values. $\dot{V}O_2$ recovery kinetics were appreciably faster than those of W′, resulting in given $\%\dot{V}O_2$ recovery being associated with a much smaller % W′ recovery. On the contrary arterialised

capillary blood [La] recovery kinetics were slower than those of W′ but no clear proportionality in the magnitude of relative recoveries was evident. This suggested that blood [La] recovery after full recovery of W′ was still continuing. More importantly, this by deduction, excludes intra-muscular [La] clearance as the exclusive mediator for W′ restitution. Ferguson et al. [87] concluded that W′ is unlikely to represent a simple 'depletable' anaerobic energy store as its complex recovery kinetics seems to reconstitute in a curvilinear manner. The study suggested that W′ is better represented by the integrated action of variables that contribute to the process of fatigue via accumulation of key metabolites, such as Pi extra-cellular K^+.

Following these findings and based on a 3-min all-out CP test, Skiba et al. [89] modelled the energy expenditure and reconstruction of W′ during exercise above CP for intermittent exercise over a range of recovery PO, using a three domain scale of moderate, heavy and severe exercise. The research utilised the $\dot{V}O_2$ slow component as an indicator of W′ expenditure, as demonstrated in the literature [86,213,214]. In particular, a possible dynamic temporal relationship between $\dot{V}O_2$ and W′ charge/discharge was investigated. Based on the key assumptions that W′ expenditure starts the moment exercise intensity exceeds CP, and that its reconstitution follows a predictable exponential time course, Skiba et al. [89] illustrated W′ kinetics with the development of a new mathematical model. Individuals had to perform an exhaustive exercise bout in the severe domain followed by intermittent exhaustive exercises at various intensities interspersed with 30 s moderate, heavy or severe recovery intervals. Results demonstrated a linear correlation between the rise in $\dot{V}O_2$ during each successive heavy interval bout and the modelled W′ net discharge. As the $\dot{V}O_2$ slow component is suggested to be linked to type II fiber recruitment [215], Skiba et al. [89] also proposed this exact link as related to the parameter of W′. Like Ferguson et al. [87], using $\dot{V}O_2$ as a proxy for PCr kinetics the study explained the progressive increase in $\dot{V}O_2$ with the W′ discharge by an associated fall in [PCr] [216,217]. Skiba et al. [89] consequently suggested W′ to be primarily representative of the relative fatigue and recruitment state of the type II fibre pool and that type I and type II fibres contribute to the depletion of W′ above CP in unequal proportions. However the absolute sum of W′ expenditure by both fibre types at exhaustion always has to equal W′.

Based on the 3-min all-out CP test, Parker-Simpson et al. [57] examined the influence of different initial metabolic rates on CP and W′ results. Individuals had to perform one all-out test without any prior exercise bout to determine CP and W′. The study also used a three domain scale of moderate, heavy and severe intensity. Prior to commencing the all-out test, participants had to perform prior exercise bouts at various intensities (6 min moderate and 6 min heavy exercise and 2 min and 4 min severe exercise). CP results were not affected by any of the prior exercises, indicating that CP is independent of different initial metabolic rates. Only the severe-intensity protocol demonstrated a significant difference in W′ when compared to the no-prior exercise value of W′. Explanations for the smaller magnitude of W′ were related to Fitts' [218] findings in that the decrease in W′ is reflective of the level of accumulated fatigue related metabolites (i.e. H^+, P_i and extracellular K^+). PCr and muscle glycogen stores are simultaneously depleted [19]. Even though W′ was not fully depleted, peak PO values during the all-out test were reduced after the prior severe exercise bouts. Parker-Simpson et al. [57] stipulated that the reduction of W′ was due to the fatigue of type II fibres initiated by the prior severe exercise bouts. Most importantly the study demonstrated the independence of CP from all factors which are seemingly detrimental to W′. Supporting these findings Johnson et al. [219] demonstrated similar results for upper body exercise. Also using the three domain scale of moderate, heavy and severe intensity, the study examined the effects of prior severe exercise on the power-duration relationship in arm cranking. For the determination of CP and W′ participants had to perform four TTE trials with and without a bout of prior severe exercise. Measuring ventilation response, blood [La], [H^+], [bicarbonate] and [K^+] for the TTE trials, the study whilst finding a significant reduction of W′ in prior severe exercise, at the same time did not identify a difference for CP. The authors concluded that the magnitude of W′ following severe upper body exercise is partially dependent on the level of prior fatigue inducing metabolite accumulation.

Most recent investigations seem to agree on an interaction between W′ expenditure, W′ replenishment and the reflection of W′ on an athlete's ability to exercise under increasing levels of fatigue caused by its own (i.e. W′) utilisation, rather than 'just' a finite amount of energy.

Coats et al. [199] addressed the uncertainty of physiological determinants of W′ by questioning if W′ is replenishable after a bout of exhaustive exercise (i.e. depletion of W′) when followed by a set duration exercise bout at 80%, 90% and 110% CP. The research was based on Fukuba and Whipp's [220] suggestion that exercise after depletion of W′ is only sustainable at an intensity below CP, i.e. a predominantly aerobic metabolism driven intensity. When performing at 110% CP post W′ depleting bout, individuals were only able to sustain ~ 30 s of exercise but all participants completed the following 20 min exercise task when performing at an 80% CP intensity [199]. Surprisingly, only two individuals completed the 20 min exercise task at 90% of CP, with four individuals reaching fatigue at submaximal ventilatory and respiratory responses compared to those seen in a prior maximal incremental test. Speculating on these diverse results, Coats et al. [199] suggested that after depletion of W′, exercise is only sustainable at a "wholly aerobic" rather than "simply below CP" intensity. However, the study clearly demonstrated that the severe intensity domain (i.e. > CP) is characterised with a progressive increase in metabolic drive.

The CP concept implies that W' does not supply energy during exercise at an intensity equal to CP and that it is possible to fully deplete it. Firstly, this does not consider oxygen kinetics as aerobic inertia delays an immediate steady-state response and at the onset of exercise energy supply is supported by anaerobic metabolism. Moreover Gastin [221] in his review suggested that during high intensity exercise trials lasting 2 – 15 min a high percentage (~ 60%) of energy contribution originates from aerobic metabolism. Secondly as demonstrated for example by Jones et al. [18], muscular [PCr] only fell to 27 ± 17% baseline value after a fatiguing above CP intensity bout of exercise. Whilst [PCr] continued to fall throughout the duration of the exercise it did not reach a fully depleted value. The CP model does not consider this more complex integration of aerobic and anaerobic energy supply during high intensity exercise but uses a clear compartmentalisation in its mathematical base.

2.6 Mathematical modelling of critical power

2.6.1 Two parameter models

The following section reviews the theoretical and mathematical basis of the power-duration relationship. It further reviews the link between the parameters derived from this modelling process, i.e. CP and W′.

A.V. Hill [222] firstly described the relationship between intensity and tolerable exercise duration. Later, this relationship was expressed by Monod and Scherrer [51] as:

$t = W'/P\text{-}CP$ (equation 1 – hyperbolic model)

In this equation, t = the tolerable duration (time to fatigue), W′ = a finite amount of energy reserve which is expressed in kJ, P = power output. CP is represented by the power asymptote (Figure 7, panel A). This model characterises CP as when exceeded leads to exhaustion in a predictable duration defined by the finite amount of energy, W′ [80,91].

Equation 1 can be transformed from a hyperbolic into a linear relationship which expresses the total work performed in relation to the tolerable duration of this work. This linear relationship is expressed as:

$P \cdot t = W' + (CP \cdot t)$ (equation 2 – linear work-time model)

In this equation $P \cdot t$ = the total amount of work performed (originally termed ‘limited work’; Wlim), t = time to exhaustion (originally termed ‘limit time’; t_{lim}). CP is denoted by the slope of the line and W′ is represented by the y-intercept (Figure 7, panel B).

Moritani et al. [52] later added the linear power-inverse time two parameter model which is expressed as:

$P = CP + W' + 1//t$ (equation 3 – linear power- 1/t model)

Here, CP is represented by the y-intercept with W' being presented by the slope of the line (Figure 7, panel C).

Each of the two-parameter models produces slightly different CP values. The linear work-time model derives the highest CP values and the hyperbolic model the lowest CP values [95,180,181]. In the hyperbolic model (Figure 7, panel A) the trapezium rule might explain an underestimation of the asymptote (i.e. CP) and an overestimating of the area under the curve (i.e. W′). TTE in the linear models is located on the x-axis, whilst work (Figure 7, panel B) is a function of TTE multiplied by mean PO. Together these individual differences produce slightly different CP values using the two-parameter models.

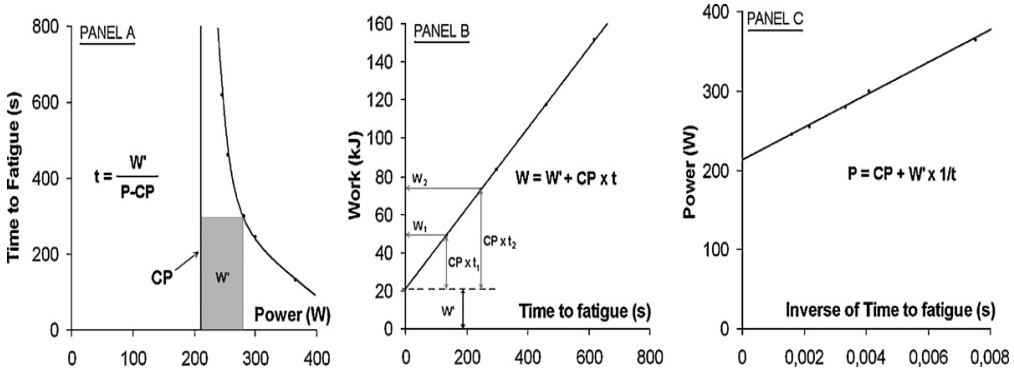

Figure 7. Schematic representation of the two-parameter critical power models. (hyperbolic time-power relationship– Panel A; linear work-time relationship – Panel B; linear power-1/t relationship – Panel C). Taken with permission from Dekerle et al. Validity of the two-parameter model in estimating the anaerobic work capacity. Eur J Appl. Phsiol. 96(3), 257-64 [223].

The power-duration relationship of the CP concept is based on a number of additional key assumptions, which provide only a simplified model of all processes that occur during high intensity exercise. Even though these assumptions contain shortcomings to varying degrees , the model has still been accepted as being robust [80,91] .

These key assumptions of the two-parameter models are as follows [79]:

1. There are only two sources of energy supply in humans, aerobic and anaerobic metabolism.
2. CP is aerobic in nature and it is rate but not capacity limited.
3. W′ is anaerobic in nature and it is capacity but not rate limited.
4. Depletion of W′ results in exhaustion and per definition exercise termination.

There are additional assumptions embedded in the above, which have been described by Morton [80] as follows:

5. CP is attainable right at the onset of exercise and it is sustainable for the entire duration of the exercise. At the point of exhaustion it coincides with the depletion of W′.
6. The power domain over which the model applies is all of $CP < P < \infty$. This implies that the anaerobic energy supply is never required, since if $P \leq CP$ the energy demand is instantly fully driven by the aerobic metabolism. Moreover the assumption also implies an unlimited (∞) magnitude of power production.
7. The time domain over which the model applies is all of $0 < t < \infty$ and that endurance at CP is indefinitely long. Even at a moderate PO, endurance time is not infinite. The model does not consider psychological or nutritional aspects, which will eventually require the athlete to terminate the performance. Similarly endurance time cannot be zero even if exercising at a maximal instantaneous PO.
8. Exercise efficiency remains constant across all power and time domains.
9. CP and W′ are constants and independent of P (and/or of t)

In summary, when applying the CP concept to cyclic activities consideration has to be given to these assumptions as they suggest a number of unrealistic physiological and mechanical scenarios [80,208,224].

CP was first believed to be indefinitely sustainable [51] which was a misinterpretation of the mathematical rather than the physiological definition. The hyperbolic and linear power- 1/time models dictate that the higher the PO, the shorter the time to exhaustion. Consequently time, when approaching zero can produce an infinitively high PO. Equally CP can also be performed at zero time when using the power-1/time model. If the given quantity of work is less than or equal to W′, the work-time relationship contrarily implies that it is possible to complete this work in zero time. Whilst this might be acceptable mathematically, such situations are not supported by the limits of human performance. Monod and Scherrer [51] stated that the work-time relationship loses linearity when performing constant load exercises of shorter than 2 minutes durations which offers some explanation for these assumptions. The loss of linearity can also be explained by a change in energetic cost over the range of TTE durations [208,224].

The power- 1/time model further suggests an 'infinite' duration when exercising below or at CP, as the energy supply is solely met by aerobic metabolism [19]. Fatigue will always occur, which compromises endurance exercise [225]. The assumption in the power-time model implies that as exercise time approaches zero, the maximal power output nears infinity. The contrary is true for the linear work-time model. Power production is limited, as total work done cannot be less than W′.

Physiologically, all CP models assume the unrealistic condition in that $\dot{V}O_{2\,max}$ is attained right at the start of exercise. In the non-elite athlete a minimum of two minutes is however required to attain $\dot{V}O_{2\,max}$ [226]. The slope of the line and the y-intercept consequently always overestimate and underestimate the true values of CP and W′ respectively [223]. However in consideration of this assumption during each trial, only TTE durations should be chosen which allows athletes to attain their $\dot{V}O_{2\,max}$ value [208] . This consequently implies that all TTE trials are located within the severe domain as the intensities are sufficiently high enough not to reach a physiological steady-state whilst being able to attain $\dot{V}O_{2\,max}$. The assumption of W′ being independent of exhaustion times and being depleted at the end of each TTE trial is likely to be true during exercises which attain $\dot{V}O_{2\,max}$ [224]. However, testing this assumption remains difficult as W′ is a theoretical construct which contains a high level of measurement error [227]. Furthermore the CP concept is based on classical fatigue theories, where a loss in power for example

is caused by an excitation-contraction coupling failure or impaired cross-bridge cycling. These classical theories solely consider peripheral and not central fatigue, which generally results in a change of central motor drive [228,229]. Billat et al. [70] and Morton [80] criticised the simplicity of CP model to describe highly complex energetic processes which are apparent during exhaustive exercises.

2.6.2 Three parameter models

Addressing some of the shortcomings of the two-parameter model, several researchers added a further parameter which resulted in the construction of a number of three-parameter models. For example, Morton [230] in response to a lack of limitation for a highest PO introduced 'instantaneous maximal power' (P_{max}) as a third parameter into the model. P_{max} can be exhibited at any instant and it is proportional to the amount of W′ remaining at that instant, depending on whether W′ is fully intact or fully exhausted. P_{max} can consequently equal the magnitude of maximal power or it can be equal to CP. However, Chatagnon et al. [231] demonstrated that P_{max} whilst giving the power-duration relationship a more accurate description, did not provide better correlations of CP and W′ with selected physiological variables than those resultant from the two-parameter models.

Based on a delayed aerobic response between the onset and attainment of steady-state exercise, Wilkie [232] proposed a correction by adding a time constant (Tau; τ) to the model, which considers the O_2 deficit. Using these 3-parameter models, CP tends to be lower and consequently is more physiologically sustainable [95,180] . The model also provides information about a maximal power production [80,179,180,197]. Whilst offering a greater level of CP sustainable accuracy, the model is more difficult to use, mathematically not straightforward and consequently of less use to coaches and athletes.

Chatagnon and Busso [233] introduced the segmented CP model by adding another correction, a second threshold (P_t) to (τ) which corresponds to the lowest power required to achieve the MAP. In this model CP remains the power asymptote for time tending towards infinity and assumes a continuous anaerobic power contribution to the energy production for PO greater than CP. The revised model assumes that the anaerobic

metabolism partially contributes to the total energy production for exercises between CP and P_t (where the O_2 demand does not necessarily exceed $\dot{V}O_{2\,max}$.), whilst exclusively provides the energy required when exercising above P_t. A parameter 'α' which accounts for the anaerobic metabolism contribution of power values between CP and P_t was also added. This extended derivative of the hyperbolic model shows a substantial contribution to the total energy production in the range between CP and P_t but the model limitations can lead to an overestimation of the anaerobic contribution [234]. Depending on the chosen exercise intensity it can further provide information on the CP and W′ contribution but as highlighted by Busso et al. [234], more studies are required, which investigate the change in efficiency which is apparent in the energy transformation using both metabolic and mechanical power across different exercise intensities and durations.

In running [235] and in cycling [182] Hill et al. compared the two-parameter with the three-parameter model to evaluate the appropriateness of either. The studies used exhaustion times between ~3 and 10 minutes and for both exercise modes the two-parameter described the relationship between velocity or power and time to fatigue well. However only the two-parameter model produced parameters of known physiological significance and CP/CV values associated with low Standard Error of Estimate (SEE) values. Using the three-parameter model resulted in high SEE values and/or unrealistic CP/CV values of no obvious physiological meaning. Hill et al. [182,235] consequently stated a preference for the two-parameter model.

A further analysis of models is beyond the purpose of this thesis, and can be reviewed elsewhere [80,179,180,233,234,236,237]. However, it should be noted, that CP values derived from non-linear two- and three-parameter models commonly result in CP which are 15-40 W lower, than those derived from linear two-parameter models [180,197]. Even with a reduction in CP value exercise duration is still not indefinite as fatigue is always imminent [218,225,238]. Research using the former models for CP determination consequently reported different outcomes for i.e. $\dot{V}O_2$ response [239] as $\dot{V}O_{2\,max}$ was not necessarily attained when performing at that particular CP intensity.

In short, according to di Prampero [208] the two-parameter linear models explain the work-time and power-1/time relationship only for intensities eliciting $\dot{V}O_{2\,max}$. Whilst

containing a number of unrealistic assumptions, the two-parameter models provide coaches with a useful testing tool. Independent of the linear relationships the two-parameter model is appropriate when describing and predicting exercise tolerance for performances above CP or CV/CS [59,172,240,241].

The above discussion demonstrates that CP and CV have attracted a significant amount of research interest not just over the past decade. Factors affecting the measurement of CP are discussed in the following section which, together with the above literature review, results in a strong argument for CP requiring standardised modelling and indeed determination procedures. This requirement contributed to the addressed research questions in this thesis and lent support to the studies in which a specific methodology was utilised throughout individual studies.

2.7 Factors affecting the determination of critical power

2.7.1 Effect of ergometer and cadence

In his review, Hill [79] emphasised that altered CP results can be caused by potential errors in the choice of ergometer and choice of cadence. When using a manual ergometer a tendency for changes in cadence is apparent. Therefore reported PO values and actual PO values are not always equal during constant load tests, unless the exact cadence is maintained. Hill [79] consequently recommended the use of electronically controlled ergometers, where the PO can be set independently of cycling cadence.

Metabolic efficiencies and $\dot{V}O_2$ at imposed PO values are sensitive to different cycling cadences [242,243] and the manipulation of cadence can directly affect both the shortening velocity [244] and the recruitment of different muscle fibres [245]. Due to increased muscle force requirements at high resistances, Ahlquist et al. [245] demonstrated that type II fibre fatigue occurs at a greater rate when cycling at low cadences. Whilst a slow cadence requires greater force production, a fast cadence requires a faster contraction velocity [246] which also has been shown to recruit type II fibres [245]. Therefore the choice of cadence can directly affect the power-duration relationship. Carnevale and Gaesser [247] using two different cadences (60 revolutions per minute [RPM] and 100 RPM) investigated the impact of values on CP and W′. At 60 RPM, CP was significantly higher (235 ± 8 W)

with no significant difference in W′ (18.9 ± 2.2 kJ) when compared to a higher cadence (204 ± 11 W; 16.8 ± 1.7 kJ). The lower CP values were speculated to result from lower produced pedal forces and from a greater cardio-vascular and blood lactate response. Furthermore the authors suggested a theoretical maximal sustainable PO in untrained men to be greater at a cadence of 60 RPM. A similar study was later performed by Hill et al. [248] who added a third cadence into their investigation. TTE trials were performed at 60 RPM, 100 RPM and at a self-selected cadence. Results were similar to those found by Carnevale and Gaesser [247] as CP results derived from the 100 RPM trials were significantly lower than those from the lower and the self-selected cadence trials (195 ± 50 W, 207 ± 50 W and 204 ± 48 W respectively). Interestingly, Hill et al. [248] also reported higher average values of W′ using the self-selected cadence trials than those from the higher and lower cadence trials (16.1 ± 6.2, 14.5 ± 5.9 and 14.6 ± 5.7 kJ respectively). Increased pedal force therefore appears to a have a greater influence on type II fiber recruitment than a fast contraction velocity. However it is not unreasonable to argue that trained road cyclists, given the choice, prefer higher cadences as fatigue occurs at a lower rate when utilising a higher portion of type I fibers [245]. Trained cyclists with common gear ratios generally self-select a higher cadence range between 70 RPM and 100 RPM [249]. Lepers et al. [249] demonstrated this as the drop in mean self-selected cadence from 89 RPM down to 69 RPM resulted in an increase in energy cost in trained cyclists. The study comprised of a 2-h constant power performance which was set at a 85% MAP intensity.

Similarly, in a group of recreational athletes McNaughton et al. [250] investigated the effects of three different cadences of 50 RPM (low), 90 RPM (intermediate) and 110 RPM (high) on values of CP and W′. The low cadence trial gained significantly longer TTE durations and significantly higher values of CP when compared to their higher cadence counterparts with no effect on W′ evident. The hypothesis of a reduced endurance performance in recreational athletes when using high cadences, even though recruiting a higher proportion of type II fibre was confirmed as the greater cardio-vascular response at higher cadences appeared to be the dominant limiting factor when determining TTE trials. Consequently the authors advised the use of lower cadences when working with this subject group.

Barker et al. [251] investigated the differences in CP and W′ in addition to $\dot{V}O_2$ response when cycling at CP intensities using 80 RPM and 100 RPM. This study used a group of trained runners and trained sprinters. CP was significantly lower (189 ± 50 W vs 207 ± 53 W) employing the higher cadence strategy with no significant difference evident for $\dot{V}O_2$response between the two cadences or groups. However CP was significantly higher in the endurance group compared to the sprinter group using both cadences. Surprisingly W′ did not result in a significant difference between groups with only a trend of higher values in the sprinter group being evident.

Investigating how end cadence impacts on CP and W′, Green et al. [252] recorded TTE durations with a cut-off point of 50, 60 and 70 RPM. The study reported unaffected CP values with a difference in W′ seen at 70 RPM. This end cut-off cadence produced significantly higher W′ values than the lower cut-off cadences.

Given the differences in resultant CP values, the choice of cadence and the choice of participants seem to significantly influence tolerable durations at given work rates. Together they are important factors when designing and comparing CP determination methods or CP results. It is therefore good practice to use lower cadence ranges in untrained individuals whilst giving trained individuals the choice of self-selecting their preferred cadence.

2.7.2 Time to exhaustion trial durations

The range of TTE durations requires careful consideration in methodological designs, as resulting CP and W′ can differ substantially if long or short durations are used. The duration of high intensity exercise to the exhaustion is inversely proportional to PO [51,222]. Poole et al. [253] explicitly avoided intensities which induced exhaustion in less than 1 min due to an impaired mechanical muscular force-generation at extremely high imposed work-rates. Similarly, due to substrate limitation and motivational issues, durations of more than 15-20 min are commonly avoided [65,79]. Additional concerns regarding appropriate TTE durations were expressed by Monod and Scherrer [51] who suggested that the work-duration relationship loses its linearity when employing very short trial durations.

Di Prampero [208] specified, that ranges of TTE durations must be such that $\dot{V}O_{2max}$ is elicited and that W′ is depleted during each trial. After the onset of exercise $\dot{V}O_{2max}$ is generally attained within 2-3 minutes, which lead Poole et al. [65] later to recommend 2 minutes as the minimum TTE trial duration. A number of authors followed this recommendations of 2 to 15 minutes [18,113,177,254] whilst others either used shorter durations [80,96,204,255,256], longer durations of i.e. 20 minutes [205,255] or up to 50 minutes [257,258]. As the slope of the power-duration relationship depends on the chosen range of TTE trials, consideration has to be given to the choice of exhaustive trials when comparing results. In his review Vandewalle et al. [224] stated that CP values which were derived from short supra-maximal efforts equates to ~ 79% MAP in trained participants and values derived from effort durations of 3.5–35 min equates to ~ 69% MAP. Addressing the issue of different relative intensities, Bishop and Jenkins [259] investigated the dependency of the CP function on the choice of TTE durations. Using five TTE trials with a minimum of 1 min and a maximum of 10 min durations, the researchers derived CP and W′ using three different combinations. Combination one included the three shortest, combination three the three longest durations with combination two comprised mixed durations. All CP and W′ results were significantly different to each other, with combination one producing the highest CP and the lowest W′ value and combination three producing reverse results. However, it has to be noted that the highest CP values also resulted in the highest SEE. In kayaking, Clingeleffer et al. [96] similarly found significant differences between CP values obtained from TTE duration ranging between 90 and 240 s and those obtained from durations ranging between 90 and 1200 s. Using four different maximal efforts of 90, 240, 600 and 1200 s CP derived from only two efforts, which incorporated the 1200 s or the 90 s effort tended to result in lowest or highest values respectively.

More recently Hill et al. [182] recommended to avoid very short duration (≤ 2 - 3 min) or very long duration (≥ 15 - 20 min) TTE trials in the CP determination process. This was to minimise aerobic inertia and to reach aerobic steady state as well as to avoid the effects of hydration, muscle glycogen depletion and reduced motivation.

In short, CP is higher when only using shorter duration efforts or lower when using longer duration efforts. Depending on the chosen TTE durations, this can possibly

distort the physiological meaning of CP and shift its relationship further away from the MLSS.

Following the recommendations made [65,182] the selection of a wider range of durations which span intervals between 3 and 15 min, with a minimum of 5 min difference between the longest and shortest effort [260], appears to be the most consistent approach when determining CP. In order to gain coherent CP results, TTE durations have to be similar when repeating or comparing tests.

2.7.3 Inter-trial recovery times

Studies in cycling ergometry have used between a 24 h [132,261] and a 15 min [262] inter-trial recovery duration for the determination of CP. Resulting CP values were therefore determined over either one day or several days. Only Bishop and Jenkins [263] and more recently Galbraith et al. [264] directly compared the conventional 24 h method with an alternative one. In cycling results suggested that a 3 h inter-trial recovery period is sufficient to determine CP and W′ in untrained subjects [263]. In running and when using trained subjects a recovery period as short as 30 min appears to be sufficient to accurately determine CV but not the ARD [264]. The question of a shortest possible recovery is of particular interest to this thesis as an overall shorter CP determination method could enhance the practical utility of CP in research and in real-world cycling. Study 3 discusses the issue of inter-trial recovery time in more detail and compares CP values determined from a 24 h, a 3 h and a 30 min recovery protocol.

2.7.4 Number of time to exhaustion trials

Similarly to the choice of time to exhaustion trial duration and choice of cadence, consideration also has to be given to the number of TTE trials. The original CP work was based on three TTE trials for resistance exercise [51] and for whole body exercise [52]. However some researchers have used as many as seven TTE trials [197] which reduces the attractiveness of the concept. Poole [253] recommended using at least four to five TTE trials to obtain the most accurate values for CP and W′. Basing their research on the linear power-duration relationship, other researchers argued for the athlete's fatigue to

be of major concern in the CP determination methodology, rather than an increased level of accuracy [97,255,265]. In other words, a case has to be made for researchers who are investigating more applied aspects of CP determination, such as Ginn [255]. However only a few researchers [97,130,260,266] utilised or validated a similar method since Ginn's original work of validating CP determination in kayaking using only two timed maximal efforts [255]. Using the linear CP models, it is clear that when employing two TTE trials a perfect linear relationship is the only possible outcome. A possible increased risk of reduced reliability is associated with this method as an error in either trial will make a potentially significant difference to CP outcome by artificially inflating or deflating the values. A 'bad' test will have less impact on outcomes if at least three trials are employed. However, according to Hill [79], when working with trained individuals who are accustomed to exhaustive exercise, as few as two TTE trials can be sufficient. Nonetheless, as a 'trade-off' between accuracy of CP values and feasibility, Hill [79] also suggested an optimal number of four to five trials. The question of a lowest number of trials is also of particular interest to this thesis as stated in the research aims. Study 5 discusses the issue of using 2 data points for the determination of CP further.

2.7.5 Fixed distance and fixed duration exhaustive trials

Basing exhaustion inducing trials within 2 min and 15 min, a number of researchers substituted a fixed intensity with a fixed distance testing method to investigate field applications of the CP model. For example Hiyane et al. [267] utilised fixed distances of 2, 4 and 6 km to model CV in cycling. Kranenburg and Smith [268] utilised fixed running distances of 907, 2267.5 and 407.5 m on an indoor 453.5 m running track to compare track determined CV with CV values determined from fixed intensity treadmill running. No significant differences were identified and the researchers noted on track CV being easier to administer in motivated athletes. Galbraith et al. [269] also in running chose fixed running distances of 1200, 2400 and 3600 m to develop a novel CV methodology using a 30 min recovery time between exhaustive efforts. These distances targeted exhaustive times of 3 min, 7 min and 12 min durations. In swimming Dekerle et al. [270] applied a similar method of fixed swimming distances gaining exhaustive times in the proposed durations between 2 min and 15 min. In summary, exhaustive trials using fixed work-

rates, fixed durations or fixed distances are suitable to measure either time, mean power or mean velocity/speed respectively.

The above discussion highlights the need for research which addresses the underpinning physiology of CP and W′ to employ a minimum of either three exhaustive trials (using either fixed work-rates, fixed durations or fixed distances) and for greater accuracy to incorporate SEE values. Research which addresses the wider application of the CP concept and/or the cumbersome nature of CP determination can legitimately employ a maximum of three exhaustive trials. However employing only two exhaustive trials might always incorporate an inevitably high risk and should only be performed by experienced and well-trained athletes.

2.7.6 Practice effects and reliability of time to exhaustion trials

The duration of each exhaustive trial is crucial and errors in measurement, lack of motivation or non-familiarity by the participant consequently influence the determination of CP results [65,178]. Work-rates during TTE laboratory trials are commonly fixed. Participants, when experiencing increasing levels of fatigue and discomfort during such trials are only left with the choice between continuing or stopping the test altogether. This potentially results in measurement errors and could lead to a different outcome during repeated trials. Reliability as one of the scientific criteria has therefore to be considered carefully when validating a new determination method.

Investigating the repeatability of TTE trials at imposed intensities, Poole et al. [65] demonstrated a significant learning effect in less experienced participants as a second trial generally resulted in longer trial durations. Moreover there was a difference in exercise tolerance increase between shorter and longer trials. Improvements in shorter TTE trials (~ 4 min) were smaller than those of longer trials (~ 8 min) (2-4% and ~ 4-6% respectively) and participants demonstrated a greater learning effect for lower intensity TTE trials. As reported by Poole et al. [65], Alberty et al. [271] in their swimming study also suggested a higher within-subject variability for lower imposed intensities and generally concluded on a reduced reliably for constant work-rate tests. This notion of altered

reliability with different exercise intensities and durations has been debated previously in the literature. [178,272]. Hopkins et al. [178] for example, demonstrated CoV values between 0.9 and 2.0% for TTE trials performed in the severe-intensity exercise domain in trained athletes which was shown to increase by additional 1.3% in non-athletes.

Gaesser and Wilson [72], for two repeated tests reported test re-test coefficients (r^2) for CP 0.92 and for W' of 0.62, indicating a higher variability of the W' parameter. When investigating repeated CP measurement using 5 TTE trials Smith and Hill [273] reported a high correlation between test re-test results (r = 0.92 for males and 0.9 for females) and a mean 5.5% difference in CP values with no difference for values of W'. Test re-test correlations for W' of r = 0.8 and 0.64 for male and females respectively caused Smith and Hill [273] to support the notion of CP to be less variable than W'.

Significant test re-test correlations were also found by Nebelsick and Housh [73] who in contrast to Smith and Hill [273] did not identify a significant difference for either CP or W' values. The study reported higher test re-test correlations for CP (r = 0.94) than for W' (r = 0.87). In cycling Jeukendrup et al. [7] compared the reproducibility of commonly used types of laboratory performance test using well-trained athletes. Test included constant work-rate tests to exhaustion, maximal work tests with an imposed intensity or a fixed duration and TT tests. The researchers stated a poor level of reliability for constant work-rate tests. Even though implementing one familiarisation trial the study reported a CoV value as high as 26.6%, whilst both other tests resulted in CoV below 3.5%. This is further supported by McLellan et al. [176] who performed 5 repeated TTE trials using 15 males of average fitness levels. The research reported a substantial variability for the repeated TTE trails with CoV values ranging from 2.8 to 31.4%. Using only highly-trained cyclists, Laursen et al. [274] also reported significantly longer second constant work-rate tests (245 ± 57 s) compared to first ones (237 ± 57 s) performed at $\dot{V}O_{2\,max}$ intensity with a reported relatively low CoV of 6%. In running Laursen et al. [275] furthermore directly compared TTE run with TT runs. Using eight endurance trained participants, the study whilst not finding a significant difference between a first and a second TTE and TT run, also reported greater levels of variability for the TTE efforts.

Contrary Hinckson and Hopkins [256], when investigating repeatability in time to exhaustion runs for CV determination found a test re-test error of less than 3%, which was deemed as representing excellent reliability. Hopkinson et al. [178] also suggested that TTE trials might require less familiarisation , as no self-selection of pace is needed. However, the literature generally agrees on a lower reliability for TTE trials with only Hopkins et al.[178] arguing the poor reliability to be an artefact between the relationship of exercise duration and PO. Moreover, Hopkins et al. [178] stated, that TTE trials appear to be more sensitive to changes in performance capabilities. A further discussion on reliability can be seen under heading 9.2.1.

2.7.7. Practice effects and reliability of time trials

A higher level of reliability appears to hold true for TTs, where the athletes are able to change the intensity according to their perception of fatigue and external motivational cues [276]. TTs have been deemed as more reliable in the literature [7,178] whilst potentially adding some variability to the measurement, as intensity fluctuates [256]. This however was argued by Jeukendrup and Currell [277] who identified pacing strategy as an inherent component of real performance which should not be excluded in performance tests. An acceptable level of variability therefore deems a test as being reliable [278]. Due to encompassing a higher level of variability, TTs whilst providing less sensitivity to changes in performance capabilities [279] offer a higher level of reliability.

To minimise random measurement errors, the above discussion emphasises the need to recruit trained participants for sport performance studies. External validation requires that the training should match the level and specifics of the performance being tested. If employing unaccustomed testing procedures, the best practice is to provide participants with the opportunity to perform a familiarisation trial. Some conjecture however surrounds the choice of exercise test and absolute recommendations cannot be made by current review findings.

2.8 Laboratory and field performance testing

The highest form of performance testing is the competitive performance itself, since it is at this juncture that all the elements involved in performance actually interplay [71]. In order to successfully bridge the gap between sports science studies/experiments and the real-world cycling, well-considered laboratory and field testing protocols are required [280]. Relevant laboratory tests commonly use standard bicycle ergometers. These ergometers simulate the sport with results generally being more reliable but less ecologically valid than field testing protocols [6]. Stationary cycle ergometers do not have the same mechanical properties, such as stiffness and damping as road cycles and the kinetic energy and crank inertial load in ergometer cycling is different to road cycling [281]. The kinetic energy in road cycling varies according to the cycling velocity and the mass of the cyclist [24]. Moreover ergometer cycling commonly uses a prescribed or freely-chosen pedal cadence which remains constant throughout testing and also commonly does not provide the cyclist with a gear changing option. In contrast, field tests can be considered as more specific as they more closely replicate what the athlete is challenged by in the natural environment of training and competition [6]. Field testing conducted away from the confines of the laboratory can however introduce 'unwanted' and 'uncontrollable' variables (i.e., 'noise'). These might influence the research design and outcomes, even if increasing the ecological validity of the study [6], Ecological validity, defined as the relationship between real-world phenomena and the outcomes of the investigation of those phenomena in a laboratory and/or experimental context [282] in field cycling contains unwanted and uncontrollable variables such as wind, weather conditions (humidity and temperature), road/track surfaces and terrains of hills or mountains. As a consequence, a certain degree of trade-off between experimental control and ecological validity is unavoidable when testing athletes. Generally, the greater the experimental control, the lower the degree of ecological validity, the less the results are likely to reflect real-world performance.

Historically, field testing procedures and data collection were less sophisticated prior to the development of mobile power meters, that is, where training and performance intensities in the field were mostly described through HR [283–285]. HR however is acutely influenced by several physiological factors, such as hypo-hydration and hyperthermia [286]. Using HR as workload feedback might result in an over- or underestimation of true

physiological demands [92]. Voigt et al. [92] for example found that compared to PO, HR underestimated the time spent below the LT intensity and the time spent above the LT plus 1 mM intensity, whilst overestimating the time spent between LT and LT plus 1 mM intensity.

Following the development of mobile power meters researchers are now able to gain the same levels of accuracy in the measurement of PO as those obtained from a stationary laboratory ergometer [44,92]. However, limited research has addressed differences between laboratory and road cycling, with inconsistent results. For example Bertucci and Taiar [10] investigated the differences in sprint performance between laboratory and field cycling. Cyclists had to perform six sprints (three seated and three standing) on a laboratory ergometer, and six of the same tests in the field using a road bicycle equipped with a mobile power meter. The ergometer recorded significantly lower maximal pedal forces (seated and standing) and significantly higher PO values in the seated position when compared to the field. Conversely, standing field sprints produced significantly higher values than the standing ergometer sprints. Bertucci and Taiar [10] explained the latter difference by zero lateral movement of the ergometer, which does not replicate real-world cycling well enough to obtain valid estimations of the maximal PO. The researchers further highlighted the necessity to perform sprint test investigations during actual cycling locomotion in order to obtain a high level of ecological validity. Gardner et al. [287] unlike Bertucci and Taiar [10] did not identify any significant differences for maximal PO, cadence or maximal torque between 6 sec ergometer and 65 m track sprint cycling. The researchers consequently suggested that maximal laboratory cycling does provide accurate means of measuring cycling performance. Bertucci et al. [24] later compared crank torque profiles, PO and rate of perceived exertion (RPE) during laboratory cycling with level ground and uphill cycling in the field. Using a Monark cycle ergometer equipped with a mobile power meter, cyclists had to perform 1 min maximal efforts in the laboratory at 60 RPM, 80 RPM and 100 RPM. Results were compared to level terrain (80 RPM and 100 RPM) and uphill (60 RPM and 80 RPM) cycling and indicated significantly different crank torque profiles between the ergometer and the field. The disparities were explained by the differences in crank inertial load. Furthermore the ergometer generated higher RPE levels, which were linked to the differences in crank torque profiles but also to the differences in exercise environment. Jobson et al. [288] found a 4% difference between laboratory and field performances of a

40-km TT. Cyclists had to complete one TT on a Kingcycle ergometer and one TT in the field. Faster times in the laboratory were explained by differences in body size whereby the larger cyclists potentially benefited from the controlled conditions. In contrast, when performing in the field, it was suggested that riders with a greater body surface area would be more likely to experience an increase in drag compared to smaller riders. Utilising a Kingcycle ergometer and a road bicycle equipped with a mobile power meter, Smith et al. [289] also demonstrated faster performance times of a 40-km TT performed in the laboratory compared to the field. Interestingly mean PO values between laboratory (303 ± 35 W) and field (312 ± 23 W) were not significantly different. Therefore it can be suggested here that a more stable laboratory environment will produce less fluctuating PO and more consistent cadence values which might result in faster performance times.

On the contrary Peveler [9] using global positioning system technology found faster field TT results. Under laboratory conditions cyclists were using their own bicycles attached to a Computrainer. The study identified a significant difference between times yielded from the laboratory (~ 35 min) and the field (~ 26 min) and the author stated that meaningful comparisons of performance cannot be made using these two environments. Different performance times might be explained by possible higher cycling velocities caused by advantageous tail wind conditions, advantages of the riders' body surface area and/or advantageous terrain conditions. Padilla et al. [290] investigated physiological responses determined from track and laboratory cycling. MAP, HR and $\dot{V}O_2$ were similar between the two environments. However blood [La] on the track was significantly higher. For a more accurate performance level prediction in the field, Padilla et al. [290] consequently recommended metabolic cost to be more appropriately expressed per unit of body surface or body mass. Jobson et al. [291] in a later study demonstrated higher field PO values when performed in an aerodynamic cycling position but there was no difference in PO values between field and the seated upright laboratory TTs. The higher aerodynamic field PO values were explained as the result of a possibly increased evaporative heat loss accompanied by a lowered body temperature and lowered RPE values. The study also found faster laboratory speeds for both cycling positions which were explained by a greater road cadence variability (9.2% vs. 1.5%) caused by differences in course topography. Most importantly, the study demonstrated

the independence of PO from environmental conditions, as the higher mean road PO values were not reflected in higher velocities, when comparing the same cycling position. Jobson et al. [291] however stated that body position actually does not affect the ecological validity of laboratory TT cycling, as the possible field aerodynamic advantages did not cause a significantly different performance outcome. Supporting these findings, no significant difference between the two environments were found by Gardner [287] who demonstrated consistent PO values between laboratory and field performances. The study compared 6 s maximal sprints performed on an SRM ergometer with 65 m standing start field sprints on road bicycles, equipped with a mobile SRM power meter. Finally Bertucci et al.[292] demonstrated 10% higher GE and cycling economy (CE) values in the field under level and uphill cycling conditions when compared to those collected in the laboratory. Results in the laboratory were determined through a software controlled simulation of level ground and uphill cycling. Cyclists in the laboratory only under uphill conditions demonstrated a preference for a higher cadence. Bertucci et al.[292] suggested that the differences might partially be due to the fact that cyclists were not able to perform their habitual side to side cycling motion when riding on a stationary ergometer. These researchers emphasised the distinct advantage of cyclists using their own bicycles as compared to the usage of a 'conventional' ergometer, such as the SRM, Lode or Monark by stressing the importance of standardisation (i.e. using the same bicycle) when comparing laboratory with field findings. The combination of differences in i.e. body position, posture, muscular activity or familiarity of the bicycle can potentially influence physiological responses. In contrast Arkesteijn et al. [293], when investigating stationary with treadmill cycling demonstrated that the type of cycle ergometer can be altered without affecting efficiency. In support of a similarity between environments Nimmerichter et al.[294] recently reported a strong relationship between maximal and sub-maximal physiological measures and acceptable levels of agreement (LoA) between a 4 minute TT and MAP (random error of -7.4 ± 14%) and between a 20 minute TT and the lactate turn point (LTP; random error of 0.02 ± 13%) obtained in the laboratory on a Lode Excalibur ergometer with those obtained in the field.

The above discussion summarizes an apparent disagreement about whether laboratory cycling generally replicates real-world cycling, but also opens a further discussion about the use of specific ergometers [292,295]. The apparent advantages of using PO as the

dependent research variable are that PO is less influenced by internal and external factors and that PO represents the most precise description of cycling performance [38,42]. However, to date only a few studies have been published using PO in a field based research setting [92,198,296,297]. The above reviewed studies are furthermore mainly concerned with the relationship between cycling speed or mechanical PO. Relatively few researchers have attempted to validate a field test against specific reference laboratory tests. Padilla et al. [290] for example validated a maximal velodrome test for competitive cyclists. No significant differences were identified between velodrome and laboratory MAP and $\dot{V}O_{2max}$ values, while maximal blood [lactate] was significantly higher in the velodrome. The study however did not report LoA or prediction error values. Another validation study was performed by González-Haro et al. [12] who compared MAP values established in the laboratory with those established from an incremental velodrome test. The respective field test nonetheless requires knowledge of the athlete's 50% MAP value (i.e. another test is required to establish this value) and is based on a load increase of 12.5 $W \cdot min^{-1}$.

According to Nimmerichter et al. [294] it is still unknown whether an uphill or flat TT of the same duration would result in different PO values. The researchers attempted to standardise the testing conditions between laboratory and field by prescribing an average gradient of ≤ 5% with a maximal difference in altitude of 10 m (TT4) and 50 m (TT20). Nimmerichter et al. [294] reported a strong correlation between TTs and performance measures established from a maximal graded exercise test (GXT) performed in the laboratory. However, significant differences were observed between PO produced during TT4 and the GXT.

Indeed there is a lack of studies that more boldly tackle the real world of road cycling. Such studies would allow road cyclists to perform not just outdoors but also in their natural environment, i.e. in hilly, mountainous or flat terrains. This in turn would provide coaches with more ecologically valid data whilst becoming more independent of a sports science laboratory and the expertise of one or more sports scientists.

2.8.1 Measurement error and reliability in performance tests

Reliability is of high importance in sports science research and of particular interest in the presented research findings. Reliability indicates the precision and therefore validity of a test to track athletic performance changes and/or to detect the effects of any intervention on performance. Hopkins et al. [178] defined reliability as the reproducibility of performance outcomes, when repeatedly performing a test. For Atkinson and Nevill [278], reliability is considered as the amount of acceptable measurement error that demonstrates the effectiveness of a test as a measurement tool.

In competitive cycling, Paton and Hopkins [298] observed mean PO changes over a 4-km performance of 6.1% and of 2.2% from base to pre-competitive and from pre-competitive to competitive season respectively. In elite cyclists, performance changes as small as 1.7 % enhance the chances of winning an event [299]. Relevant performance tests should be able to detect changes of that magnitude, i.e. the measurement error requires to be of a smaller magnitude than the change in performance increase/decrease. Performance test reliability can be expressed in a number of ways, such as CoV, which is the standard deviation of the measure divided by the mean of the measure. Lamberts et al. [300] for example demonstrated a low variability for laboratory 40 km TT performances (CoV of 0.7 % for time and 1.7% for mean PO) in well-trained cyclists. Similarly, during severe exercise Hopkins et al. [178] demonstrated a test-retest CoV value of ~ 0.9–2.0% in trained athletes. Upper levels of acceptable measurement error, i.e. CoV values of 5% [301] or 10% [302] have been proposed. Other reliability measures include the Pearson's product moment correlation (r) which can identify significant, i.e. reliable correlations. However the value cannot detect changes in the mean and it is sensitive to the heterogeneity of values between participants [303]. Intra-class correlation coefficient (ICC) values are also commonly used in reliability studies. An ICC value > 0.9 indicates high reliability, a value > 0.8 moderate reliability with a values of 0.7 or less questioning the reliability of a testing protocol [304]. The ICC however can be sensitive to systematic bias and is affected by the sample heterogeneity [305] but it is sensitive to the order and the magnitude (i.e. mean difference) of repeated values and therefore meaningful statements can be made about the reliability of a measure [304]. LoA have also been used in research to express the reliability of a testing protocol. An advantage of using CoV as the expression of reliability is that of the widely accepted upper limits as proposed by

Nevill and Atkinson [302] and by Hopkins et al. [178]. Based on these upper limits researchers can make clear statements about the reliability of a testing method and put them into the context of a meaningful change in performance.

For the purpose of the presented research, the decision was taken to accept measurement errors below 5% as indication of reliability [178]. The presented research consistently uses CoV values, if repeated trials were performed. Study IV of this thesis additionally expresses reliability using ICC values with an indication of the level of reliability as described above.

Different types of performance tests are associated with different levels of typical error of measurement (TEM), where the TEM is divided into systematic and random error. In cycling, the systematic error is associated with inaccuracies related to the measurement device to for example measure PO accurately [8]. Measurement devices therefore can be reliable whilst being invalid if providing inaccurate values. The systematic error can be reduced through the choice of an appropriate measuring device, examples of which the presented studies are the SRM ergometer and the PowerTap power meter. Both have shown to measure PO accurately (i.e. valid) and reliably [44,46]. The random error includes test biological re-test variability caused by the athletes [8]. Lamberts et al. [22] demonstrated a lower level of random error when using experienced or highly trained cyclists. Keeping biological variability low, the experimental studies therefore attempted to recruit experienced cyclists. The random error furthermore contains ergometer variations, which for example can be the cause of a calibration drift. The level of random error can be reduced by the choice of appropriate tests. Generally fixed duration or distance TT appear to demonstrate greater reliability than TTE tests [305]. Commenting on a lower levels of TTE reliability, Laursen et al. [275] pointed towards possible contributing factors such as boredom or lack of motivation associated with open-end exercise tests, when participants are blinded towards an 'expected duration'. It is however doubtful that CP relevant TTE trials which are located within the severe domain would evoke boredom. Hinckson and Hopkins [256] furthermore argued that the poor reliability of TTE trials might be an artefact between the relationship of exercise duration and PO rather than variability caused by an athlete per se. A small random change in a subject's ability to output power from test to test by e.g. 1% can result in a

much larger random change in TTE duration (~ 10-20%). Therefore a small change in performance capability can be detected by this type of test.

2.9 The need for refined/novel methods to determine critical power

The key characteristics of most commonly used performance markers in sports and their advantages/disadvantages as a testing tool are summarized in table 4. Whilst satisfying a number of scientific requirements, such as being objective, valid, sensitive and reliable, CP determination to date is not athlete friendly or available as a field protocol.

Up to now the conventional determination of CP requires an incremental maximal exercise test in addition to a minimum of three TTE trials, each performed on different days. Moreover, CP determination to date is performed in the laboratory, making CP less accessible to athletes but also imposing a demanding and time consuming protocol onto the athlete. Given these factors the practical utility of CP is low [306], and on this basis CP is not routinely assessed in research or clinical exercise testing [19]. However CP determination does provide a more meaningful index of aerobic fitness, when compared to other indices, such as the LT or $\dot{V}O_{2max}$ [55].

Table 4. Commonly used methods for evaluating endurance fitness

Test	Strong Evidence 'Validity'	Strong Evidence 'Reliability'	Strong Evidence 'Sensitivity'	Objective	Field Testing	Athlete Friendly
Lactate Threshold	√	√	√	x	√	√
Ventilatory Threshold	√	√	√	x	x	√
Maximum Lactate Steady State	√	√	√	√	√	x
Critical Power	√	√	√	√	**x**	**x**
Onset of Blood Lactate Accumulation	x	x	√	√	√	√
Individual Anaerobic Threshold	x	?	√	√	x	√
Lactate Minimum Speed	x	√	x	√	√	√
Velocity at $\dot{V}O_{2max}$	√	√	√	√	x	√
Externally Valid Time Trial	√	√	√	√	√	√

Reproduced and adapted from Jones, AM., and Doust, J. (2003) Limitations to submaximal exercise performance[152]. In Reilly and Eston (eds) *Kinanthropometry and Exercise Physiology Laboratory Manual: pp 235-262*

None of these commonly used endurance fitness markers to-date fulfils all criteria of scientific robustness whilst providing coaches with an athlete-friendly, non-invasive field testing application. Moreover it is only CP that fully applies to the mode of cycling without the reliance of sophisticated analytical equipment, such as a gas analyser. It is its significance (see 2.2) and its role as a physiological intensity marker (see 2.3) rather than a physiological event that makes CP potentially highly attractive for coaches and athletes. Thus, if CP was to be validated as an athlete-friendly field protocol, it could offer an easy and accessible tool for coaches to systematically monitor the endurance capacity of their athletes.

An additional important area to be considered is that despite the potential importance of the CP concept to sports performance, the practical features have rarely been communicated between science and coaching [55]. A number of reasons are likely to be responsible for this lack of practical application. For coaches to adapt the concept, an easier 'grasp' is required. However the concept is commonly engulfed in physiological specific jargon and mathematical terms hindering its progression into the real word of sport [55]. Furthermore an apparent agreement in the scientific literature over which mathematical model to be the most suitable one is missing. Whilst a number of models have been extensively discussed in the literature (see 2.6), a decision over a 'best practice' model is outstanding. Therefore it is left to the coach to take that decision or to reject the adaptation of the concept based on lack of clarity. It appears however, that a number of researchers, possibly due to a more simple use of mathematical model have chosen the power-1/time model as their preferred analytical model [177,185,307] whilst others are using more than one model for their result analysis [88,205].

As recently highlighted by Argyris and Savvas [308], the use of an appropriate trial number in combination with the right mathematical model only leads to the CP calculation which complies with its definition of a prolonged sustainable exercise intensity. The overall inconsistencies in the literature potentially contribute to the confusion for coaches on how to apply the CP concept into a real-world setting.

Furthermore standardisation presents a challenge. Like the appropriateness of the applied CP model, the number of trials is not standardised, potentially adding to a decrease in interest in the practicality of the concept. Unlike other performance tests, the number of trials and the durations of those will lead to different CP results [79].

The robust nature of the power-duration relationship has widely been accepted in research yet has failed to make a substantial impact in the real world [19]. Whilst its ability to characterize exercise tolerance over durations of ~ 2-30 min has real relevance to sporting performances [80], an application of the concept in wider sporting context is yet to be made. Vanhatalo et al. [55] stated that the concept cannot predict exercise tolerance at or below CP intensity and that it is only applicable for sports generally in which a significant period of performance time is spent in the severe-intensity domain. The latter

applies to such a variety of sporting events for which to date CP has failed to impact on. By definition, performing in the severe domain results in expenditure of W'. W', possibly due our incomplete understanding of its exact nature [223] and due to inherited difficulties in accurate and reliable measurement [197] has not yet made any impact in the real world either, whilst interestingly attracting ever increasing research attention. In cycling it is also less meaningful as expressed in kilo Joules or Joules and not in Watt, making it less user-friendly for coaches and athletes to interpret. Little is known about the exact rate expenditure, only that the intensity above CP performances dictates the rate at which W' is being expended, i.e. the higher the wattage above CP the faster the expenditure. There is also a lack of clearly defined context in which W' could be integrated into training and performance. A major attempt was made by Skiba et al. [89] who recognised the theoretical importance of W' for pacing strategies and dynamically modelled W' utilisation. However, knowledge of how W' can be used in training to-date is limited.

In cycling, riders with higher W' values generally have greater sprint capabilities and it is widely known that there is a 'conflict' between W' and CP in that well- trained athletes cannot have high values in both [55]. In particular highly endurance trained athletes have CP values which are close to their $\dot{V}O_2$max value. However they also have only modest W' values [55]. With the exception of sprint coaches, W' therefore only has a limited value to coaches who are mainly interested in increasing the endurance performance capabilities of their athletes. However since CP and W' originate from human bioenergetics and are both part of the power-duration concept, more research is required to fully understand their multiple and interrelated factors and how to apply the resultant knowledge to performance improvements. This however was not the main focus of this research thesis. Whilst reporting values of W' for all studies, it was CP with its significant advantages (see 2.2) that was used as a tool for field performance testing.

Besides any practical issues (i.e. being 'user-unfriendly'), the combination of the choice of model, choice of number of trials and choice of trial durations and potentially some inconsistencies over the meaning of W' might be incomprehensible to coaches. As suggested by Vanhatalo et al. [55] refined determination methods are required. Before

reaching its full real-world potential, a focused approach on standardised determination methods however is also needed.

2.9.1 Statement of the research challenge

Reliably monitoring training adaptations requires regular testing to ensure that the training performed is achieving targeted adaptations [7,8,22,305]. Therefore, to be of use in the real world, any approach to the measurement of PO in cycling must be sufficiently sensitive to detect very small changes in PO that occur in the well-trained athlete reliably [21]. It is not unreasonable to argue that currently available field tests neither fully satisfy scientific requirements nor are they particularly athlete-friendly [6,7]. Furthermore, practical issues such as limited access to a testing laboratory or financial resources might prevent regular testing, and repeated laboratory tests can compromise the training, recovery and performance of athletes. A solution to this problem would be the adoption of measurement technologies, such as mobile power meters and the integration of performance testing into routine training and/or competitive activity. Ecologically valid field tests could provide coaches with more relevant data, whilst technological developments such as mobile power meters, independently of training regimes, such as interval training provide immediate real time intensity feedback. With limited research available relating to the translation of laboratory tests in the field to measure field performance indices [43,309–311], future challenges are apparent for valid and reliable testing applications.

The apparent research questions raised in this thesis attempt to address the above need for a cycling field performance test. Requirements of such test are that it has to be standardised, easy to use, athlete-friendly and that it determines a valid performance marker, such as CP. Moreover this performance test should have the potential to achieve some real-world impact.

3.2 Research questions:

Study 1 addressed the validity and reliability of the 3-min all-out cycling test by Vanhatalo et al. [312] . The null hypothesis was of a non-significant difference between CP values determined from the 3-min all-out test and conventionally determined CP values when using a different exercise mode and ergometer.

Study 2 addressed the viability of CP field determination. Field testing was located on an outdoor tarmac cycling track. The null hypothesis was of a non-significant difference between CP values determined in the laboratory and in the field.

Study 3 addressed the elaborate laboratory CP determinaton method by comparing conventionally determined CP values using a 24 h recovery period between TTE trials with those, using a 3 h and a 30 min inter-trial recovery period. The null hypothesis was of a non-significant difference between all CP values.

Study 4 addressed the validity and reliability of CP determined from exhaustive trials performed on the road. Based on the previous findings, the study comprised laboratory CP determination using a 30 min recovery method and three novel CP protocols. Two of these comprised intentional pre-defined efforts with a third one extracting non-intentional and intentional highest efforts from training and racing data. The null hypothesis was of a non-significant difference between CP laboratory values and those determined in the field, i.e. on the road.

Study 5 utilised collected CP values from study two and study four and investigated the agreement between CP determined from 3 data points with those determined from 2 data points. Laboratory values were compared with laboratory values, field with field values and laboratory with field values. The null hypothesis was of a non-significant difference between 3 data points and 2 data points determined CP values.

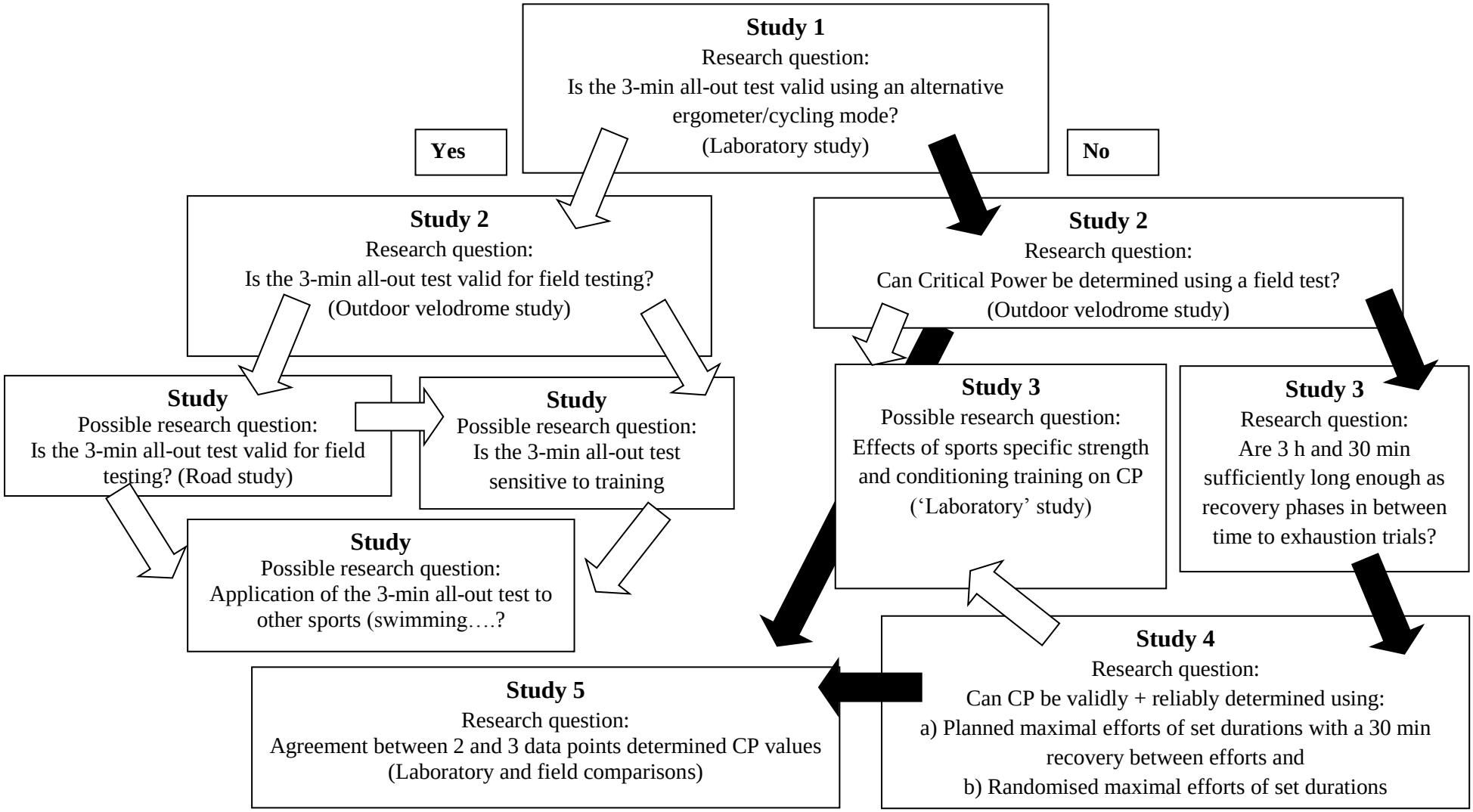

Figure 8. Overview of research process. Black arrows indicate the actual research route. The white arrows indicate hypothesised/projected alternative routes.

CHAPTER 3: GENERAL METHODS

3.1 Ethics, health and safety

Prior to each study, full ethical approval was obtained from the University of Greenwich Ethics Committee (UREC). All UREC applications considered health and safety aspects relating to participants as well as aspects associated with the use of the laboratory, laboratory testing, analysis equipment, use of testing equipment in the field and general field testing. All procedures were performed in accordance with the Declaration of Helsinki (1964).

To be safely used by human participants the research laboratory was fully prepared prior to and post each laboratory based test. These preparations were in accordance with the safety guidelines established by the Department for Life and Sports Science. Preparations included equipment cleaning (i.e. ergometer), equipment calibrations (i.e. gas analyser) and the assembly of required testing equipment (sufficient amount of wipes, blood sampling containers, etc.). Potentially hazardous equipment such as facemasks and turbines were cleaned by immersion in a container filled with disinfectant for a minimum of 30 min post-tests in accordance with the manufacturer' guidelines. The Sports Science laboratory uses a binary disinfection system to prevent equipment building immunity which could increase the risk of contamination. Researchers alternate between a sodium hypochlorite tablet (Milton, Procter & Gamble, Weybridge, UK) and a potassium peroxymonosulfate powder (Virkon, DuPont, Bristol, UK). Before reuse, this equipment was left to air dry.

3.1.1 Blood Sampling

During blood sampling and blood analysis disposable latex gloves were worn by the PhD candidate at all times and continuously changed to avoid any possible cross contamination. Prior to taking fingertip blood samples, the sampling site (non-dominant hand) was cleaned with alcohol. Disposable Unistik lancets (Mumford, Oxford, UK) were used to penetrate the skin. Initial drops of blood were wiped off and samples of ~ 20 µL of arterial blood collected in capillaries and emerged in cups

filled in with haemolysing solution. Blood samples were analysed using the Biosen C_line analyser (EFK Diagnostics, Barleben, Germany). During studies 1, 3 and 4 post analysis cups were immediately disposed of in sharps bin containers in accordance with the Human Tissue Act (2004). Other hazardous material such as lancets and blood sampling cubs were also disposed of in sharps bins. Wipes and tissues and latex gloves were appropriately disposed of into clinical waste sacks for incineration.

The second study was performed on an outdoor track (560 m circumference). Testing included blood sampling with was performed in accordance with the safety guidelines established by the Department for Life and Sports Science for field testing. Blood samples were stored in haemolysing solution cups, which in accordance with the manufacturers' guidelines, when stored in a fridge are stable for up to 5 days. Samples were analysed three times per week. Hazardous material such as lancets, wipes and latex gloves were appropriately stored and on blood analysis days disposed for incineration.

3.2 Specific methods

3.2.1 Subject recruitment and test preparations

All testing was performed using trained, recreational cyclists with a minimum of two years training and competition experience. Generally trained cyclists are accustomed to sports specific maximal and high intensity efforts through training and racing. Even though trained, none of the participants were elite-level athletes. Prior to participation cyclists were given verbal and written information (*Appendix* 1) containing detailed descriptions of all testing procedures. These were reiterated prior to the initiation of individual tests and any remaining questions were fully answered before test commencement. Associated risks and benefits were clearly stated. Following a standardised medical questionnaire, suitable participants provided written informed consent to participate in the study (*Appendix 2*).

For each testing procedure, participants were requested to be:

i) rested (no strenuous exercise in the preceding 24 hours),
ii) well hydrated,
iii) to refrain from consuming alcohol for 24 hours,
iv) to refrain from consuming food or caffeine in the 3 hours before each test.

3.2.2 Feedback and test familiarisation

To maintain motivation and to ensure maximal efforts, participants were given strong verbal encouragement throughout all laboratory and track testing procedures. Feedback regarding remaining test times was also proved for track TT tests. Feedback on performance and individual study outcomes were provided after study completion.

Some of the testing protocols contained a degree of unfamiliarity for participants, even though being accustomed to exhaustive training. In particular the utilisation of a novel 3-min all-out exhaustive test in the first study required participants to perform one familiarisation trial. Further, participants were instructed to perform one unsupervised familiarisation trial of each effort (Study 2) and of protocol 1 and protocol 2 (Study 4), which were not included in the data. The familiarisation trials in Study 2 and Study 4 reduced the impact of learning effects associated with TTs [274] whilst ensuring that participants were comfortable with the exact nature of each maximal effort test.

3.2.3 Measurement procedures

3.2.3.1 Anthropometrics

Anthropometric measurements were taken prior to each study. These included age, body height and body mass. Body height was measured using a portable stadiometer (Seca GmBH, Hamburg, Germany), measured to the nearest 0.1 cm. Body mass was measured using a laboratory digital scale (Seca 861, Seca GmbH, Hamburg,

Germany) to an accuracy of 0.1 kg. During both measurements participants wore their cycling clothes but no shoes.

3.2.3.2 Cycle ergometer and power meters

All laboratory testing was performed in a seated position. Study 1 was performed on a SRM ergometer (Schober Rad Messtechnik, Jüelich, Germany), Studies 2 and 3 were performed on a road bicycle (Raleigh Airlight 100, Bishops Stortford, UK) equipped with a PowerTap Elite wheel (CycleOps, Madison USA). Study 4 was performed on participant's personal road bicycles, which were equipped with a PowerTap Elite wheel.

Prior to each test, the SRM ergometer and PowerTap wheels were zero-offset according to the manufacturers' instructions. The saddle and handlebar of the SRM ergometer and the Raleigh road bicycle were adjusted to suit each participant and settings were replicated exactly during each subsequent test. Participants in all studies were permitted to use their own pedals.

3.2.3.3 Pulmonary gas exchange

The first two studies measured pulmonary gas exchange breath-by-breath using a MetaMax gas analyser (Metamax 3B, Cortex Biophysik, Leibzig, Germany) and the final two studies used a Cortex MetaLyzer 3B gas analyser (Cortex Biophysik, Leipzig, Germany). Participants wore a face mask (Hans Rudolph, Shawnee, USA) which was attached to a mouthpiece and triple V volume transducer turbine.

The inspired and expired gas volume and concentration signals were continuously sampled using electro-chemical (O_2) and infrared (CO_2) analysers via a capillary line connected to the mouthpiece. Prior to each test the analyser was calibrated against gases of know concentration (16% O_2, 5% CO_2; Viasys, Hoechberg, Germany), and the turbine volume transducer was calibrated using a 3 L syringe (Hans Rudolph, KS). $\dot{V}O_2$, carbon dioxide output ($\dot{V}CO_2$) and minute ventilation ($\dot{V}_E$) were calculated using standard formulae [147].

3.2.3.4 Maximal incremental test

All studies required participants to perform an incremental maximal exhaustive test performed in the laboratory. The incremental protocol consisted of a 5 min warm-up period at a set intensity (males: 150 W; females: 120 W) which was followed by an increase of 20 $W \cdot min^{-1}$ until volitional exhaustion. $\dot{V}O_2$ was continuously measured breath-by-breath and participants typically reached exhaustion between 12-15 min. $\dot{V}O_{2max}$ was calculated as the highest mean oxygen consumption over a 30-s period. MAP was determined as the highest mean PO during this same period.

3.2.3.5 Blood lactate and heart rate recording

Blood samples were sampled prior and post each laboratory TTE trial and prior and post each fixed-duration TT in Study 2. With the exception of the field part of Studies 2 and 4, HR during laboratory testing was continuously monitored using the Cortex gas analyser and recorded second by second. HR during the field testing part in Study 2 and 4 was measured using short-rate telemetry and recorded second by second through the Garmin head unit (Garmin, Olathe, Kansas, USA) which was attached to the handle bar of the road bicycle. Tests were subsequently downloaded for analysis of HR response using the software of the Garmin Training Centre (Garmin, Olathe, Kansas, USA). A number of training files in Study 4 revealed that participants did not consistently adhere to wearing the HR monitor during maximal efforts. HR consequently was excluded from the field data analysis in Study 4.

3.2.4 Standardisation of field tests

3.2.4.1 Track testing

Study 2 was performed in an outdoor track with a 640 m circumference. Participants had to perform on different days a 3 min, 7 min and 12 min maximal effort. A 5 min warm-up period was performed prior to each test by participants cycling around the track at a self-selected pace. Testing commenced with a standing start, allowing maximal acceleration. After sitting down participants were required to remain in this position but were allowed to change gear. Verbal feedback and encouragement was

provided after each completed lap. Participants were required to provide a maximal effort and to avoid pacing during each test. Once completed participants were required to continue cycling after a test to allow for a cool-down phase.

3.2.4.2 Road testing

Study 4 has a lesser degree of standardisation in that participants performed all testing unsupervised on regular roads. Cyclists were required to perform all maximal efforts in a rested and warmed-up stage. Individual exhaustive test durations were identical with Study 2 but also included sets of all maximal efforts using a 30 min recovery between efforts and non-planned maximal efforts. The recovery periods were either performed passively resting or participants continued to cycle at a recovery intensity. If rested passively cyclists were required to continue cycling after 25 min to allow for a 5 min warm-up period. Furthermore no instructions were given about a seated or standing position but participants were requested to avoid freewheeling.

3.2.5 Standardisation of laboratory tests

In all studies CP was determined from three constant work-rate tests at power settings equivalent to 80%, 100% and 105% MAP. After a 5-min warm-up at a work-rate of 150 W, the test resistance was set and cyclists were instructed to maintain their self-selected preferred cadence for as long as possible. Consistent with previous CP research [312] strong verbal encouragement was provided throughout the tests. Tests were terminated when cadence dropped by 10 $rev \cdot min^{-1}$ below preferred cadence for more than 10 seconds.

3.3 Choice of critical power model and exhaustive test durations

The two-parameter model has been one of the first physiological models applied to human performance [70]. The advantage of the two-parameter linear model is that it provides an accessible and simple application, which enables the characterisation of an individual work- time or power-1/time relationship. Billat et al. [70] recommended the two-parameter linear model to coaches as a valuable and easy to use testing tool

for profiling athletes' potentials. On the contrary, the hyperbolic and in particular the three-parameter models appear to be too complex to be used in training. Furthermore the effectiveness as training or testing tool has not yet been studied in any of the three-parameter models.

Another important consideration was the TTE and maximal effort durations, as CP is highly depended on the exhaustion times used [208,313]. Poole et al.[65] recommended durations between 2 min and 15 min as suitable to fulfil requirements of the CP models such as reaching $\dot{V}O_{2\,max}$ whilst avoiding substrate or motivational limitation. Another advantage is that of a low sensitivity of CP to larger errors in TTE of these durations has been reported in the literature [224,241]. The choice of maximal durations of 3 min, 7 min and 12 min for field testing were consequently justified. These were performed in the format of fixed-duration TT events. Individual TTs are unique races in that riders are performing against the clock and not against other competitive riders. Classic tactics, such as drafting where a rider 'hides' in the slipstream of another rider to preserve energy is not permitted in this race event. Therefore TT type tests presented an exceptional field research opportunity for which study and real-world conditions are closely matched.

The aim of this research thesis is to present applied sport scientists and cycling coaches with an athlete friendly but also user friendly field CP determination protocol. As CP is simply calculated and well researched for cycling ergometery using the linear two-parameter models of work-time and power-1/time, these were chosen within current studies to determine CP. If CP determined in the field proved to be valid, then result analysis had to be accessible to coaches and athletes. Having been identified in the literature as suitable testing methodology in trained cyclists [248] studies permitted participants to self-select their preferred cadence. Finally the number of trials was set as three, allowing SEE values to be calculated.

In summary, to produce consistent results for CP, which are comparable with the literature, the following criteria were used throughout all studies.

- CP was determined using the linear work-time and the linear power-1/time relationship,
- three TTE trials were used to determine CP in the laboratory,
- three maximal efforts were used in the field to determine CP,
- laboratory TTE trails span a duration between 2 – 15 min,
- field TTs included maximal efforts over 3 min, 7 min and 12 min,
- where appropriate familiarisation trials were required prior to data collection,
- TTE laboratory trials during conventional CP determination used each participant's preferred cadence,
- TTE trials were terminated when participants were unable to sustain their preferred cadence by more than 10 RPM for more than 10 s.

3.4 Presentation of research results

Bland and Altman [314] proposed an analysis of agreement, when comparing a new testing method against an established one. If no calibration of the measurement instrument is possible, Bland and Altman stated that neither the new or the established testing method, provide an unequivocal correct measurement and only a significant level of agreement can indicate if the established testing method can be replaced. This new analysis of agreement was developed in response to their criticism of the product-moment correlation coefficient (r) as the indicator of agreement. According to Bland and Altman [314] whilst a significant high correlation does indicate a strong relationship between two measurements, it does not include an indication about the level of agreement unless all data points lie exactly on the line of equality and give the same result every time it is being re-measured. Hence a significant high correlation whilst providing a very strong relationship might substantially lack in agreement when comparing two measurement methods. Bland and Altman [314] consequently proposed to plot the difference between the measurement methods against their mean which also allows for investigating possible relationships between the measured and the 'true' established value. A lack of agreement is summarised by

calculating the bias, if no relationship between the difference and the mean is evident. The bias is estimated by the mean difference and SD of the differences. In normally distributed differences, 95% of differences will fall between the mean difference ± 1.96 SD, which Bland and Altman refer to as 'limits of agreement' (LoA). LoA in this method replaces the SEE. Two measurement methods can be used interchangeably if the differences fall within the mean difference ± 1.96 SD, i.e. between the LoA. The Bland-Altman analysis has been widely used in medical and sports science research. The resulting plots, which illustrate LoA are the presentation of validity or method comparison research and have also been used to present results in reliability studies. Contrarily, Hopkins [315] advocates the use of the correlation approach and a further discussion on this issue can be found in chapter 9. Aforementioned results of this research thesis included both, Bland-Altman and linear regression analysis.

EXPERIMENTAL CHAPTERS

CHAPTER 4: THE 3-MIN ALL OUT TEST DOES NOT PROVIDE A VALID MEASURE OF CRITICAL POWER USING THE SRM ISOKINETIC MODE

4.1 Introduction

CP is traditionally determined via repeated, multi-day, exhaustive exercise tests. This arguably reduces its practical utility [306]. Several authors have investigated the validity of single 'all-out' tests to determine CP [142,254,316]. Given that any exercise bout performed above CP should lead to the gradual expenditure of W′, a sufficiently long all-out exercise bout should lead to the attainment of CP [312].

Based on evidence that W' depletion takes <60 s [102,317], Brickley et al.[316] hypothesized that power output at the end of a 90-s all-out test would be equivalent to CP. However, the final power output reported by Brickley et al. was significantly higher than CP. Subsequently, Vanhatalo et al. [312] investigated the efficacy of a 3-min all-out cycling test and reported that mean power output for the final 30 s (End Power or EP) matched CP. Burnley et al. [90] further demonstrated the reliability of EP using three 3-min tests. These results led Poole [142] to state that "the 3 min test promises to herald a new era for experimental exercise physiology". Indeed, EP has already been used successfully in a range of settings [56,57,318].

The work of Vanhatalo and colleagues [312] suggests that the power profile of all-out cycle exercise has a fundamental physiological basis. If this is true, similar levels of agreement between all-out end-test muscle performance and CP should be observed irrespective of the mode of measurement [19]. However until very recently, published studies of the 3-min test in cycling were conducted using the linear mode setting of the Lode Excalibur Sport ergometer exclusively [56,90,102,319,320]. The degree to which the high level of agreement between parameters reported by Vanhatalo et al. [90] is mechanistic or coincidental has not been independently established. Recently, Bergstrom et al. [321] performed the 3-min test using a Quinton ergometer, also using the linear mode, as well as with the Monark ergometer with 3.5% and 4.5% of body

weight as the set resistance. No agreement between EP or work done above EP (WEP) values using the Quinton and Monark ergometer were observed.

The aim of the present study was to investigate whether EP using the SRM isokinetic mode will provide reliable values of CP. Based on the findings by Vanhatalo et al. [312] no significant differences between conventionally and 3-min all-out test determined CP and W′ values were hypothesised.

4.2 Methods

4.2.1 Participants

Twelve males and one female recreational cyclists (mean ± SD: age 33 ± 7 yr, body mass 78 ± 14 kg, height 1.79 ± 0.1 m, MAP 345 ± 54 W, $\dot{V}O_{2max}$ 5.2 ± 0.9 L·min^{-1}) participated in this study.

Exercise testing was conducted on an electronically braked SRM cycle ergometer (Schober Rad Messtechnik, Jülich, Germany). Participants visited the laboratory seven times. During visit 1, participants completed an incremental test to determine $\dot{V}O_{2max}$ and MAP, as well as a 3-min all-out test for familiarisation. In visits two to seven participants completed three constant work rate trials and three 3-min trials randomly assigned. A standard warm-up of 5-min at 100 W followed by 5-min passive rest and 3-min of unloaded cycling [90] was used prior to each trial. During tests the investigator provided consistent and strong verbal encouragement. A post-test blood [lactate] of ≥ 8 mM HR within 10 beats of age-predicted HR maximum was taken as an indicator for attainment of $\dot{V}O_{2\,max}$ and accepted as a successful test [322]. All visits were separated by a minimum of 24 h and were completed within a maximum period of 21 days. Each participant completed each of their seven tests at the same time of day.

4.2.2 Critical Power determination

CP was determined from three constant work rate tests. at power equivalent to 80%, 100% and 105% MAP. Linear regression was used to provide values of CP and W′ using the work-time ($W = CPt + W'$; equation 1) and the power-1/time ($P = W'(1/t) + CP$; equation 2) model. Values using equation 1 or 2 were consequently termed CP1 and CP2.

4.2.3 3-min all-out tests

During the 3-min test the resistance on the pedals was provided by the SRM ergometer in isokinetic mode, and cadence was therefore maintained at the participants' preferred level throughout. Participants were instructed to attain peak power as quickly as possible from the start, and to maintain maximum power throughout the 3 min. To facilitate this, during the final 10 s of the standard warm-up participants increased cadence by 10-20 $rev \cdot min^{-1}$ above preferred cadence. Consistent with Vanhatalo et al. [312] participants were not informed of elapsed time. EP was calculated as the mean power output over the final 30 s of the test. WEP was calculated as the power-time integral above EP. Blood lactate was sampled and analysed at rest before the test and immediately after its completion.

4.2.4 Statistical analysis

Data were examined using the Shapiro-Wilks' normality test. CoV were derived from log-transformed data [301]. 95% confidence intervals were calculated for each CoV. Repeated measures ANOVA was used to test for significant differences between 3-min trial one and trial two and between trial two and trial three. Consistent with Vanhatalo et al. [312], agreement between: EP and CP1, WEP and W′1, EP and CP2 and WEP and W′2 for both models was assessed using a paired-samples *t*-test and LoA [314,323]. Relationships were assessed using Pearson product moment correlation coefficients. Additionally, linear regression was used to calculate values for SEE to estimate error associated with predicting EP and WEP values. Statistical significance was accepted at $P < 0.05$. Results are reported as mean ± SD unless otherwise stated.

4.3 Results

ANOVA indicated no significant differences in EP between pairs of trials, F (2, 26) = 0.83, P > 0.05. CoV for EP was 4.45% between trials one and two and 4.29% between trials two and three Bland-Altman plots of the test-retest data are presented in Fig. 9. The EP 95% LoA for trials one-two was -2 ± 37 W (0.99 */÷ 1.14 as a ratio) and for trials two-three it was -4 ± 35 W (0.98 */÷ 1.13 as a ratio). The intraclass correlation coefficient (ICC) for EP values was 0.97 (95% CI = 0.92 - 0.99).

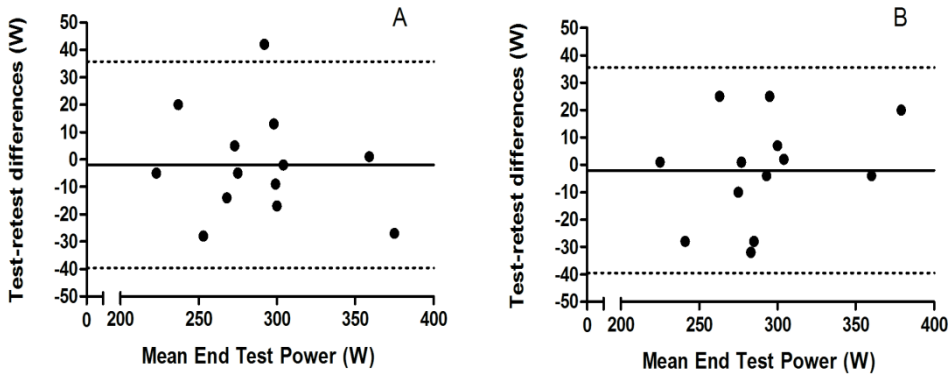

Figure 9. Bland-Altman plots of the End Power test–re-test differences between trials one and two [A] and trials two and three [B]. The solid horizontal lines represent mean bias, whilst the dashed lines represent the 95% limits of agreement

CP and mean EP were normally distributed. Contrarily to the hypothesis, statistically significant differences were observed between EP and CP1 (EP = 290 ± 41 W vs. CP1 = 253 ± 41 W, *t(12)* = -6.16, P < 0.001) and between EP and CP2 (EP = 290 ± 41 W vs. CP2 = 259 ± 38 W, *t(12)* = -4.65, P < 0.001). The SD of the differences for CP1 versus EP was 19 W, providing 95% LoA of 25 ± 48 W (Fig. 10C; 0.87 */÷ 1.16 as a ratio) and for CP2 versus EP the SD of the difference was 18 W, providing 95% LoA between 20 ± 41 W (Fig. 10D; 0.89*/÷ 1.14 as a ratio). The correlation coefficient for EP and CP1 was r = 0.89, P = 0.001 (Fig. 10A) and for EP and CP2 r = 0.90, P = < 0.001 (Fig. 10B). Mean r^2 values for equation 1 were 0.99 ± 0.01 (SEE 2.94 ± 2.23) and for equation 2 0.94 ± 0.06 (SEE 11.96 ± 6.55). The SEE value for the linear relationship between CP1 and EP was 19.49 W, CL (14.49 – 30.22) with an average error prediction of 7.7% and for

CP2 and EP it was 17.10 W, CL (12.79 – 26.52) with an average error prediction of 6.6% (Table 5).

Table 5. Mean Differences, correlation, SEE and LoA for EP and CP

	CP1	**CP2**
Mean Difference EP-CP (W)	37 ± 19	31 ± 18
Correlation (r) with EP	0.89	0.90
SEE (W)	19	17
SEE (%)	7.7	6.6
LoA (W)	25 ± 48	20 ± 41

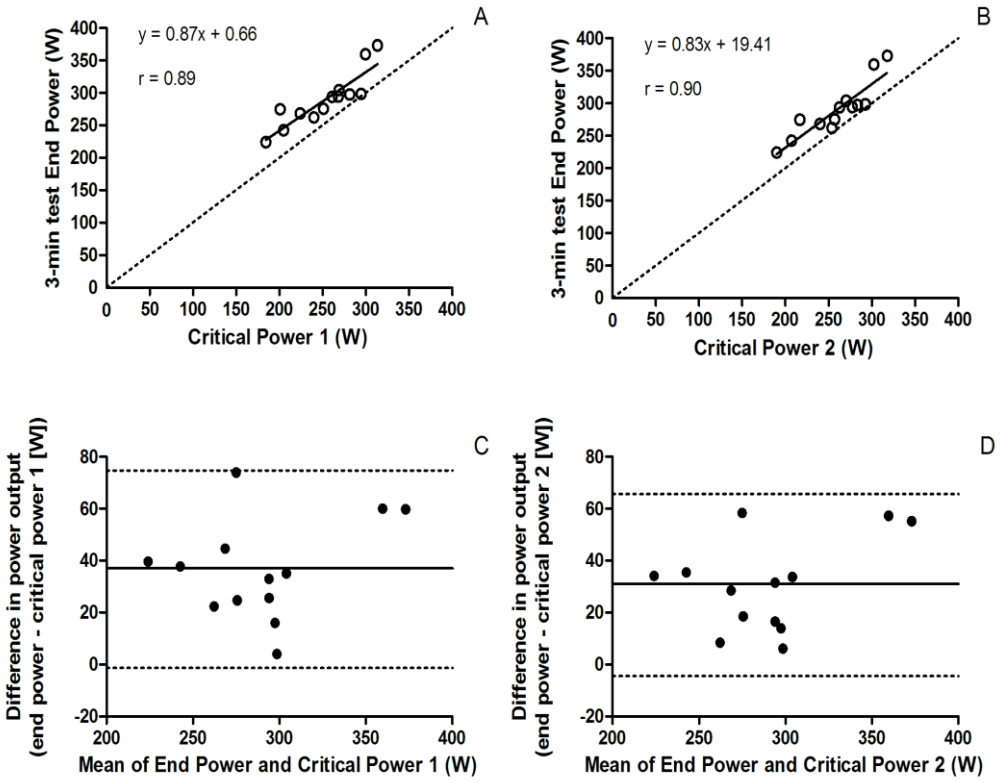

Figure 10. Bland-Altman plots of the relationship (panel A and B) and limits of agreement (panel C and D) between End Power and CP1, and between End Power and CP2. In panel C and D the solid horizontal line represents the mean difference between End Power and Critical Power 1 and 2, and the dashed lines represent 95% limits of agreement.

Significant differences were observed between WEP and W′ 1 (WEP = 12.5 ± 4.3 kJ vs. W′ 1 = 18.6 ± 4.8 kJ, t(12) = -4.65, P = 0.001) and between WEP and W′ 2 (W′ = 16.6 ± 4.8 kJ, t (12)= -3.3, P = 0.006). Therefore the set hypothesis also has to be rejected. The SD of the differences was 4.78 kJ for W′ 1 versus WEP, providing 95% LoA of 3.27 ± 9.1 kJ (Fig. 11C; 0.64 */÷ 1.96 as a ratio) and for W′ 2 versus WEP the SD of the differences was 4.53 kJ, providing 95% LoA of 1.43 ± 6.9 kJ (Fig. 11D; 0.73 */÷ 1.93 as a ratio). The correlation coefficient for WEP and W′ 1 was r = 0.43, P = 0.14 and for WEP and W′ 2 r = 0.48, P = 0.10 (Fig. 11A and 11B). The SEE value for the linear relationship between W′1 and WEP resulted in 4.5 kJ, CL (3.37 – 6.98) with an average error prediction of 24.2% and for W′2 and WEP it was 4.37 kJ, CL (3.27 – 6.78) with an average prediction error of 26.3% (Table 6).

Table 6. Mean Differences, correlation, SEE and LoA for WEP and W′

	W′1	W′2
Mean Difference WEP- W′ (kJ)	-6 ± 4.8	-4 ± 5
Correlation (r) with WEP	0.43	0.48
SEE (kJ)	4.5	4.4
SEE (%)	24.2	26.3
LoA (kJ)	3.27 ± 9.1	1.43 ± 6.9

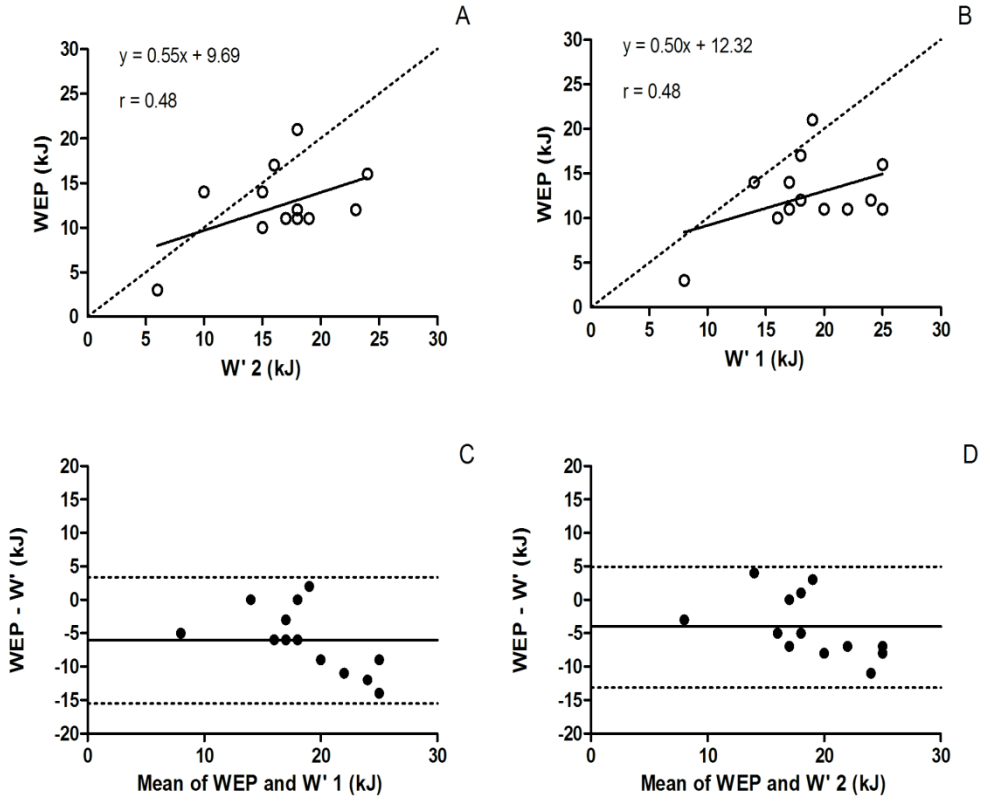

Figure 11. Bland-Altman plots of the relationship (panel A and B) and limits of agreement (panel C and D) between WEP and W′1 and between WEP and W′2. In panel C and D the solid horizontal line represents the mean difference between End Power and CP 1 and 2, and the dashed lines represent 95% limits of agreement.

4.4 Discussion

The above results suggest that a 3-min all-out cycling test using the SRM isokinetic mode does not provide a valid measure of CP. Specifically, the mean power output during the final 30 s of the 3-min all-out test appears to be significantly higher than CP values derived from both work-time and power-1/time models. The 3-min test also appears to underestimate the 'anaerobic' parameter of the CP model (i.e. W′). The above results also suggest that the 3-min all-out test is a reliable measure of EP when studying a trained athletic population.

Given that the CoV values observed were below the proposed boundary of 5% [301] the EP from a 3-min all-out cycling test can be considered to be reliable. Indeed, Burnley et al. [90] suggested that EP is a reproducible measure when reporting a CoV (typical error as a percentage of the mean) only a little lower than that reported here (3% vs. ~ 4.9%). Johnson et al. [324] reported a CoV of 6.7% for the 3-min all-out EP results, and even given this accepted the test as reliable. However, caution should be taken as such a level of variation is unlikely to be acceptable when evaluating the relatively small training-induced changes seen in well-trained athletes [21]. Such a conclusion is supported by limits of agreement analyses which suggest that, with an approximate 95% probability, the differences between the test and re-test of EP in a well-trained cyclist will lie between -40 W and +36 W. Assuming that the bias is negligible, ratio limits of agreement suggest that, between any two tests, EP will differ by as much as 14% in a positive or negative direction. Using a magnitude based analysis, Paton and Hopkins [299] identified that a change of 1.7% in performance impacts on the chances of an elite road TT cyclist winning an event. With an average SEE value for EP-CP 1 and EP – CP 2 of 7.7% and 6.6% respectively, the discrepancy between the two measurement methods in the present study would therefore result in substantial performance differences.

In a heterogeneous group of cyclists, runners and fitness trained participants, Vanhatalo et al. [312] reported no differences between EP (287 ± 55 W) and CP (287 ± 56 W). In contrast, in the present study EP was significantly higher than CP1 and CP2 (37 W and 31 W respectively). Several factors might explain this lack of agreement. Firstly, it is possible that the use of three constant work rate trials resulted

in an inaccurate CP and W′ values. Vanhatalo et al. [312] used five trials, whilst research seeking to model the power-exhaustion time relationship commonly uses four or more trials [234,251]. However, several recent investigations have used three tests for CP and W′ determination [102,254]. According to Hill [79] the decision as to the number of trials used depends on the fitness level of participants as well as their familiarity with all-out exercise. Participants in the present study were accustomed to all-out exercise, a fact which arguably justified the use of three trials in line with Hill's proposal. Strong correlation and low SEE values observed for each participant and model used lend further support to this decision (mean r^2 value for equation 1 was 0.99 ± 0.01/SEE 2.94 ± 2.23 and for equation 2 it was 0.94 ± 0.06/SEE 11.96 ± 6.55). Secondly, as pulmonary gases were not recorded during the 3-min all-out tests, it might be suggested that the study did not meet all three conditions outlined by Jones et al. [19] for the attainment of a successful 3-min test (i.e., that participants did not reach sufficiently high intensity). However, the post-test blood [La] (12.3 ± 3.8 mM) were higher than those reported by Vanhatalo et al. [312] (10.2 ± 2.2 mM). It is fair to assume that participants did perform at an appropriate intensity given that all participants also reached values within 10 beats per minute of their age-predicted maximal HR. Furthermore, the group mean power profile suggests both the very high intensities achieved during the first 60 s of the all-out trials and the subsequent plateau, both of which are vital to the proposed efficacy of the 3-min test (Fig. 12).

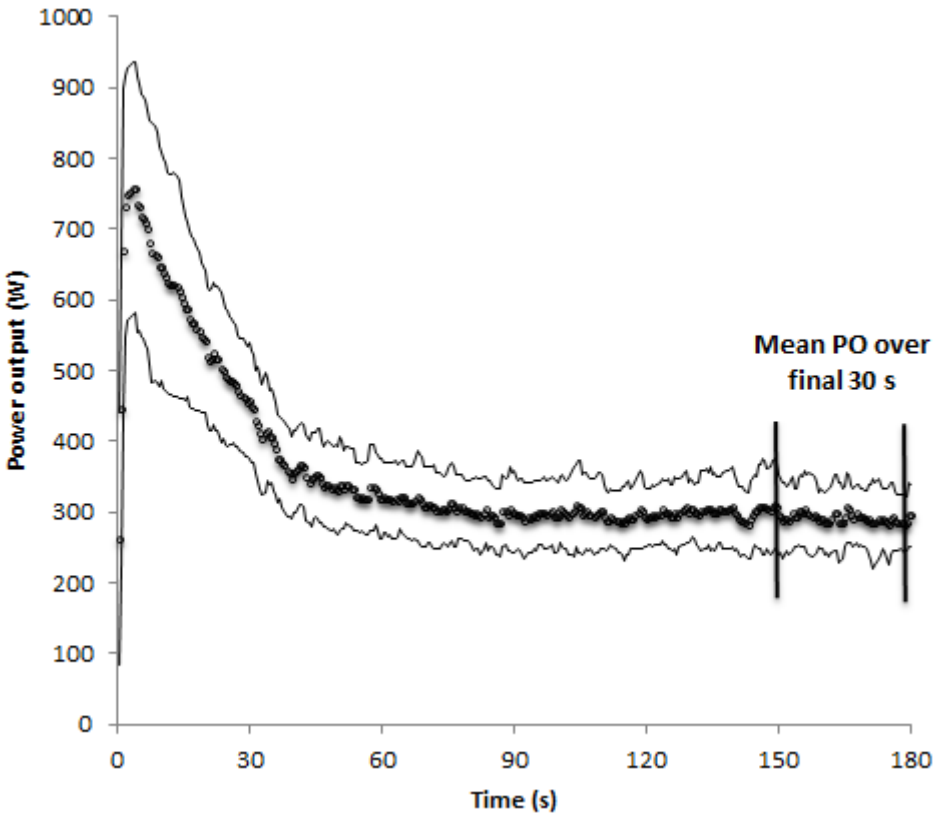

Figure 12. Group mean power profile of the 3-min all-out cycling test. Solid lines represent the standard deviation.

It is also possible that the discrepancy between the present results and those of Vanhatalo et al. [312] relate to the use of different ergometers. The isokinetic mode of the SRM allows the cyclist to maintain a fixed cadence whilst the resistance adapts to any change in pedal force. In contrast, in the linear mode of the Lode the applied resistance is cadence dependent, and in the early stages of the 3-min test, the high power output necessitates a very high cadence. As a participants' ability to produce power declines, so too does cadence. In order to ensure that cadence does not fall to unacceptably low levels, the researcher must adjust the Lode's power/cadence settings. This is done by adjusting the 'linear factor' α in the equation $Power=\alpha*RPM^2$. To date, researchers have adjusted the linear factor such that preferred cadence is reached at $GET+0.5*(\dot{V}O_{2max}-GET)$ (i.e. 50% Δ), where GET is the gas exchange threshold. Given that 50% Δ is very close to CP (46.7% Δ in Vanhatalo et al. [312]), it is possible that the use of a Lode ergometer biases the 3-min all-out test towards an End Power close to GET and therefore to CP.

Values of EP, CP1 and CP2 reported in the present study may have been influenced by the selection of participants. Whilst previous studies [90,102,312] utilised a range of athlete abilities, the present study was conducted on a relatively homogeneous sample of trained cyclists. This suggests that participants in the present study, who are accustomed to high intensity cycling performances, may have been better able to sustain their 3-min effort to ensure that W′ was not depleted.

Mean W'1 (18.3 kJ) and mean W'2 (16.6 kJ) were also higher than in the subject group investigated by Vanhatalo et al. [312] (16 kJ). It is possible that participants with a higher W' take longer to fully expend W' than those with a smaller W' using the isokinetic mode, a mode in which resistance is modulated according to fatigue level whilst maintaining cadence. This might suggest the need for an all-out test longer than 3 min. However, this does not appear to be supported by the power profile in the present study in which power declined towards a relative plateau over a similar time course to that described by Vanhatalo et al. [312]. Bergstrom et al. [321] recently reported 150 s EP derived from a similar method as the 3-min test using a Lode ergometer and which did not significantly differ from EP observed in the original 180 s test duration.

Whilst it is not clear whether or not W' describes a true 'anaerobic work capacity'[223], if valid, the 3-min test would nevertheless provide a valuable tool for the assessment of this parameter. However, the data reported in the present study suggest that the anaerobic parameters derived from the 3-min test significantly underestimate W'. This supports Vanhatalo et al. [312] who reported a WEP markedly below W' in six of ten participants. Vanhatalo et al. [312] suggested that the discrepancy might be the result of different acceleration profiles of the flywheel during all-out and constant work rate exercise when using the Lode ergometer. The suggestion is supported by the results in the present study as the SRM ergometer uses flywheel technology similar to the Lode ergometer.

The generalization of the CP concept to all-out exercise is dependent upon the capacity of the all-out trial to fully deplete W'. Despite satisfying the requirements of the 3-min test [19], it might be possible that the present participants were unable to fully deplete W'. This is surprising given that a maximal accumulated oxygen deficit has been demonstrated following 60-90 s of all-out exercise [306,325]. Such observations led

Brickley et al. [316] and Dekerle et al. [254] to evaluate whether a 90-s all-out test could determine CP in adults and children, respectively. As in the present study, in these studies testing was completed on an SRM ergometer using the isokinetic mode and EP was significantly higher than CP. Despite a plateau being apparent in the final 10 s of the 90-s test, Dekerle et al. [254] suggested that power output continues to decline at the end of the test. This led to the hypothesis that a test of longer duration would allow CP to be attained [316]. The hypothesis is refuted by the observation that the results of the current investigation agree so closely with those obtained when using the 90-s test to derive CP.

Following the protocol proposed by Vanhatalo et al. [312] whilst using an isokinetic mode might explain different outcomes between EP and CP1/CP2. To investigate the robustness of the 3-min all-out test Vanhatalo et al. [320] manipulated the flywheel resistance for participants to achieve EP cadences which were ± 10 rev·min^{-1} different from the original investigation. The authors reported no differences in EP for reduced cadence values and a reduced EP when applying a higher cadence strategy. Consistent with the standard protocol, participants in the present study applied their preferred cadence throughout testing but on average had a higher cadence (95 ± 8 rev·min^{-1}) when compared to Vanhatalo et al. [312] (88 ± 6 rev·min^{-1}). The standard all-out protocol requires participants to adopt their preferred cadence, but the standard test conditions can be sensitive to minor variations in the ergometer resistance settings. Carnevale and Gaesser [247] and Barker et al. [251] investigated the impact of pedalling speed on the power-duration relationship. Both studies reported a lower CP and an unaffected W' when employing a high (100 rev·min^{-1}) vs. a low (60 rev·min^{-1}) cadence strategy. The differences in cadence between the present study and Vanhatalo et al. [312] could be partly responsible for the observed discrepancies between EP and CP1/CP2.

In this study values for EP were consistently higher and values for WEP consistently lower than values for CP and W' respectively and a systematic error can also be suggested as cause of the differences in outcomes. The likely source of this systematic error could therefore be the choice of ergometer.

Based on the reliability of the all-out test, a training intervention study with an average EP as presented of 290 W and a SD value of 43 W would require a sample size of 195 cyclists in order to track a 1.7% performance improvement. However when investigating a performance increase of 6.6%, as suggested by Paton and Hopkins [298] between base and pre-competitive season would require a sample size of 19 cyclists.

4.5 Conclusion

The findings of the present study suggest that the CP concept might not be generalisable to the use of all ergometer models or modes. The 'aerobic' (EP) and 'anaerobic' (W') parameters derived from 3-min all-out cycle test are significantly different to the 'aerobic' and 'anaerobic' parameters derived from the standard work-time and power-1/time CP model. Using only cyclists with a preferred cadence ≥ 90 rev·min^{-1}, or validation studies using rowing, self-powered treadmill ergometers, or track running or cycling might shed some further light into the different outcomes of our study.

The generalisation of the CP concept to all-out exercise based on the findings in this first study had to be questioned. Therefore an investigation of a different all-out duration (i.e. longer) or of possible causes of the 3-min test not providing a valid measure of CP when using the isokinetic mode was not justified to make sufficient progress in the PhD research process. With an overarching aim to develop an athlete-friendly field CP protocol, the research process consequently led to pursue the conventionally accepted laboratory determination method being compared with a similar method adapted to the field.

CHAPTER 5: HIGH AGREEMENT BETWEEN LABORATORY AND FIELD CRITICAL POWER IN CYCLING

5.1 Introduction

In cycling CP is traditionally determined under laboratory conditions by using TTE trials at fixed intensities [60,62,172]. An estimation of the MAP is required to calculate the intensity in question. The total number of trials required to model CP ranges between three and five [72,172,247,260,273,326], although it is usual for at least three trials to be performed, especially in non-elite athletes. Given this, laboratory estimation of CP can be time consuming and potentially disruptive to an athlete's training programme.

The previous study demonstrated that the recently developed 3-min all-out test [312] does not result in a valid determination of CP using a different ergometer or different testing mode. Consequently the test was not considered in the present study for a potential field testing application. Following fundamental scientific principles, the research process instead led to compare a novel field CP protocol with the valid and reliable laboratory-based CP determination method.

Other sports have used field-based determination of the related phenomenon of CV. In swimming, Wakayoshi [327] and Dekerle [270] suggested that the field estimation of CV in swimmers requires only two performances (200m and 400m). In running Kranenburg and Smith [268] determined lab CV using constant load tests on a treadmill that induced exhaustion within 3, 7 and 12 mins, and employed three set distances, each run on an indoor track, to determine field CV. The authors reported that this field-based method of CV determination proved robust, and that field CV was significantly related to 10 km race speed. Again in running, Galbraith et al. [269] developed a field protocol to determine CV also using set distances yielding finishing times between 2 and 12 minutes. Both studies used three trials of durations ranging between 3 and 12 min, and trained participants. Hiyane et al. [267] determined CV using all-out cycling tests over distances of 2, 4 and 6 km resulting in testing times between 1 and 10 minutes.

Data suggest that laboratory and field tests might produce different findings. For example, Jobson et al. [288] reported higher power output values in the field than in the laboratory at given $\dot{V}O_2$ values, whilst Bertucci et al. [292] found an increased gross efficiency and cycling economy in the field when compared to the laboratory. Whilst conditions in the laboratory are more controllable, providing greater reliability, field tests have the advantage of providing greater ecological validity [178,328]. Such validity might be a function of many factors. For example, field tests allow the athlete to perform in an environment consistent with that in which they usually compete, permitting previously acquired effort regulation skills to be employed, therefore reducing the need for habituation to laboratory protocols. Field tests are also relatively unconstrained by the mechanical limitations often imposed by laboratory equipment. Contrast for example cycling on a velodrome with riding a mechanically stable ergometer; in the former the bicycle moves laterally under the rider, and the rider is likely to have developed a handling technique that both controls for this and in doing so optimises the contribution to forward motion of various synergistic and stabilising components of the skeletal- and neuro-muscular systems. These components are less likely to be employed in all but the most ecologically valid laboratory settings. These factors are especially pertinent if the performance in question is measured over a pre-set time, as opposed to time to exhaustion. The former better replicates the characteristics of most sports events, which take place over fixed distances or times and which rarely entail performance to the point of volitional exhaustion. A further benefit of field testing is that it widens access to the techniques and knowledge base of traditionally laboratory-based sports sciences, especially to athletes and coaches with low financial resources.

Whilst TTE protocols have frequently been used in sports research [329,330], they are often associated with low reliability. For example, using untrained participants, Krebs and Power [331] and McLellean et al. [176] reported CoV values ranging between 5.2–56.0% and 2.8–31.0% respectively. Even using well-trained cyclists, Jeukendrup et al. [7] reported CoV values ranging between 17 and 40%. In contrast with TTE protocols, testing protocols that employ a fixed quantity of work, distance or time are reported to be more reliable [5,7,332–335]. However, it was recognised that in conducting

the present study field determination of CP was based on laboratory determination derived through TTE protocols.

Whilst all of the above advantages hold true for many settings, the major limitation with field testing is the lack of control over environmental variables. Even in relatively controlled environments such as indoor athletics tracks, velodromes and swimming pools, variations in temperature and humidity, and disturbances in air or water flow caused by other athletes, can reduce reliability of measurement. This of course becomes a far more serious problem in outdoor road or track cycling where wind and temperature conditions can vary substantially within minutes. In modelling cycling performance in varying wind conditions, Swain [336] used a circuit course which contained equal-length segments of headwind and tailwind. The modelled time for trials was greater in wind conditions compared to no-wind conditions. These greater times resulted from the slowdown of the cyclist into headwinds, which were greater than time saved with tailwinds. Counter to this suggestion Quod et al. [11] compared values of CP observed in the laboratory with those observed in competition, and reported no significant differences between the two ($p = 0.09$, relative difference -0.8%).

To date, only two studies have employed field settings for the estimation of CV [267] and CP [11] in cycling. The purpose of the present study was to use a method similar to that of Kranenburg and Smith [268] and of Quod et al. [11] and to compare values of CP derived through laboratory-based TTE trials with values of CP derived through field tests using trials of set durations. A non-significant difference between CP values determined in the laboratory and CP values determined in the field was hypothesised.

5.2 Methods

5.2.1 Participants

Twelve male and two female recreational cyclists were recruited from local cycling clubs (mean ± SD: age 40 ± 6 yrs; body mass 70.2 ± 6.5 kg; $\dot{V}O_{2max}$ 3.8 ± 0.5 L·min^{-1}; MAP 311± 32.5 W).

5.2.2 Protocol

The study used a within-subject design. During the first laboratory session $\dot{V}O_{2max}$ and MAP values were established. Participants then performed three laboratory-based ergometer TTE tests and three field TTs all randomised (below). To minimise training effects each participant completed all seven sessions within 21 days. A minimum of 24 hours rest was required between individual tests [60,113].

A 24 speed road bicycle (Raleigh Airlite, UK), equipped with a PowerTap Elite wheel (CycleOps, Madison, USA) and a magnet for direct cadence measurement was used to measure work in both laboratory and field tests [42]. The saddle and handlebar were adjusted to suit each participant and settings were replicated exactly during subsequent tests. For laboratory testing the bicycle was attached to a Computrainer (RacerMate, Seattle, USA). To ensure the most accurate power reading the PowerTap was zero-offset prior to each test according to the manufacturer's instructions. According to Bertucci et al. [44] the PowerTap provides a power output accuracy of 1.2 ± 1.3 % and coefficient of variation values of 0.9 to 2.9%. The authors deemed it a valid and reliable measure of power output at submaximal intensities. The same road bicycle and PowerTap Elite wheel was used for all participants and tests.

5.2.3 Laboratory based tests

Cyclists completed three tests to exhaustion on the equipment described above. Capillary fingertip blood samples were collected at rest, immediately post-test and 3 min post-test and analysed for [La]. Consistent with published guidelines [322] a post-test blood [La] of ≥ 8 mM or HR within 10 beats of age-predicted HR maximum was taken as an indicator for attainment of $\dot{V}O_{2max}$ and accepted as a successful test.

5.2.4 Field based tests

Participants were tested over fixed times of 3, 7 and 12 min rather than over set distances on an outdoor velodrome. Tests were completed on separate days and in randomised order. Capillary fingertip blood samples were taken at rest, immediately

post-test and 3 min post-test. A post-test blood [La] of ≥ 8 mM or HR within 10 beats of age-predicted HR maximum was taken as an indicator for attainment of $\dot{V}O_{2\,max}$ and accepted as a successful all-out test [322].

5.2.5 Control of environmental factors

As suggested above, environmental conditions are a major concern in field testing. Consistent with the data reported by Swain [336], it was initially decided that field testing would not take place in wind speeds above 6.6 $m \cdot s^{-1}$, or in rain or otherwise wet conditions. The latter scenario was relatively straight forward to address. However, wind speed so frequently exceeded the 6.6 $m \cdot s^{-1}$ level that cancelling tests on the basis of this criterion would have extended data collection beyond the 21-day criterion and might have introduced other sources of error (e.g., training/de-training effects). Cancelling on the basis of wind speed – which would have led to several tests being abandoned once underway – would likely have led to participants dropping out of the study. Therefore testing went ahead irrespective of measured wind speed, and this issue and decision is discussed further below.

5.2.6 Calculation of critical power and W′

Linear regression was used to provide values of CP and W′ from the results of the laboratory and the field trials using the work-time model [P = W' + (CP · t)] are consequently termed CP1 and W′′1 and using the power-1/time model [*P* = (W' / t) + CP] are consequently termed CP2 and W'2.

5.2.7 Statistical analysis

The distribution of each variable was examined with the Shapiro-Wilks' normality test. Pearson product moment correlation analysis was used to provide an indication of the strength of any relationship between field- and laboratory-derived CP1 and CP2 and W'1 and W'2. Agreement between laboratory and field CP1 and CP2 and W'1 and W'2 was assessed using a paired samples *t*-test and LoA[314,323]. Paired samples *t*-tests were conducted to identify any differences in laboratory and field

based CP TTE trials, in maximal [La], and maximal HR for each equivalent test (80% and 12 min, 100% and 7 min, 105% and 3 min) and for differences between relative percentages of MAP achieved during the laboratory- and field-based CP1 and CP2 tests. Additionally, linear regression was used to estimate error associated with predicting field CP and W′ values [301] Statistical significance was accepted at $P < 0.05$. Results are reported as mean ± SD unless otherwise stated.

5.3 Results

No significant differences were observed between field-based and laboratory-based CP1 (234 ± 24.4W vs. 234 ± 25.5W respectively; t (13) = 0.97, p = 0.924) and CP2 (235 ± 24.1W vs. 236 ± 29.1W respectively; t (13) = 0.81, p = 0.435). Data recorded in the two environments were highly correlated (r = 0.976; p < .05 (CP1) and r = 0.973; p < .05 (CP2)). Mean difference between laboratory- and field-based values for CP1 was 0.17 ± 5.72 W (95% CI, - 3.14-16.61; limits of agreement [LOA], −10.98 to 10.8 W) and for CP2 it was 2 ± 7.72 W (95% CI, -2.28 -25.35; [LOA], -13.88 to 17.3 W) (Table7; Fig. 13).

Table 7. Mean Difference, LoA and SEE for values of laboratory and field CP

	Mean Difference (W)
CP1 lab – CP1 field	0.17 ± 5.72
CP2 lab – CP2 field	2 ± -7.72
	Limits of Agreement (W)
CP1 lab – CP1 field	-10.98 – 10.8
CP2 lab – CP2 field	-2.28 – 25.35
	Standard Error of Estimate (%)
CP1 lab – CP1 field	1.9
CP2 lab – CP2 field	2.5

Significant differences were observed between laboratory- and field-based W'1 (12.2 ± 2.7kJ vs 17.3 ± 5.4kJ respectively, t (13) = -3.98, p = 0.02) and W'2 (11.6 ± 2.7kJ vs. 16.5 ± 4.8kJ respectively; t (13) = -3.93, p = 0.02). The mean difference in W'1 was -5.1 ± 4.8kJ (95% CI, -7.86 – 9.14; [LOA], -14.5 to 4.3 kJ) and in W'2 it was -4.9 ± 4.7kJ (95% CI, -7.58 – 8.94; [LOA], -14.0 to 4.2 kJ) (Table 8; Fig 14). The predication error associated with the laboratory-based and field-based

values of CP/W' was 1.9% (CP1) 2.5% (CP2) and for W′ it was 26.3% (W′1) and 27.6% (W'2).

Table 8. Mean Difference, LoA and SEE for values of laboratory and field W′

	Mean Difference (kJ)
W′1 lab – W′1 field	-5.1 ± 4.8
W′2 lab – W′2 field	-4.9 ± -4.7
	Limits of Agreement (kJ)
W′1 lab – W′1 field	14.5 – 4.3
W′2 lab – W′2 field	-14 – 4.2
	Standard Error of Estimate (%)
W′1 lab – W′1 field	26.3
W′2 lab – W′2 field	27.6

Analysis of blood [La] (mM) revealed significantly higher concentrations for field-based testing when comparing the 100% TTE trial versus the 7 min test ($t(13) = -2.12$, $p = 0.035$) and the 105% TTE trial versus the 3 min test ($t(13) = -2.36$, $p = 0.009$) whilst the 80% TTE trial versus the 12 min test did not result in a statistically significant but low p-value (0.054) (Table 10). Mean PO values for laboratory and field exhaustive trials are shown in table 9. Table 11 illustrates the differences in mean initial 10 s and 30 s power values for field and laboratory-based tests. Table 12 shows mean durations (± SD) and mean distances (± SD) for laboratory and field tests respectively.

Table 9. Mean PO (W) values for laboratory and field exhaustive trials

Participant	Mean 3 min PO (W)	TTE 105% PO (W)	Mean 7 min PO (W)	TTE 100% (W)	Mean 12 min PO (W)	TTE 80% (W)
1	254	251	222	235	210	203
2	349	332	281	307	270	270
3	312	304	263	275	250	244
4	362	356	306	335	296	288
5	241	245	221	215	209	208
6	362	316	303	285	260	245
7	371	325	291	286	270	263
8	292	302	287	285	231	230
9	353	361	314	340	293	293
10	336	323	267	304	253	251
11	315	304	275	265	268	240
12	299	301	249	273	245	238
13	352	323	300	318	284	268
14	370	328	308	309	277	273
Mean (± SD)	**326 ± 42 W**	**312 ± 33 W**	**278 ± 30 W**	**288 ± 35 W**	**258 ± 27 W**	**251 ± 27 W**

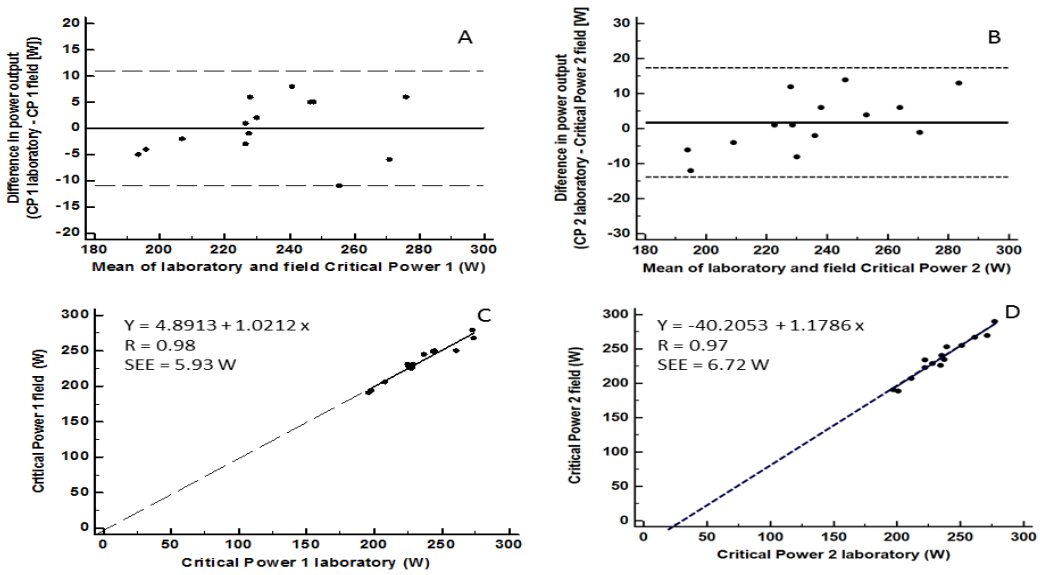

Figure 13. Illustration of the correlation and LoA between CP values derived from laboratory and field tests (C and D) using the Bland Altman test for the relation and bias (solid line) ± 95% limits of agreement (dashed lines) between laboratory-based CP and field-based CP (A and B).

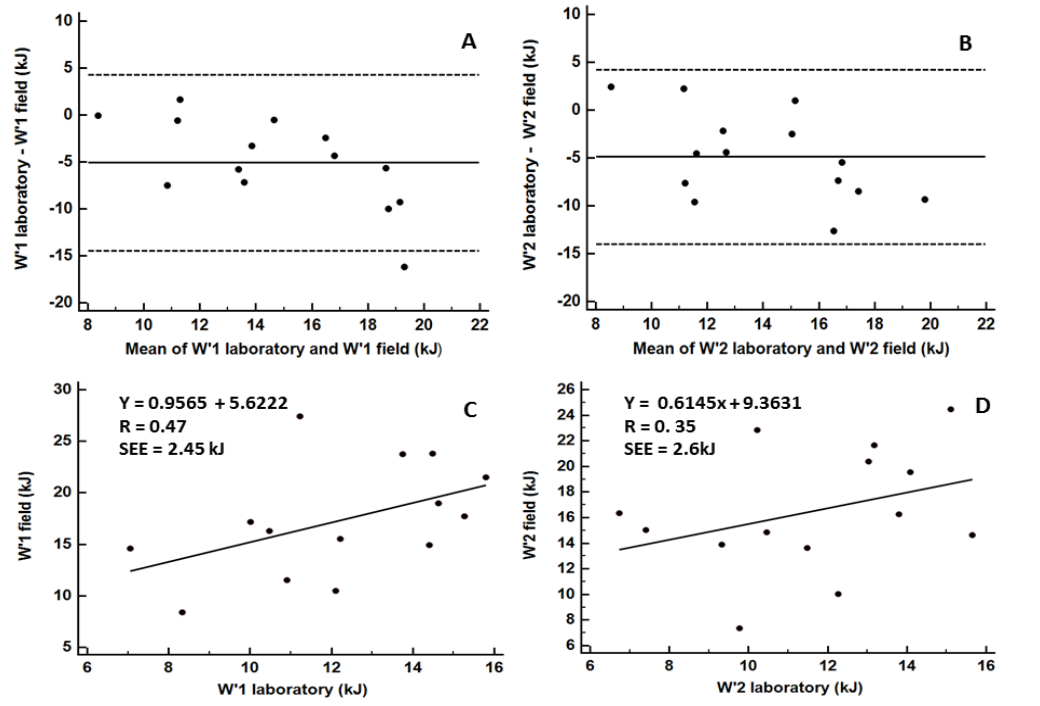

Figure 14. Illustration of the correlation and LoA between W′ values derived from laboratory and field tests (C and D) using the Bland Altman test for the relation and bias (solid line) ± 95% limits of agreement (dashed lines) between laboratory-based W′ and field-based W′ (A and B).

Table 10. Group maximal blood [La] (mM) results, p-values and confidence intervals of the difference

	105% MAP	**3 min**	**p-value**	**Lower- upper 95% confidence intervals of the difference**	
Lactate (mM)	12.26 (± 2.29)	14.22 (± 2.98)	0.009[a]	- 3.34	-0.58
	100% MAP	**7 min**	**p-value**	**Lower- upper 95% confidence intervals of the difference**	
Lactate (mM)	13.55 (±1.99)	13.55 (± 1.99)	0.035[b]	-3.14	-0.14
	80% MAP	**12 min**	**p-value**	**Lower- upper 95% confidence intervals of the difference**	
Lactate (mM)	13.84 (± 3.30)	14.95 (± 3.09)	0.054	- 2.25	-0.021

[a] = sign. different from the mean 105% constant work-rate lactate values (P <0.05). [b] = sign. different from the mean 100% constant work-rate test (P <0.05).

Table 11. Mean initial 10 s and 30 s P values (W) for field and laboratory tests

Field Test	**Initial 10 s P (W)**	**Lab Test**	**Initial 10 s P (W)**
Test 1	12 min = 532 ± 184 W	Test 1	80% TTE = 179 ± 38 W
Test 2	7 min = 624 ± 133 W	Test 2	100% TTE = 174 ± 38 W
Test 3	3 min = 633 ± 148 W	Test 3	105% TTE = 204 ± 34 W
Field Test	**Initial 30 s P (W)**	**Lab Test**	**Initial 30 s P (W)**
Test 1	12 min = 451 ± 132 W	Test 1	80% TTE = 212 ± 45 W
Test 2	7 min = 496 ± 108 W	Test 2	100% TTE = 230 ± 40 W
Test 3	3 min = 524 ± 95 W	Test 3	105% TTE = 279 ± 45 W

Table 12. Mean durations of laboratory TTE trials and mean distance covered of field fixed durations trials (±SD)

	Mean time elapsed (s)	SD time elapsed (s)	Mean distance covered (metres)	SD distance covered (metres)
Lab trials				
80% MAP	725	141	-	-
100% MAP	239	48	-	-
105% MAP	152	30	-	-
Field trials				
180 s	-	-	1858	157
420 s	-	-	4118	233
720 s	-	-	7030	261

Ferguson et al. [86,87] in their CP research added another TTE trial if individual SE values for CP fell above or below that of 3 W. Interestingly individual SE values of ± 3 W in the present study fit well for the linear work-time model of laboratory and field-based CP but lie above (~ 8 W) of the recommended value in the power-1/time-power model.

5.4 Discussion

A mean difference between laboratory- and field-derived CP values of 0.2 ± 5.7 W, suggests that field testing might provide valid determination of CP in cycling and the hypothesis of a non-significant difference has to be accepted. Results support those of Quod et al [11]and Kranenburg and Smith [268].

Using a magnitude based analysis, Paton and Hopkins [299] identified that a change of 1.7% in performance impacts on the chances of an elite road TT cyclist winning an event. With an average SEE value for laboratory-based CP1/2 versus field-based CP 1/2 of 1.9% and 2.5 % respectively, the discrepancy between the two measurement methods in the present study is deemed to be acceptable, considering that a group of elite cyclists would have likely produced lower biological variability [300,337]. The study however did not investigate the reliability of the field CP determination as participants performed each field exhaustive trial on one occasion only. Therefore the

error associated with the reliability of the field method and its sensitivity to track meaningful performance changes is unknown.

Whilst in designing the study, the research student was optimistic that the field-based determination of CP held some promise. However differences between laboratory-based and field-based values of CP were lower than anticipated, especially given that the velodrome used for TT field testing provided no shelter and wind speeds above the 6 $m \cdot s^{-1}$ criterion suggested by Swain [336] were frequently observed. Given the linear function between work completed and time, any deviation of this linearity due to unequal headwind and tailwind speeds would have been identified in the individual CP1/CP2 field-based plots (the mean r-value for field-based CP1 was 0.99 ± 0.001 and for field-based CP2 it was 0.99 ± 0.008). Therefore our data do not appear to support those of Swain, and individual SEE values reported above appear to support this position. Of course, given the relatively small number of participants there is the possibility that the findings are due to chance. Therefore results will need to be tested on different, and ideally larger, samples.

A greater variance in either of the protocols will display heteroscedasticity, i.e. a non-uniform error. This appears to be present in the results (Figures 13 and 14). When correcting for heteroscedasticity, i.e. log-transforming results, this however does not cause any different outcomes. To identify which of the protocols, the laboratory or the field protocol has greater variance, repeated tests are required. This presents a limitation to the study as each protocol was only performed on one occasion.

Another aspect of the data worthy of discussion are the significant differences between laboratory and field-based values of W'1 and W'2. Field-based values of W′1 were on average 5.09 kJ and for W′2 4.89 kJ higher than the respective laboratory values. This is accompanied by overall higher blood lactate responses for field testing (Table 10) and by a difference in power profiles between laboratory and field. Table 6 illustrates the initial 10 and 30 seconds of the all tests. Testing in the field began from a standing start with an initial acceleration phase whilst constant load testing was performed at a constant cadence with the resistance increasing to the required intensity at the beginning of each TTE trial. This difference in power profile is most pronounced in the shorter field trials (3 and 7 min). It can be speculated that during

the acceleration phase in the field participants utilised a higher portion of type II muscle mass resulting in significantly higher power and blood [La] values [338,339] compared to the constant load tests. The relative rate of field-based W' (kJ) expenditure therefore also seems to be greater when compared to the laboratory testing. Skiba *et al.* [89] suggested that *W′* may be primarily a representative of exercising type I and type II muscle mass but that the sum of W' expended at exhaustion is equal to the known total W′′. If this is true than the difference between laboratory- and field-based W' might be explained by the difference in environmental or testing conditions (i.e. standing start, acceleration against air resistance or use of body weight during the acceleration phase) .The research student acknowledges this limitation to the field-based approach, and recognises that a rolling start with paced lap times might provide a more reliable value for W′.

5.5 Conclusion

CP has traditionally been determined in the laboratory. Results of the present study whilst suggesting a significant difference in W′ between the laboratory and the field, also suggest a high agreement in CP between the same environments. The field determination of CP may offer a more ecologically valid and less expensive alternative to traditional approaches, making it a more widely available test. However the data above are from a small sample, and the researcher advises a replication of the study, ideally with a larger subject group.

Results in this study provided a first indication that CP can be determined in the field, i.e. on an outdoor tarmac track. These findings led to the development of the third study, which addressed the elaborate and cumbersome nature of testing to determine CP.

CHAPTER 6: COMPARISON OF INTER-TRIAL RECOVERY TIMES FOR THE MEASURMENT OF CRITICAL POWER IN CYCLING

6.1 Introduction

The previous study suggested that it is possible to obtain a high agreement between CP values determined in the laboratory with those determined in the field, specifically from set duration trials performed on an outdoor velodrome. Furthermore veldrome CP determination as presented does not require a MAP test. It does however conventionally require a 24 h recovery between multi-day maximal efforts and therefore still presents a time consuming protocol, potentially disruptive to an athlete's training and race preparation. Given the above, the practical utility of CP is low [306] such that CP is not routinely assessed in research, clinical exercise testing or athletic performance capacity evaluation [19].

Over and above the 3-min all-out test described above, a number of investigations have utilised alternative inter-trial recovery methods in the estimation of CP and W′. For example, Carter et al. [177] determined CP with an inter-trial recovery period of 4 h between TTE trials, whilst Jenkins et al. [62,68,340,341], Dekerle et al. [254] and Barker et al. [102] all used 3 h inter-trial recovery periods. Housh et al. [66,260] employed a 30 min inter-trial recovery period between two TTE trials performed on the same day, which later replicated by Hinckson and Hopkins [256], who performed three TTE trials. Quod et al. [11] utilised maximal efforts, lasting 6 to 600 s with active recovery periods of 54 to 600 s of cycling at ~100 W between efforts for the estimation of CP and W′. Finally, Bishop and Jenkins [263] were the only researchers who directly compared a 24 h with a 3 h inter-trial recovery time and results demonstrated that a 3 h recovery period provided non-significant different values of CP and W'.

The main concerns in shortening inter-trial rest periods relate to whether a reduced recovery time allows for full W′ restoration [87], whilst avoiding a subsequent performance enhancing primed $\dot{V}O_2$kinetics effect [342]. Investigating the duration of

primed $\dot{V}O_2$ kinetics, Burnley et al. [343] observed an increase in primary $\dot{V}O_2$ amplitude and a reduced $\dot{V}O_2$ slow component in the 2nd of two bouts of heavy exercise separated by 30 – 45 min passive recovery. Burnley's study further showed an association between a significantly elevated baseline blood [La] and primed $\dot{V}O_2$ kinetics. The presence of an underlying mechanistic basis for this association was however questioned. For example, transitioning from rest to exercise, significantly elevated resting blood [La] appears to have little effect on a second bout of maximal effort performance of 5 min duration [344]. Indeed several authors [209,343,345] suggested either no effect or an enhanced effect of elevated muscle [La] on subsequent performance. Supporting this argument, Westerblad et al. [171] posited increased inorganic phosphate levels [P_i] as the major cause of muscle fatigue, as research has demonstrated little direct effect of metabolic acidosis on muscle function at physiological temperatures [346–348]. Ferguson et al. [87] investigated the effects of recovery duration from prior exhaustive exercise and demonstrated that W′ after a 15 min period of cycling at 20 W was restored to ~ 82% (Figure 15). Deciding on a shortest possible inter-trial recovery period therefore provides a challenge for research as priming effects enhances performance [342] whilst elevated [P_i] and incomplete recovery of W′ can contribute to early fatigue, resulting in a performance decrease [347].

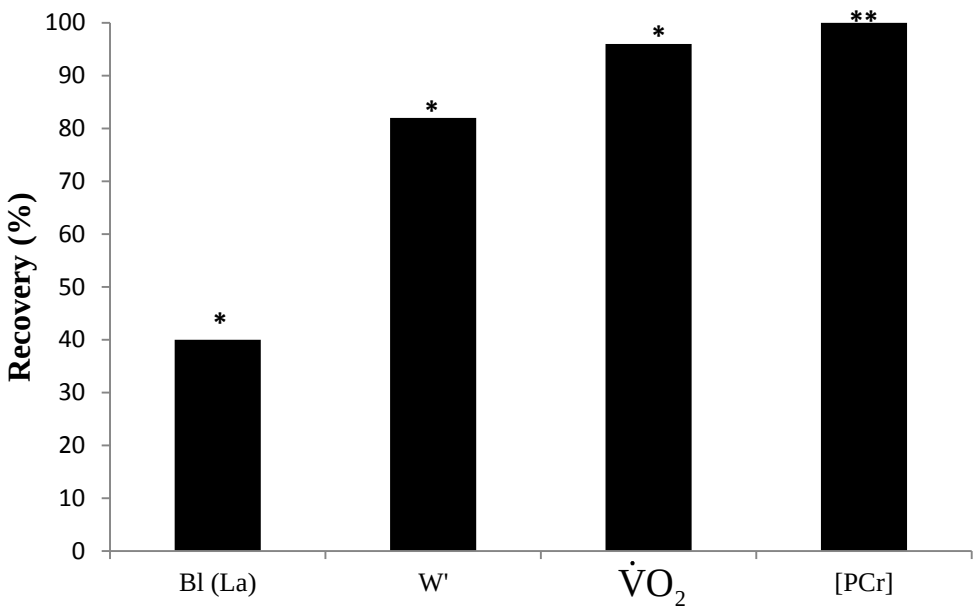

Figure 15. Percentage recovery of Bl [la], W' and $\dot{V}O_2$ following an exhaustive bout of exercise * and [PCr] recovery following high intensity exercise ** after 15 min. Diagram adapted from Ferguson et al. [87] and Forbes et al. [211].

Based on the literature on recovery kinetics of physiological variables such as $\dot{V}O_2$ or Bl [La]) and based on common practice in CP investigations, the present study compared values of CP and W′ derived using the conventional period of 24 h recovery (protocol A), with an alternative 3 h (protocol B) and 30 min (protocol C) recovery. A high level of agreement between CP derived from the different protocols and W′ was hypothesised.

6.2 Methods

6.2.1 Participants

Nine competitive, recreational road cyclists (mean ± SD: age 33 ± 8 yr, body mass 78 ± 10 kg, Maximal Aerobic Power (MAP) 358 ± 35 W, $\dot{V}O_{2max}$ 3.9 ± 0.4 L·min^{-1}) participated in this study. The study was approved by the University Ethics Committee of the host institution. Prior to providing written informed consent and participation, cyclists were fully informed of the nature and risks of the study.

6.2.2 Protocol

During the first visit $\dot{V}O_{2max}$ and MAP values were established. In randomised order, each cyclist then completed three CP protocols. Protocol A used a 24 h inter-trial recovery (three visits), protocol B a 3 h inter-trial recovery (one visit) and protocol C a 30 min inter-trial recovery (one visit). During all tests, participants were blinded to TTE trial intensities and elapsed times. Participants refrained from heavy exercise in the 24 h prior to all tests and from food intake in the 3 h prior to all tests. To minimise training effects, all visits were separated by a minimum of 24 h and were completed within a maximum period of 14 days (Figure 16). Each cyclist completed each of their six visits at the same time of day.

A road bicycle equipped with a PowerTap Elite wheel (CycleOps, Madison, USA) and a magnet for direct cadence measurement was used in this study [42]. The road bicycle was attached to a Computrainer (RacerMate, Seattle, USA). The saddle and handlebar were adjusted to replicate each participant's own bike settings as closely as possible. Settings were replicated exactly during subsequent tests. The PowerTap device was zero offset prior to each test according to the manufacturer's instructions.

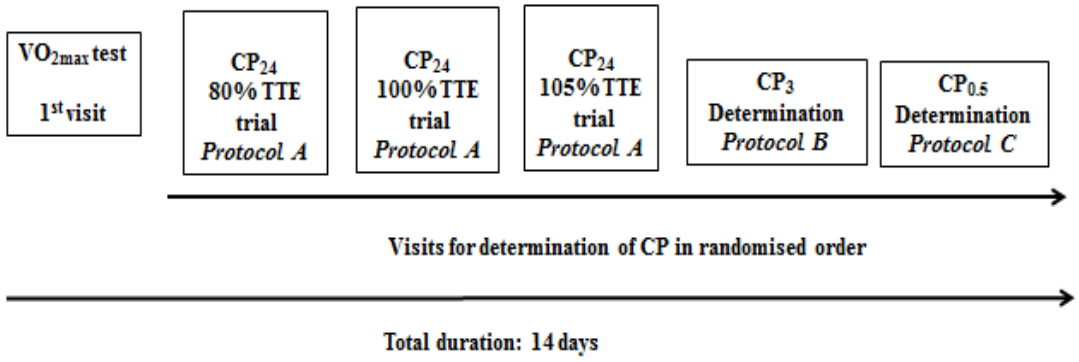

Figure 16. Illustration of testing protocol

6.2.3 Critical power determination

Each protocol required cyclists to complete three TTE trials on the equipment described above. Protocol A used a randomised TTE trial order, with protocol B and C utilising a lowest (80% MAP) to highest work rate (105% MAP) order. Under protocol C, participants were allowed to continue unloaded cycling for 3 minutes before dismounting the bicycle and resting passively in a seated position at the end of the 80% and 100% MAP TTE trials. HR ($b \cdot min^{-1}$), PO (W) and cadence ($rev \cdot min^{-1}$) were recorded continuously via the PowerTap, and expired gases were continuously sampled through the gas analyser. Fluid intake was permitted ad libitum, with cyclists being allowed to consume a minor meal of their choice immediately post TTE trials under conditions of protocol B or snack, such as a piece of fruit under protocol C conditions. All cyclists reached their individual $\dot{V}O_{2max}$ value ($\pm$ 0.08$L \cdot min^{-1}$), a post-test blood [La] of $\geq$ 8 mM and a HR within $\pm$ 5 beats of their maximal HR values established during the $\dot{V}O_{2max}$ test.

6.2.4 Calculation of critical power and W′

Linear regression was used to calculate CP and W' using the work-time ($W = CPt + W'$; equation 1) and the power-1/time ($P = W'(1/t) + CP$; equation 2) models. Results using equation 1 or 2 were consequently termed CP1/W′1 and CP2/W′2. Results using the 24 h inter-trial recovery method were termed CP_{24}/ W'_{24} and for the shorter inter-trial recovery durations of 3 h and 30 minutes were termed CP_3/ W'_3 and $CP_{0.5}$/ $W'_{0.5}$ respectively.

6.2.5 Statistical analysis

Data were examined using the Shapiro-Wilk normality test. Pearson product moment correlation analysis was used to provide an indication of the strength of relationship between the different inter-trial protocols for CP or W′ using equation 1 (CP1/W′1) and equation 2 (CP2/ W′2). Agreement between different testing protocols for CP_{24}/W'_{24}, CP_3/W'_3 and $CP_{0.5}$/$W'_{0.5}$ was assessed using a repeated measures ANOVA test and LoA[314,323]. A repeated measures ANOVA test was also used to assess

differences between the protocol specific durations of TTE trials and resting and post-exercise blood [La] between and within different protocols. Linear regression was used to calculate values for SEE to estimate error associated with predicting CP and W′ values. Statistical significance was accepted at $P < 0.05$. Results are reported as mean ± SD unless otherwise stated.

6.3 Results

CP and W' were normally distributed. Repeated measures ANOVA demonstrated no significant differences between CP1 and CP2 derived through the three inter-trial recovery protocols ($P > .05$). Significant differences were observed between $W'1_{24}$ and $W'1_{0.5}$, between $W'1_{3}$ and $W'1_{0.5}$ and between $W'2_{3}$ and $W'2_{0.5}$ ($P < .05$). Mean SEE values for $CP1_{24}$ were 2 ± 3 W, for $CP1_{0.5}$ 3 ± 1 W, and for $CP1_{0.5}$ 1 ± 1 W. Table 13 and 14 illustrates mean difference and 95% LoA for all results and models with Table 15 and 16 illustrating mean CP and W′, SEE (± Confidence Limits) and average prediction errors for each protocol. Using equation 1, mean r^2 for protocol B was 0.99 ± 0.02 (SEE 3 ± 1 W) and for protocol C it was 0.99 ± 0.01 (SEE 1 ± 1 W).

Mean duration for 80% TTE trials was 619 ± 33 s, for 100% TTE trials 230 ± 18 s and for 105% TTE trials 165 ± 16 s. Reported for each protocol, 80% TTE, 100% TTE and 105% under protocol A resulted in 650 ± 237 s, 251 ± 81 s and 179 ± 59 s. Respectively under protocol B mean durations were 623 ± 213 s, 222 ± 81 s and 169 ± 49 s and under protocol C they were 578 ± 170 s, 210 ± 79 s and 143 ± 23 s. Significant differences ($P < 0.05$) were observed for mean resting blood [La] in protocol C between 80% TTE trials and both 100% and 105% TTE trials but also between protocol C 100% and 105% TTE trials and their protocol B and C counterparts. For post blood [La], significant differences were observed between protocol A 80% TTE trial and 105% TTE trials in protocol B and C (Table 17).

Table 13. Mean Difference (±SD), 95% Limits of Agreement between CP results

	Mean Difference (W)	95% LoA (W)
$CP1_{24}$ vs. $CP1_{3}$	3 ± 6	-2 ± 8
$CP1_{24}$ vs. $CP1_{0.5}$	- 2 ± 9	-9 ± 6
$CP1_{3}$ vs. $CP1_{0.5}$	- 5 ± 8	-11 ± 1
$CP2_{24}$ vs. $CP2_{3}$	3 ± 6	-2 ± 8
$CP2_{24}$ vs. $CP2_{0.5}$	-2 ± 12	-12 ± 7
$CP2_{3}$ vs. $CP2_{0.5}$	-5 ± 10	-14 ± 3

Table 14. Mean Difference (±SD), 95% Limits of Agreement between W′ results

	Mean Difference (kJ)	95% LoA (kJ)
$W'1_{24}$ vs. $W'1_{0.5}$	0.1 ± 3.5 *	-2.6 ± 2.9
$W'1_{24}$ vs. $W'1_{3}$	3.7 ± 4.6	0.1 ± 7.2
$W'1_{3}$ vs. $W'1_{0.5}$	3.6 ± 3.5 *	0.9 ± 6.2
$W'2_{24}$ vs. $W'2_{30}$	0.2 ± 3.9	-2.8 ± 3.2
$W'2_{24}$ vs. $W'2_{3}$	3.9 ± 5.7	-0.5 ± 8.3
$W'2_{3}$ vs. $W'2_{0.5}$	3.7 ± 4.2 *	0.5 ± 7.0

***** Significantly different ($P < 0.05$)

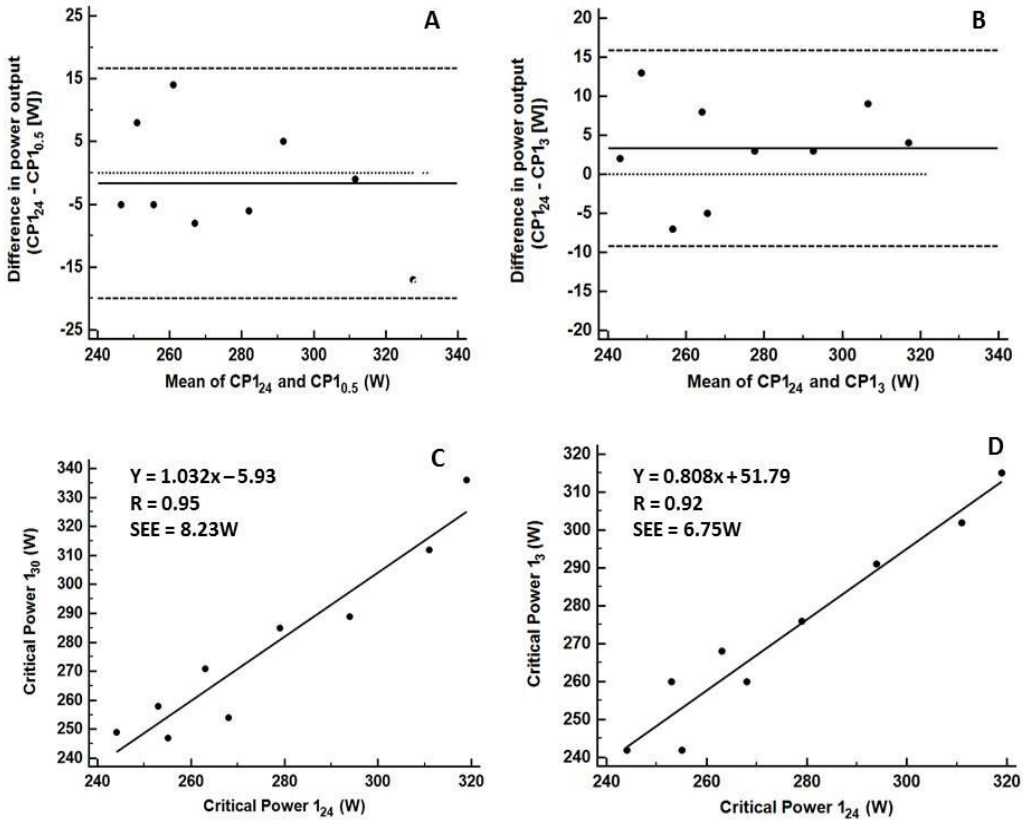

Figure 17. Illustration of the correlation and LoA between CP_{24} and $CP_{0.5}$ (C) and between CP_{24} and CP_3 (D) and the residuals between $CP1_{24}$ and $CP1_{0.5}$ (A) and between $CP1_{24}$ and $CP1_3$ (B) using the Bland Altman test for the relation and bias (solid line) ± 95% limits of agreement (dashed lines).

Table 15. Mean CP (±SD), Standard error of estimates and average prediction errors (%)

	Mean (W)	**SEE (W)**	**Lower CL**	**Upper CL**	**Average pred. error (%)**
$CP1_{24}$ vs. $CP1_3$	276 ± 27 vs. 273 ± 26	6.75	4.76	12.12	2.45
$CP1_{24}$ vs. $CP1_{0.5}$	276 ± 27 vs. 278 ± 31	8.23	5.80	14.79	2.98
$CP2_{24}$ vs. $CP2_3$	277 ± 26 vs. 274 ± 25	6.68	4.71	12.01	2.41
$CP2_{24}$ vs. $CP2_{0.5}$	277 ± 26 vs. 279 ± 33	10.05	7.09	18.07	3.63

Table 16. Mean W′ (±SD), Standard error of estimates and average prediction errors (%)

	Mean (kJ)	SEE (kJ)	Lower CL	Upper CL	Average pred. error (%)
$W'1_{24}$ vs. $W'1_{30}$	15.3 ± 4.6 vs. 15.2 ± 4.4	3.59	2.53	6.45	23.46
$W'1_{24}$ vs. $W'1_{0.5}$	15.3 ± 4.6 vs. 11.6 ± 3.0	4.66	3.29	8.37	30.46
$W'2_{24}$ vs. $W'2_{3}$	15.2 ± 4.7 vs. 15.0 ± 4.2	3.89	2.74	6.99	25.59
$W'2_{24}$ vs. $W'2_{0.5}$	15.2 ± 4.7 vs. 11.3 ± 3.5	5.00	3.53	8.99	32.89

***** Significantly different to W'_{24} ($P < 0.05$)

Table 17. Group mean resting blood [La] (mM) results for all protocols

Prior TTE trial	**Lactate (mM) 80% TTE trial**	**Lactate (mM) 100% TTE trial**	**Lactate (mM) 105% TTE trial**
Protocol A	1.5 ± 0.6	1.5 ± 0.7	1.4 ± 0.6
Protocol B	1.5 ± 0.5	1.8 ± 0.8	1.5 ± 0.5
Protocol C	1.2± 0.3	3.5 ± 0.8$^{*/**}$	4.1 ± 1.3$^{*/**}$
Post TTE trial	**Lactate (mM) 80% TTE trial**	**Lactate (mM) 100% TTE trial**	**Lactate (mM) 105% TTE trial**
Protocol A	12.5 ± 1.5	11.8 ± 3.0	10.5 ± 2.8
Protocol B	13.2 ± 2.7	11.0 ± 2.6	10.1 ± 2.3$^{¥}$
Protocol C	11.5 ± 3.1	10.4 ± 2.2	9.2 ± 2.0$^{¥}$

* Significantly different to protocol C 80% TTE trial resting value ($P < 0.05$)
** Significantly different to respective protocol A and protocol B 100% and 105% TTE trial resting values ($P < 0.05$)
$^{¥}$ Significantly different to respective 80% TTE trials ($P < 0.05$)

6.4 Discussion

This study investigated whether 3 h or 30 min inter-trial recovery times are sufficiently long enough to provide accurate determination of CP and W′, when compared to the standard 24 h inter-trial recovery values. Results suggest that inter-trial recovery periods as short as 30 min provide valid results of CP, but not of W′. These findings are supported by Galbraith et al. [264] who in running demonstrated that CV but not the ARD can be determined accurately when using the same between exhaustive trial recovery duration of 30 min. A 5% [301] and 10% [302] CoV have been

cited as an acceptable upper limit in sports science reliability studies. Assuming, that the different protocols measure the same variable, a CoV of 1.93 (± 0.8%) for CP1 and a CoV of 2.29 (± 1.1%) for CP2, support a suggested acceptance for the inter-changeability of protocols. Small mean differences and 95 % limits of agreement for $CP1_{24}$ vs. CP_3 (-2 ± 8 W), $CP1_{24}$ vs. $CP_{0.5}$ (-9 ± 6W), $CP2_{24}$ vs. $CP2_3$ (2 ± 8 W), and $CP2_{24}$ vs. $CP2_{0.5}$ (12 ± 8 W) also suggest an acceptable level of agreement between the 24 h and shorter recovery duration protocols (Table 13) and the hypothesis of non-significant differences between CP values can be accepted. These findings are supported by Bishop and Jenkins [263] who after a familiarisation trial determined CP in untrained individuals. The researchers did not find any significant CP values between a 24h and a 3 h recovery period. W′ resulted in an unacceptable low level of agreement (Table 14) and high average prediction errors for both alternative protocols (~27% for W'_3 and ~29% for $W'_{0.5}$; table 16) which is inconsistent with the findings by Bishop and Jenkins [263] and the hypothesis of non-significant differences between W′ has to be rejected. Paton and Hopkins [299] identified that a performance change of 1.7% impacts on the chances of an elite road TT cyclist winning an event. Table 15 presents the average prediction error for $CP1_{24}$ vs. $CP1_3$ of 2.45% and for $CP1_{24}$ vs. $CP1_{0.5}$ of 2.98% (for CP2 2.41% and 3.63% respectively). It is fair to assume that, in comparison with the participants in the present study, lower biological variability in elite cyclists would likely result in even lower SEE values and negligible differences between protocols [22].

With only one study addressing CP inter-trial recovery test manipulation in cycling [263], pertinent investigations focus on the effects of prior exercise bouts on consequent performance. Whilst minimal effects are evident in the current study, previous research suggests that prior exercise such as a TTE trial can be detrimental to subsequent exercise, when it is too intense [349], or when recovery periods are too short [86,350]. Alternatively, an enhanced performance effect on severe exercise tolerance has been observed after moderate and heavy prior exercise with the application of a resting period of ~ 10 min between efforts [213,351] or no rest provision [177]. However, this effect is not present after prior sprint exercise [352], 8 min after low and high intensity exercise [353], and after 10 min of heavy exercise [213]. Employing a 20 minutes recovery period Bailey et al. [342] found a ‘large’ 30% TTE performance increase in a

second bout of severe exercise. Basing their research on a performance enhancing priming effect duration of 30-45 min [343], Bailey et al. [342] found that primed $\dot{V}O_2$ kinetics per se did not seem to have caused the performance improvement. Faster $\dot{V}O_2$ kinetics were also associated with a decrease in exercise tolerance when applying a 3 min recovery period between the same two bouts of severe exercise. It appears that TTE trials in the present study did not cause such 'priming' effect on CP. A number of authors [86,87,213,350] suggested that prior severe exercise alters W′, but not CP, during subsequent high intensity exercise. These suggestions confirm our findings as no performance enhancements, but an alteration of W′ was observed. Both 100% and 105% TTE trials were located in the severe domain but did not alter TTE durations. Low agreement levels (Table 14) and high prediction errors for protocol B and C, plus high individual SEE values (Table 16) for each participant under protocol C conditions, confirm the hypotheses for W′.

Resting blood [La] was significantly elevated for both, the 100% and 105% TTE trials in protocol C (3.5 ± 0.8 mM and 4.1 ± 1.3 mM respectively; table 17), but did not seem to exhibit any performance enhancement as suggested by Burnley et al. [343]. Even though not reaching statistical significance, protocol C 105% TTE trial durations on average were ~ 36 s shorter when compared to protocol A. This was similar when comparing protocols B and C 105% TTE trial durations. According to Burnley et al. [343], an elevated blood [La] indicates primed oxygen kinetics which not just result in a decreased oxygen deficit, but also sparing of substrate level phosphorylation, and a reduced slow component, causing a performance enhancement in subsequent exercise bouts [213]. Nielsen et al. [345] suggested that acidosis caused by elevated blood [La] actually protects the muscle from fatigue which is due to the loss of muscle K^+. This is supported by Bangsbo et al. [354] who further suggested that neither muscle glycogenolysis nor glycolysis are reduced because of acidosis and that it is the accumulation of K^+ in the muscle interstitium which is a major factor in the development of fatigue. This accumulation results in a change of membrane potential affecting the excitability and consequent performance of the muscle [345]. An optimal [La] of ~ 2-3 mM has been suggested [213], which through the preservation of muscle K^+ may enhance performance, conversely levels of ~ 6 mM do not seem to significantly alter time to exhaustion [353]. Blood [La] prior

to some TTE trials were ~3-4 mM, and so even though elevated, were unlikely to significantly affect the subsequent performance. However, Ferguson et al. [87] suggested that lactate recovery kinetics are slower than those of W′, resulting in continued lactate processing after full replenishment of W′. This implies that full recovery was not evident in protocol C, since values of W′ in this protocol were significantly smaller when compared to protocol A ($p = 0.02$).

Forbes et al. [210] investigated the effects of recovery time on PCr kinetics during repeated bouts of heavy-intensity exercise, and found that 6 – 15 min was long enough for the full restoration of [PCr]. In the present study, a 30 min inter-trial recovery period therefore should have been sufficiently long enough for the restoration of [PCr] to resting levels and also for the removal of elevated [P_i], as highlighted by Westerblad et al. [171]. Whilst [PCr] are only one constitute of W′ [51] results in the present study suggest an incomplete restoration, which might have fractionally contributed to the differences in W′ between the protocols.

There appears to be a lack of consensus as to the true nature and role of W′. W′ defined as a finite amount of energy [200] and originally thought to be comprised of energy derived from substrate-level phosphorylation utilizing intramuscular high-energy pools and anaerobic glycolysis, with an additional contribution from myoglobin- and haemoglobin-bound oxygen stores [51] was believed to result in exhaustion, when depleted [52,65]. More recently W′ has been suggested to represent the accumulation of fatigue- related metabolites, such as [P_i], [H^+] and [K^+], to some critical tolerable limit [18,199,200]. According to Coats [199], depletion of W′ resulting from a prior bout of severe exercise influences subsequent performances around CP intensity. This was seen in the present study in that 100% and 105 % test durations under protocol C conditions were shorter than those of their protocol A counterparts with no significant difference in CP evident.

Challenging a finite capacity-based explanation for tasks to failure, Ferguson et al.[87] explored the effects of an exhaustive conditioning bout on CP and W′. Identifying a multi-variable character of W′ with complex recovery kinetics Ferguson et al [87]

demonstrated that W′ reflects an ability to exercise under increasing levels of fatigue caused by its own utilisation. Ferguson et al.[87] found no differences for CP but for W′, when employing protocols of 2, 6 and 15 min recovery between a one 6 min W′ depleting exercise bout followed by TTE trials. Consistent with these findings Parker Simpson et al. [57] found that prior exercise at intensities above CP, i.e. severe intensity significantly reduces W′ whilst not affecting CP in all-out exercises. As suggested by the results of the present study, the robustness of CP means that the reductions in time to fatigue after prior exhaustive exercise seem to be solely dependent on W′, resulting in less than stable values for this variable.

Investigating the influence of moderate hypoxia on high intensity exercise tolerance, Dekerle et al. [84] found that the ranges of TTE did not differ between normoxic and hypoxic conditions. However CP was significantly affected (mean – 13%) under hypoxic conditions with W′ not demonstrating a significant difference but exhibiting large intra-individual responses (-36 to + 66%). Like Vandewalle et al. [355], Dekerle et al. [84] consequently questioned whether the two-parameter model allows a valid estimation of W′ and, as suggested by Ferguson et al. [87] if W′ actually represents a finite energy store, as CP and W′ did not seem to be entirely independent using the two-parameter model. At the present time it can only be speculated whether it is our incomplete understanding of W′ or an insufficient mathematical model which is responsible for the different outcomes.

The present study used an arbitrary inter-trial recovery period of 3 h and 30 min, basing these durations on previously published studies [66,260] and on studies addressing issues such as W′ [342] restoration and primed $\dot{V}O_2$ kinetics [261,343]. W′ under protocol B and C conditions resulted in a low level of agreement and high prediction errors, possibly identifying a non-complete restoration of this ‘finite amount of energy’. It can however only be speculated whether or not a longer than 3 h recovery period would have resulted in smaller differences between W′ values.

6.5 Conclusion

CP has traditionally been determined using 24 h inter-trial recovery periods. Results of the present study, whilst suggesting a significant difference in W′ between the protocols, also suggest a high agreement and a low prediction error for CP using 3 h and 30 min inter-trial recovery periods. With the W′ conundrum requiring further investigation, CP appears to be robust to the manipulation of TTE recovery times. A substantially reduced inter-trial recovery period as low as 30 min consequently widens the practical utility of CP determination for scientists, coaches and athletes.

Together with study 2, sufficient evidence was accumulated to investigate a more athlete-friendly field, i.e. road determination method of CP. The final study of this research thesis addressed the gap between track and road CP determination using the newly developed shortened method presented in this current study.

CHAPTER 7: VALIDITY AND RELIABILITY OF DIFFERENT FIELD TESTING METHODS FOR THE DETERMINATION OF CRITICAL POWER

7.1 Introduction

The previous studies established that CP can be determined in the field, i.e. on the track and that it is possible to perform CP laboratory testing with a 3 h and with a 30 min TTE inter-trial recovery method. The latter however requires a prior MAP test to calculate relevant TTE intensities. Whilst providing some greater practical utility than the conventional CP determination method these findings yet do not fully bridge the gap between research laboratory and real-world road cycling.

Study two was conducted on an outdoor cycling velodrome [356]. Results therefore do not indicate whether the agreement between CP values also holds true for road cycling, where the terrain can be flat or undulated. When testing in the field, several confounding issues have to be considered. For example, high wind resistance can account for as much as 80–90% of the metabolic cost of cycling [357]. Additionally, drag force increases as the square of the riding speed and as the cube of power output [358], and is a function of a riders frontal surface area [359], which will influence cycling velocity. However, during uphill cycling, gravitational resistance, which is proportional to body mass, becomes the more dominant resistance force as velocity is low [360]. In short, cycling velocity in the field is significantly influenced by the above factors that are not present during laboratory testing.

However, unlike velocity, PO is independent of external conditions, such as wind and potentially offers a more appropriate testing variable when designing field testing protocols [294]. A significant contribution to this research topic was made by Quod et al. [11] who investigated the differences in PO values produced in the laboratory with those produced during road races. The study recruited 10 experienced cyclists who were assessed in their maximal capacity to produce power over set durations. These durations which cyclists typically encounter during road races were set at 6, 15, 60, 60, 240 and 600 s. The final three maximal efforts were

also used to model CP and W′. Road race data were downloaded from individual SRM power meters and analysed using the WKO TrainingPeaks software (Peaksware LLC, v3+, Boulder, USA). Each pair of laboratory and field data were analysed and did not reveal a significant difference. The same results were found for laboratory and field determined CP and W′ but the study failed to report values of LoA for these parameters. It has to be noted that the lowest duration of 60 s used in the modelling process of CP and W′ does not comply with the requirements of CP determination as set by DiPramperio [208], i.e. attainment of $\dot{V}O_{2max}$. Furthermore the power profile testing, which included relevant CP and W′ efforts was not validated against conventional determination standards, i.e. a 24 h recovery in between maximal efforts. Quod et al. [11] for the purpose of power profiling cyclists instead utilised an active recovery performed at 100 W for individual break periods of 330 s, 480 s and 600 s between relevant maximal efforts of 60 s, 240 s and 600 s respectively.

Based on Study 2 of determining track CP, the purpose if this final study was to investigate whether these findings also hold true for road cycling. Using the method investigated in Study 3, laboratory CP was determined using a 30 min intra-trial recovery testing method. Laboratory CP was compared with that determined from maximal road efforts of 12 min, 7 min and 3 min duration. The study further aimed to compare CP obtained from the highest 12, 7 and 3 minute power outputs recorded during a five week training period, with that from the laboratory. Finally the reliability of each respective CP field protocol was investigated. Based on the previous study findings and an independency of power to field environmental conditions a non-significant difference and high level of agreement for CP values but not for W′ was hypothesised. Further a good level of reliability across repeated field trials was hypothesised for CP and for W′.

7.2 Methods

7.2.1 Participants

Participants in this study were competitive, recreational road cyclists with a minimum of two years racing experience [minimum of 250–300 km or 10 h training

volume per week]. Eleven moderately trained cyclists (mean ± SD: age 32 ± 8 yr, body mass 76.9 ± 14.9 kg, MAP 351 ± 37 W, $\dot{V}O_{2\,max}$ 51.4 ± 9.8 mL·kg^{-1}·min^{-1}) completed protocol one. Due to one drop out, 10 participants completed protocols two and three (mean ± SD: age 32 ± 8.9 yr, body mass 75.3 ± 15.1 kg, MAP 346 ± 36 W, $\dot{V}O_{2\,max}$ 51.9 ± 10.3 mL·kg^{-1}·min^{-1}).

7.2.2 Method

Participants' road bicycles were equipped with a PowerTap Elite wheel (CycleOps, Madison, USA) and a magnet for direct cadence measurement [42]. For the laboratory tests, the same road bicycle was attached to a Computrainer (RacerMate, Seattle, USA). The PowerTap device was zero offset prior to all trials according to the manufacturer's instructions (15). During two visits to the laboratory, $\dot{V}O_{2\,max}$ and MAP values and laboratory CP (protocol A) were determined. Participants refrained from heavy exercise in the 24 h prior to tests and from food intake in the 3 h prior to tests. For both laboratory visits participants were instructed to arrive at the same time of the day. The field study contained three different protocols. Protocol 1 (N = 11) required participants to complete three individual field-based tests to determine CP with protocol 2 (N = 10) requiring participants to complete individual efforts during single but randomised training sessions. Protocol 3 (N = 10) used the highest three PO values of all training files to determine CP.

7.2.3 Laboratory tests

Participants completed three TTE trials on the equipment described above. Work rates were equivalent to ~80%, ~100% and ~105% MAP, using a lowest to highest work rate order with a 30 min inter-trial recovery period. During rest periods fluid intake was permitted ad libitum. During each TTE trial, participants were cooled using an electric fan. Laboratory conditions were stable in a range of 18–22 C$^{\circ}$ with 45–55% humidity. After each TTE test termination participants were allowed to continue cycling at a recovery intensity of ~100 W for 5 minutes before dismounting the bicycle and resting passively. HR, PO and cadence were recorded continuously via the PowerTap, and expired gases were continuously sampled through the gas

analyser. Participants were blinded to TTE trial intensities and elapsed times. All cyclists reached their individual $\dot{V}O_2$max value (± 0.09 $L \cdot min^{-1}$), a post-test blood [lactate] of ≥ 8 mM and a HR within ± 5 beats of their maximal HR values established during the $\dot{V}O_2$peak test.

7.2.4 Field tests

Within the racing season and over the duration of 5 weeks, cyclists were required to record their training and racing activities using the PowerTap. Participants were instructed to avoid freewheeling during 'purposeful' efforts. Environmental conditions were not standardised and no instructions for the choice of road, gradient or cycling position were given.

Protocol 1 (N = 11);

CP and W′ were determined using 3 field-based tests. These comprised of a 12 min, followed by a 7 min and a final 3 min maximal effort using a recovery period of 30 min. Between maximal efforts cyclists either rested passively or continued cycling at a low, i.e. recovery intensity. Protocol 1 consequently resulted in three CP and three W′ values. Cyclists were instructed to perform these series of maximal efforts fully rested.

Protocol 2 (N = 10);

CP and W′ were determined using 3 field-based tests, which were performed individually during single but randomised training sessions. Participants in total had to complete three sets of required efforts of each 12 min, 7 min and 3 min maximal efforts over 9 individual training sessions. Cyclists were instructed to perform any of these maximal efforts fully rested. The completion of one set, i.e. a 3 min, a 7 min and a 12 min effort were used in the CP and W′ modelling process and protocol 2 consequently resulted in three CP and three W′ values.

Protocol 3 (N = 10);

As some of the intentional efforts were lower than 'non-intentional' efforts, protocol 3 used the highest three PO values (12, 7 and 3 minute durations) of all training and

racing files for the determination of CP and W′. Protocol 3 consequently resulted in three CP and three W′ values.

7.2.5 Calculation of critical power and W′

Training and racing sessions were recorded via a Garmin Edge 500 head unit (Garmin International, Kansas, USA). Participants were required to download sessions daily and to share the files with the PhD researcher using Dropbox (Dropbox, Inc., California, USA), a cloud storage file hosting service. Files were imported into WKO training software via the device agent into the 'athlete function' which was personalised for each participant. Using the performance management chart function, the specified efforts were exported into Microsoft Excel to model CP and W′ for protocols 1, 2 and 3. For all protocols, linear regression was used to determine CP and W′ using a work-time ($W = CPt + W'$; equation 1) and power-1/time ($P = W'(1/t) + CP$; equation 2) model. Values using equation 1 or 2 were consequently termed CP1/ W′1 and CP2/ W′2.

7.2.6 Statistical Analysis

Data were first examined using the Shapiro-Wilk normality test. Both, the validity and the reliability of field CP and W′ values were assessed within each protocol. To assess the variability of results from protocols 1–3, the within subject variation, expressed as a CoV and ICC were used. Repeated measures ANOVA was used to test for significant differences between repeated trials. Pearson product moment correlation analysis was used to provide an indication of the strength of any relationship between the laboratory values for CP and W′ and the different field values using equation 1 (CP1/W′1) and equation 2 (CP2/W′2). Pearson product moment correlation analysis was used to provide an indication of the strength of any relationship between the laboratory values for CP and W′ and the different field values. Agreement between the laboratory values and all mean field values of CP and W′ was assessed using LOA; [314,323]. Linear regression was used to calculate values for SEE for CP and W′ for each protocol and for laboratory values of CP and W′.

Differences of statistical significance between laboratory and mean field values of CP and W′ were tested using paired samples t-tests and accepted at $P < 0.05$. Results are reported as mean ± SD unless otherwise stated.

7.3 Results

Agreement between laboratory and field CP and W'. Laboratory CP values from all protocols were significantly correlated with field CP values ($P \leq 0.01$). Laboratory TTE trials durations were 667 ± 176 s, 256 ± 105 s, and 143 ± 44 s at 80%, 100% and 105% MAP respectively. The paired samples t-tests did not reveal any significant differences between laboratory and field CP values for all protocols ($P > 0.05$). Significant differences were demonstrated for protocol 2 and 3 values of W′ ($P < 0.05$). LoA and SEE values for CP and W′ values and protocols are presented in table 18 and 19 respectively; with figures 18 to 20 illustrating Bland-Altman plots of laboratory and mean field values of CP for all protocols.

Table 18. Mean values, mean differences, limits of agreement and standard error of CP

	Protocol 1	**Protocol 2**	**Protocol 3**
Mean Values CP1 field (W)	275 ± 35	271 ± 47	272 ± 44
Mean Difference CP1 lab (W)	0 ± 6	7 ± 17	-5 ± 14
95% CI	-3.69 - 3.69	-18.97 - 4.95	-14.59 - 5.40
LoA (W)	-11 - 11	-26 - 40	-23 - 32
SEE (%)	1.7	3.9	3
SEE (W)	4	12	9
	Protocol 1	**Protocol 2**	**Protocol 3**
Mean Values CP2 field (W)	277 ± 38	271 ± 44	276 ± 46
Mean Difference CP2 lab (W)	-2 ± 14	10.37 ± 21.80	-5 ± 20.00
95% CI	-11.19 - 7.74	-26.06 - 5.06	-19.31 - 9.31
LoA (W)	-26 - 29	-32 - 53	-34 - 44
SEE (%)	4.5	5.8	5.2
SEE (W)	11	17	14

Table 19. Mean values, mean differences, limits of agreement and standard error of W'

	Protocol 1	**Protocol 2**	**Protocol 3**
Mean Values W'1 field (kJ)	13 ± 3.67	17 ± 4.63	21 ± 5.48
Mean Difference W'1 lab (kJ)	0.47 ± 3.46	-4.40 ± 4.48*	8.04 ± 4.08*
95% CI	-1.86 - 2.80	0.94 - 7.86	5.11 - 10.96
LoA (W)	-7.26 - 6.32	-13.88 - 5.08	-16 - 0.03
SEE (%)	29	31.8	27.7
SEE (kJ)	3.06	3.61	2.73
	Protocol 1	**Protocol 2**	**Protocol 3**
Mean Values W'2 field (kJ)	12 ± 3.37	17 ± 4.7	20 ± 4.86
Mean Difference W'2 lab (kJ)	-0.14 ± 3.36	-4.62 ± 5.69*	7.79 ± 3.15*
95% CI	2.40- 2.12	0.54 - 8.69	5.53 - 10.04
LoA (kJ)	-6.44 - 6.72	-15.77 - 6.54	-14 - -1.6
SEE (%)	31.4	39.4	31.8
SEE (kJ)	3.08	4.03	2.83

Values are mean (± SD) - * = significantly difference from laboratory W'1/ W'2

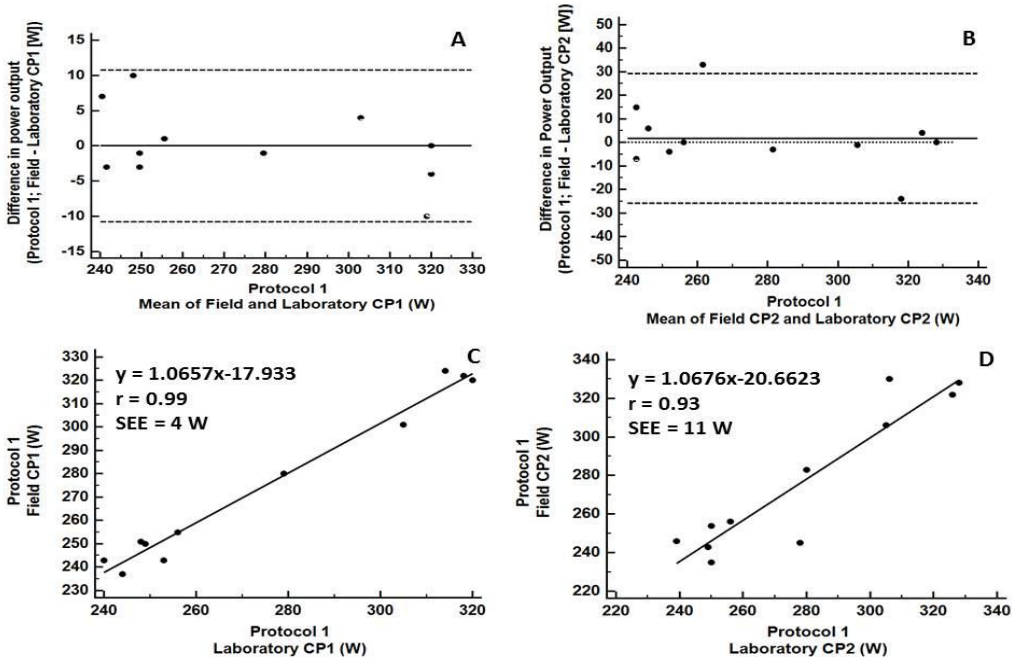

Figure 18. Protocol 1; Bland-Altman plots of the limits of agreement (panel A and B) and the relationship (panel C and E) between laboratory CP and field CP. In panel A and B the horizontal line represent the mean difference between laboratory CP and field CP, and the dashed line represents 95% LoA.

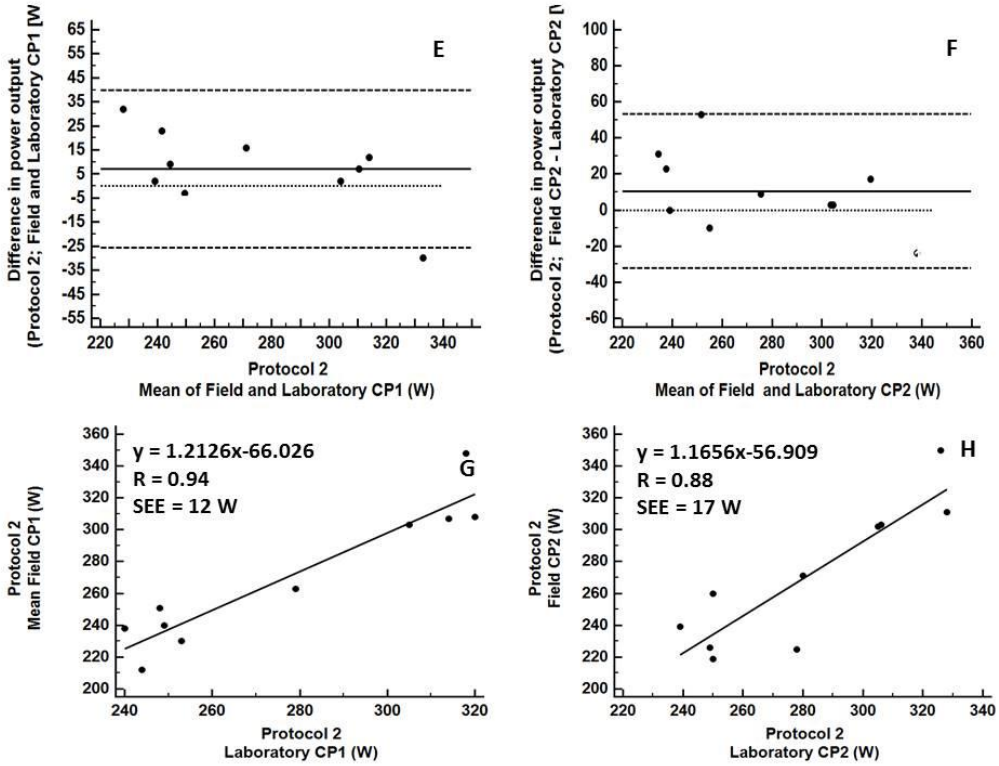

Figure 19. Protocol 2; Bland-Altman plots of the limits of agreement (panel E and F) and the relationship (panel G and H) between laboratory CP and field CP. In panel A and B the horizontal line represent the mean difference between laboratory CP and field CP, and the dashed line represents 95% LoA

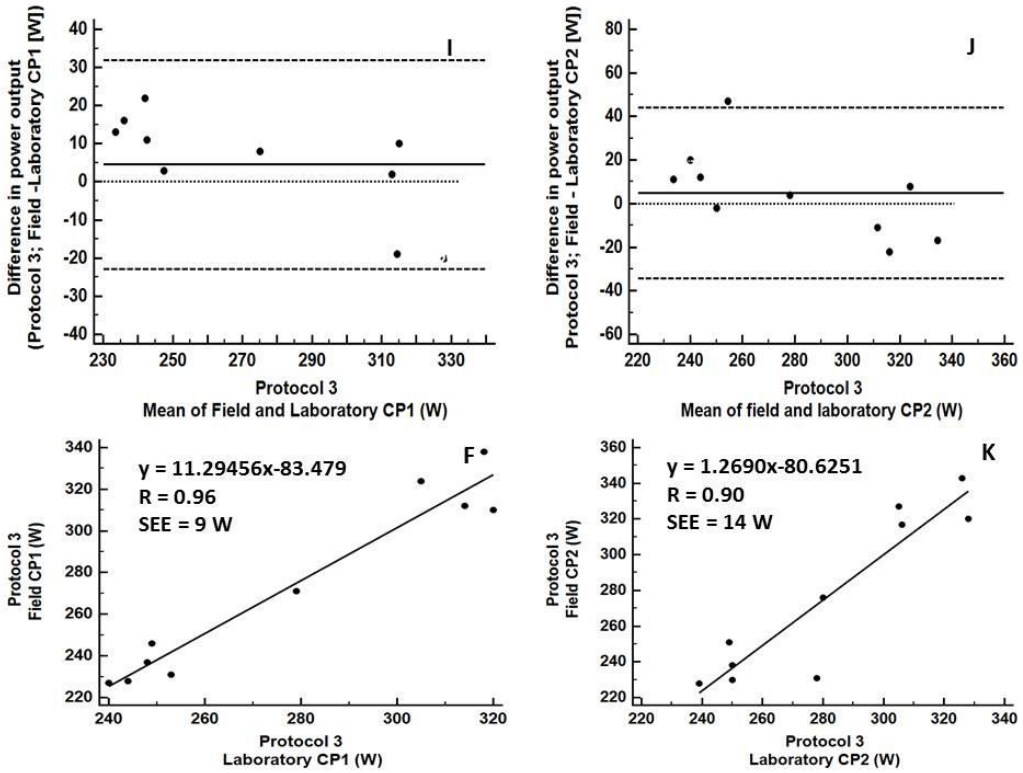

Figure 20. Protocol 3; Bland-Altman plots of the limits of agreement (panel I and J) and the relationship (panel F and K) between laboratory CP and field CP. In panel A and B the horizontal line represent the mean difference between laboratory CP and field CP, and the dashed line represents 95% LoA.

Reliability of protocols. For all protocols, repeated measures ANOVA identified no significant differences (i.e. bias) in CP between trials. (Protocol 1, CP1, $F(2, 9) = 0.74$, $P > 0.05$. CP2, $F(2, 9) = 1.64$, $P > 0.05$; Protocol 2, CP1, $F(2, 8) = 0.46$, $P > 0.05$. CP2, $F(2, 8) = 0.20$, $P > 0.05$; Protocol 3, CP1, $F(2, 8) = 0.32$, $P > 0.05$. CP2, $F(2, 8) = 3.33$, $P > 0.05$). CoV values for protocol 1 ranged between 1.6% and 2.5%, for protocol 2 the range was between 5% and 7% and for protocol 3 it was between 2% and 3.6%. ICC for all protocols for CP ranged between 0.96 and 0.99 (95% CI 0.90 – 0.99) (Table 20).

Table 20. Coefficient of Variation values (CoV), Intraclass Correlation Coefficient (ICC) values and 95% Confidence Intervals (CI) of all field CP results

	Protocol 1		**Protocol 2**		**Protocol 3**	
	CP1	**CP2**	**CP1**	**CP2**	**CP1**	**CP2**
CoV (%) Trials 1 vs 2	1.6	2.5	6	7	2	3.6
CoV (%) Trials 2 vs 3	1.9	2.1	5	5.9	2.5	3.3
ICC	0.99	0.99	0.97	0.96	0.98	0.99
95% CI	0.98 – 0.99	0.98 – 0.99	0.91 – 0.99	0.90 – 0.99	0.95 – 0.99	0.96 – 0.99

Table 21. Coefficient of Variation (CoV) values, Intraclass Correlation Coefficient (ICC) values and 95% Confidence Intervals(CI) of all field W' results

	Protocol 1		**Protocol 2**		**Protocol 3**	
	W′1	**W′2**	**W′1**	**W′2**	**W′1**	**W′2**
CoV (%) Trials 1 vs 2	43	47	41	48	10.3	15.6
CoV (%) Trials 2 vs 3	48	46	33	42	20	17.9
ICC	0.14	0.16	0.17	0.02	0.66	0.63
95% CI	-0.20 – 0.58	-0.82 – 0.81	-0.17 – 0.62	-0.29 – 0.44	0.31 – 0.89	0.23 – 0.89

7.4 Discussion

The main findings of this final study were a good level of agreement between laboratory and field determined values of CP for all protocols. Furthermore laboratory CP strongly correlated with field CP and CP field protocols generally had a very high test-retest reproducibility (Table 20). Table 18 demonstrates low mean, non-significant differences between field and laboratory CP values, acceptable LoA [314] and low SEE values. Gonzalez-Haro et al. [12] accepted their incremental velodrome field test as being valid with reported LoA of 130 W to -24 W and a random error of 77.1 W (13.9%). The study demonstrates LoA values which are considerably higher and SEE's that are considerably lower than those reported by Gonzalez-Haro et al. [12]. Study 2 reported similar mean differences of 2 ± 8 W with LoA between 11 W and 17 W and SEE values of 2.5% to those in this current study when comparing CP determined in the laboratory with CP determined from the track. It therefore can be suggested that the field protocols can be considered to be acceptable when determining CP in the field. In particular the field method used in protocol 1 provided the best agreement between laboratory and field CP values *(*Fig. 18, panel A and B). This is not surprising given an almost equal protocol in that CP determination was performed within a maximum testing duration of 2.5 hours, using the same order of maximal efforts and a 30 min recovery period between those efforts.

As hypothesised, low levels of agreement were found for field determined W′ values (Table 19). Moreover, protocol 2 and 3 identified significant differences between laboratory and field W′ with high prediction errors (≥ 29%) for all field values being evident. Previous research has questioned the reliability of W′[223,224]. Although likely to be multifactorial, differences for W′ under protocol 1 as suggested in Study 2 might be due to differences in standing or rolling start or change of cadence with a change in terrain [360]. Adding to these influences and due to having performed relevant efforts on different days, protocol 2 and 3 might contain more environmental (for example changes in weather condition or humidity), time and circadian rhythm influences, which can impact on anaerobic power [361]. By contrast Quod et al. [11], did not find any effect of location on W′ when comparing laboratory and race

determined values. Moreover W′ in the present study appears to exhibit a lower test re-test reproducibility (Table 21) which further compromises the validity of this parameter**.** Another issue to consider is that of ground level and gradient cycling. Padilla et al. [362] investigated differences between level and uphill TTs in professional cyclists. Mean PO was generally higher during uphill cycling and the authors suggested that higher PO can only be achieved during uphill cycling. Given that no instructions were provided on how to perform the maximal efforts nor where to perform them, an undulated terrain and possible changes in cycling position might have contributed to the differences in W′ due to an increased portion of type II fibre recruitment and the resultant higher PO values associated with greater blood lacatate concentrations [338,339].

A CoV of 10% has been suggested as the criterion value commonly used to define an acceptable level of test reliability [302]. To verify a reliable test Atkinson et al. further suggested an ICC > 0.8. Hopkins [301] later defined a lower 5 % CoV as the acceptable upper limit in sports science reliability studies. Given that the CoV values for CP observed using protocol 1 and protocol 3 (Table 20) were below the lower boundary as defined by Hopkins [301] the respective testing protocols can be deemed as being reliable. High interclass correlation coefficients (i.e. > 0.9; table 20) further demonstrate the repeatability of all protocols with a small bias ± random error, which are considerably lower than those reported by Gonzalez-Haro et al. [12]. Protocol 2 resulted in mean CoV values of 5.6% and 6.5% for CP1 and CP2 respectively, which according to Atkinson et al. [302] can also be deemed as acceptable. However, poorer LoA and higher associated prediction errors (Table 18) means that it is reasonable to question whether protocol 2 is as good as protocols 1 and 3 in its ability to accurately monitor the small changes in CP typically seen in trained athletes [21]. Furthermore, the hypothesis of W′ demonstrating a good level of reliability across repeated field trials has to be rejected. CoV and ICC values for all protocols were higher than the defined values by Atkinson et al. [302] or Hopkins [301] and it is questionable whether this parameter of the power-duration relationship is either valid or reliable in field testing.

The present study collected data over the duration of a 5-week period, towards the end of the racing season. Whilst assuming that CP would remain stable over this time period, small performance changes which may have affected results cannot be eliminated [363]. Cyclists were required to conduct a total of 18 purposeful efforts of 12, 7 and 3 minute durations during the period. Attempting to have a minimal impact on regular training, cyclists were not required to conduct the efforts in any order or at any specific time point. Interestingly, the results of this study are supportive of the previous work conducted in the outdoor velodrome, where the cyclist were performing within a consistent and more predictable environment [356]. Using a similar approach as in protocol 2, cyclists had to perform maximal efforts of fixed durations of 12, 7 and 3 min on separate days, and in a randomised order. A high agreement for CP but not for W′ was found when comparing laboratory and velodrome environments. However reported values for LoA of CP in the present study (Table 18, protocol 2) are not as high as in the velodrome study, which possibly demonstrates an influence of terrain on CP.

Under protocol 3, the single highest 3 min, 7 min and 12 min efforts from all of the training and racing files were extracted. Cyclists were not given instructions as to where to perform or how to perform these maximal efforts (i.e. seated or standing). Whilst laboratory trials were solely performed in a seated position results demonstrated a high level of agreement with field CP values (table 18) Using a similar approach to the current study, Quod et al. [11] extracted maximal efforts of fixed durations over 1 min, 4 min and 10 min to model CP and W′ from race data. In agreement with the study findings , Quod et al. [11] did not find a significant difference between laboratory and field CP results. However, it has to be noted that the lowest duration of 60 s used by Quod et al. does not comply with the requirements of CP determination as set by DiPramperio [208], i.e. attainment of $\dot{V}O_{2max}$. Furthermore the power profile testing, which included relevant CP and W′ efforts was not validated against conventional CP determination standards, as the researchers utilised an active recovery performed at 100 W for individual break periods (330 s, 480 s and 600 s between relevant maximal efforts of 60 s, 240 s and 600 s respectively). Interestingly, the data demonstrate a trend for higher mean field PO's under protocol 3, compared to those of protocol 1, 2 and in the laboratory. Training files revealed,

on a number of occasions, that cyclists produced higher mean PO of the set duration efforts under protocol 3, i.e. during efforts extracted from regular training and racing data. However, the higher mean PO values did not appear to greatly influenced values of CP, just W′. Deemed as being reliable (mean CoV 3.5%; ICC 0.99 and mean CoV 2.5%; ICC 0.97 for CP1 and CP2 respectively), the protocol used in protocol 3 could therefore provide a valid other method of assessing CP from 'normal' training efforts during which the cyclist does not have to provide pre-defined 'intentional' efforts.

7.5 Conclusion

Built on findings of Study 2 and Study 3 this the final empirical study demonstrates that CP can be determined in the field under 'controlled' (i.e. planned maximal efforts for a given protocol) and 'uncontrolled' (i.e. extraction of data from training and performances) situations. In particular protocol 1 resulted in a high level of agreement and low prediction errors, whilst providing a more ecologically valid testing environment when compared to laboratory testing. When applying protocol 2 and 3, lower LoA values and higher prediction errors have to be acknowledged but in spite of this, both protocol 2 and 3 have the advantage of being more easily integrated into the training schedule of riders. Each proposed CP field protocols can therefore be recommended to coaches and athletes as routine assessment. Future research studies are recommended to analyse training related changes in CP throughout the racing season, in particular applying field CP protocols 1 and 3, which provided the lowest CoV values.

The data sets collected in Study 2 and Study 4 allowed the research student to address an additional research question: that of 2 data points determined CP values using laboratory and field data from conventional and novel CP protocols. This analysis forms Study 5.

CHAPTER 8: EFFICACY OF CRITICAL POWER DETERMINATION FROM 2 VS 3 DATA POINTS

8.1 Introduction

The overarching aim of this research thesis was to investigate CP field protocols with a further focus on the athlete-friendliness of such protocols. Study 4 demonstrated that CP can be determined from training and racing files. However, some coaches might prefer to physically test their athletes as an important part of the coaching process. Performance tests with a lowest possible impact on the athletes training and racing schedule might further extend the practical utility of such tests. The collected CP results presented in this thesis therefore provide an ideal opportunity to re-analyse some data by comparing 2 data points vs 3 data points determined CP values.

Housh et al. [260] suggested that as few as two TTE trials are sufficient for the estimation of CP. Whilst this option might be very attractive to applied sports scientists and coaches, it contains a high risk of deriving inaccurate CP values if one the TTE trials is not fully exhaustive. Using combinations of two, three or four trials, Housh et al. [260] found a correlation value of r = 0.99 and low SEE between two TTE trials and four TTE trials determined values of CP. In line with their own recommendation [260], time to exhaustion between the two trials differed by ≥ 5 minutes and only time to exhaustion durations between 1–10 minutes were employed in this research. Using the linear CP models, it is clear that when employing two TTE trials a perfect linear relationship is the only possible outcome. However, as discussed under heading 2.7.4 a possible increased risk of reduced reliability is associated with this method as a 'bad test' will change the slope of the line for the work-time CP model or the y-intercept for the power-1/time CP model, having a potentially significant impact on the calculated CP. However, according to Hill [79], when working with trained individuals who are accustomed to exhaustive exercise, as few as two TTE trials can be sufficient. Nonetheless, as a 'trade-off' between accuracy of CP values and feasible testing, Hill [79] also suggested an optimal number of four to five trials. The notion of only using two TTE trials was taken up by Clingeleffer et al. [97] who performed a similar study in elite kayakers. The method included four maximal timed efforts. Combinations for CP determined from any two

such efforts were compared against CP obtained from all four efforts. A significant difference was only found for the combination of the shortest two efforts (90/240 s). All other combinations (90/600 s; 90/1200 s; 240/600 s; 240/1200 s and 600/1200 s) did not result in a significant difference though a magnitude based analysis might have revealed a different outcome. The second largest difference between two trial and four trial determined CP values was ~ 24 W, which in kayaking can result in different performance outcomes [255]. Furthermore a trend was evident in that CP values, that incorporated the longest or shortest maximal efforts, resulted in either lower or higher CP outcomes respectively. Clingeleffer [97] and Ginn [255] both argued that individuals might be more motivated by shorter and/or a lower number of trials to ensure more true maximal efforts. The only recent work investigating the number of exhaustive trials comes from Toubekis et al. [266] in swimming. Using young, trained swimmers the researchers measured the relationship between the LT and CV and physiological responses during interval training relative to CV. CV was modelled using two, three and four maximal efforts and the researchers consequently supported the notion of using fewer trials for CV determination as being more practical. Ferguson et al. [86,87] defined the number of trials with a systematically chosen acceptable SEE cut-off value. Individuals had to perform a minimum of three TTE trials but an additional trial, if SEE related CP values were > ± 3W. This arguably ensures greater accuracy, but decreases the feasibility of the testing for athletes who carefully plan their training and testing schedule. To date most research has employed a recovery period of 24 h between TTE trials. This approach potentially adds another day to the testing schedule, making the method even less practical.

The aim of this final data analysis was to compare laboratory and field determined 3 data points determined CP values with the 2 data points laboratory or field 2 data points CP values using the same data sets. In line with the recommendations made by with Housh et al. [260] data analysis only included exhaustive trials with a minimum of ≥ 5 minutes differences, (i.e. 12 min and 3 min or 80% and 105% MAP). A non-significant difference and high levels of agreement for CP values was hypothesised.

8.2 Methods

Data were examined using the Shapiro-Wilk normality test. A comparison between field 3 data points and field 2 data points CP values, laboratory 3 data points and laboratory 2 data points CP values and laboratory 3 data points and field 2 data points CP values was performed. Field results were further divided into velodrome and protocol 1 CP values. Agreement between the 3 point data values and 2 point data values of CP1 and CP2 was assessed using LOA [318,328]. Linear regression was used to calculate SEE values for each protocol. Differences in statistical significance between 3 data points and 2 data points CP values was tested using paired samples t-tests and accepted at $P < 0.05$. Results are reported as mean ± SD unless otherwise stated.

8.3 Results

CP values were normally distributed. Paired samples t-tests demonstrated no significant differences between CP1 and CP2 derived through the 3 data points and 2 data points method ($P > .05$) for all comparisons. Table 22 illustrates Study 2 mean differences, SEE and LoA for all 3 versus 2 data points comparisons between laboratory and field (i.e. velodrome) CP results. Table 23 illustrates Study 4 mean differences, SEE and LoA for all 3 versus 2 data points comparisons between laboratory and field (i.e. protocol 1) CP results.

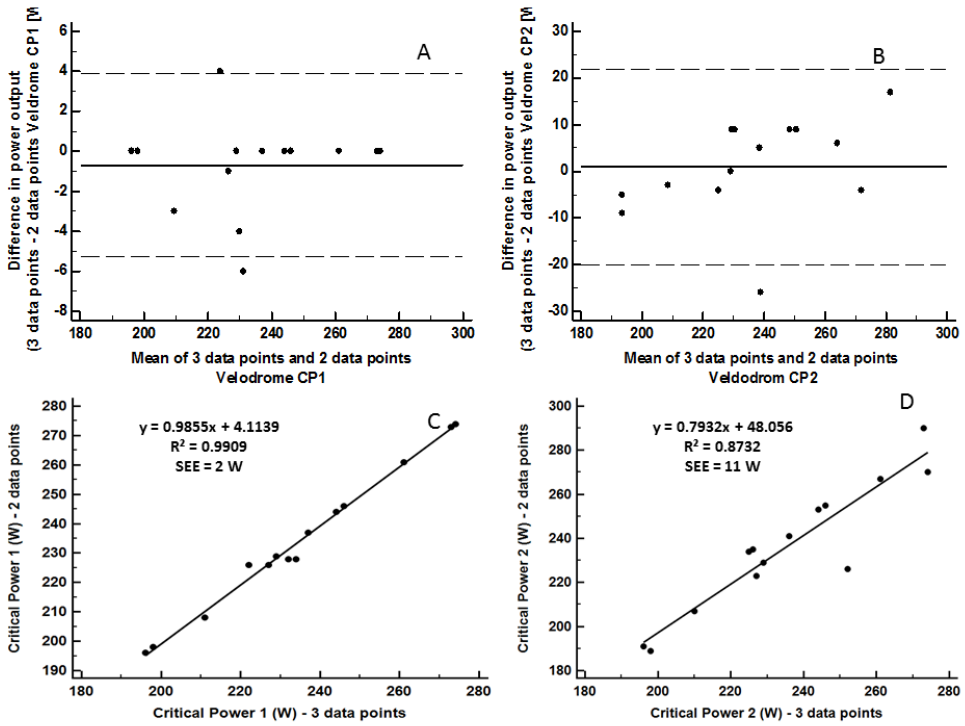

Figure 21. Bland-Altman plots of the limits of agreement (panel A and B) and the relationship (panel C and D) between velodrome CP using 3 or 2 data points. In panels A and B the solid horizontal line represents the mean difference between 3 data points and 2 data points CP1 and CP2 values.

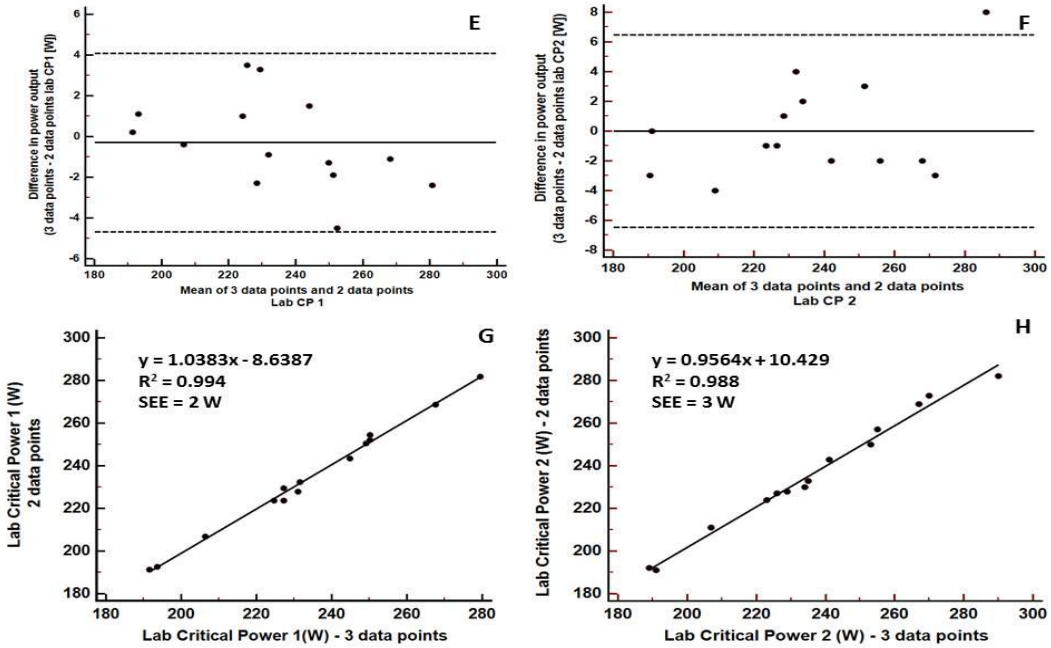

Figure 22. Bland-Altman plots of the limits of agreement (panel E and F) and the relationship (panel G and H) between laboratory CP using 3 or 2 data points. In panels E and F the horizontal solid line represents the mean difference between 3 data points and 2 data points CP1 and CP2 values.

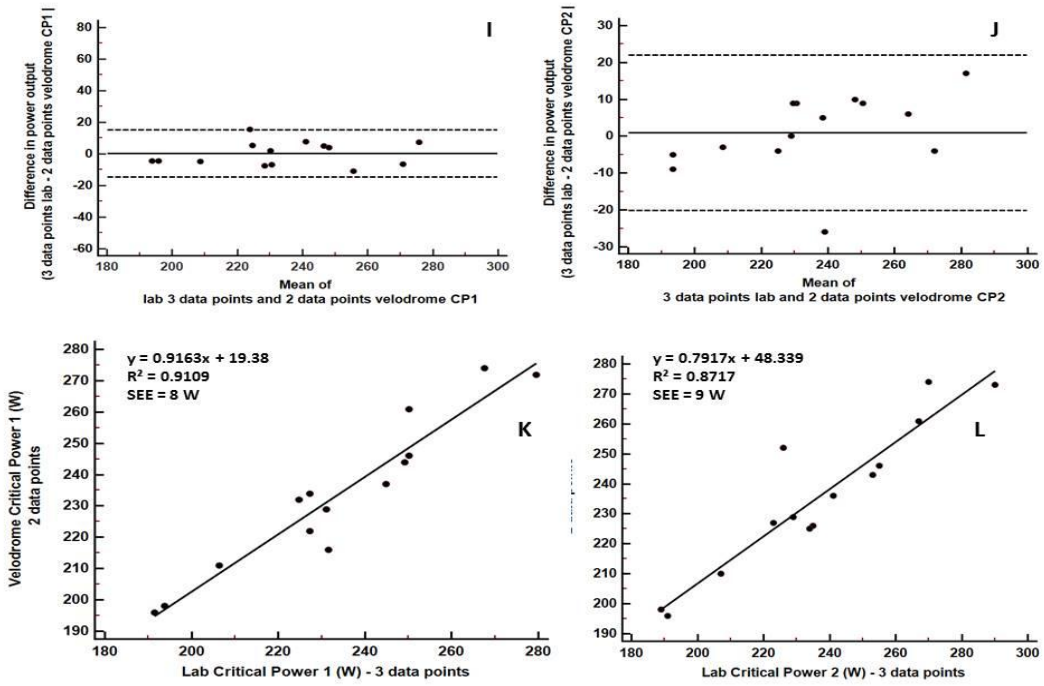

Figure 23: Bland-Altman plots of the limits of agreement (panel I and J) and the relationship (panel K and L) between laboratory CP values using 3 data points and velodrome CP using 2 data points. In panels I and J the horizontal solid line represents the mean difference between velodrome CP using 3 data points and field CP using 2 data points.

Table 22. 3 data points versus 2 data points determined CP values analysis (Study 2)

Study 2 results	**Mean Difference (W)**	**SEE (W)**	**SEE CI**	**SEE (%)**	**LoA (W)**
CP1 (Velodrome 3 data points vs. 2 data points)	1 ± 2	3	2 - 4	1.12	-5 to 4
CP2 (Velodrome 3 data points vs. 2 data points)	1 ± 11	11	8 –18	4.75	- 20 to 22
CP1 (Lab 3 data points vs. 2 data points)	0 ± 2	2	1 – 3	0.86	- 5 to 4
CP2 (Lab 3 data points vs. 2 data points)	0 ± 3	3	2 –5	1.3	- 7 to 7
CP1 (Lab 3 data points vs. Velodrome 2 data points)	0 ± 8	8	6 - 13	3.4	- 15 to 15
CP2 Lab 3 data points vs. Velodrome 2 data points)	1 ± 11	11	8 - 18	4.6	- 20 to 22

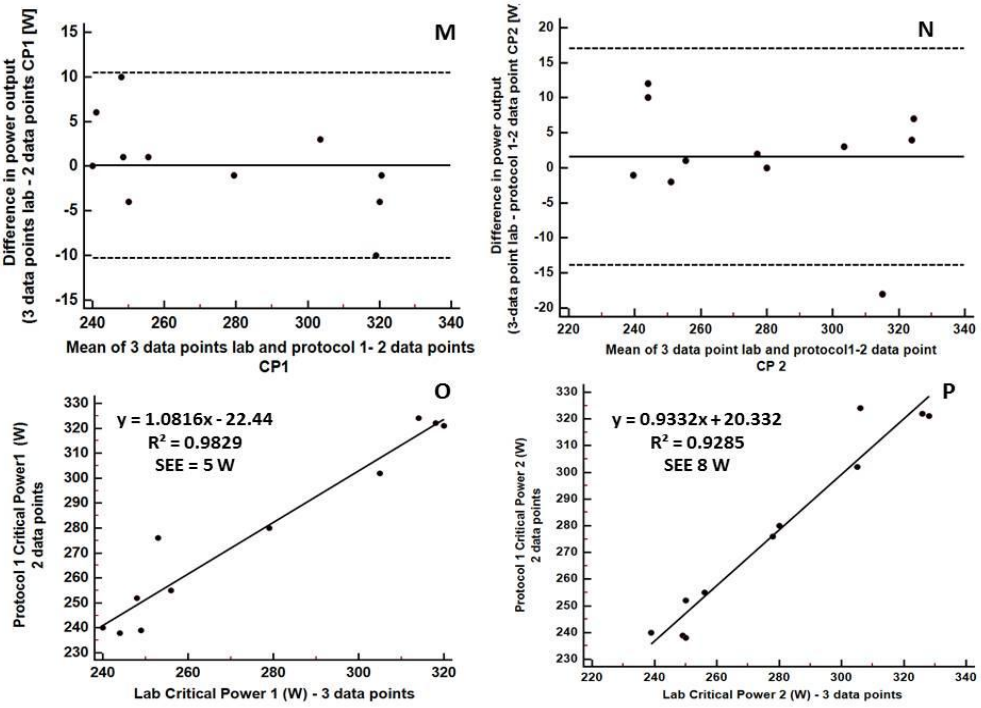

Figure 24: Bland-Altman plots of the limits of agreement (panel M and N) and the relationship (panel O and P) between protocol 1 CP using either 3 or 2 data points. In panels M and N the horizontal solid line represents the mean difference between protocol CP using 3 data points and field CP using 2 data points.

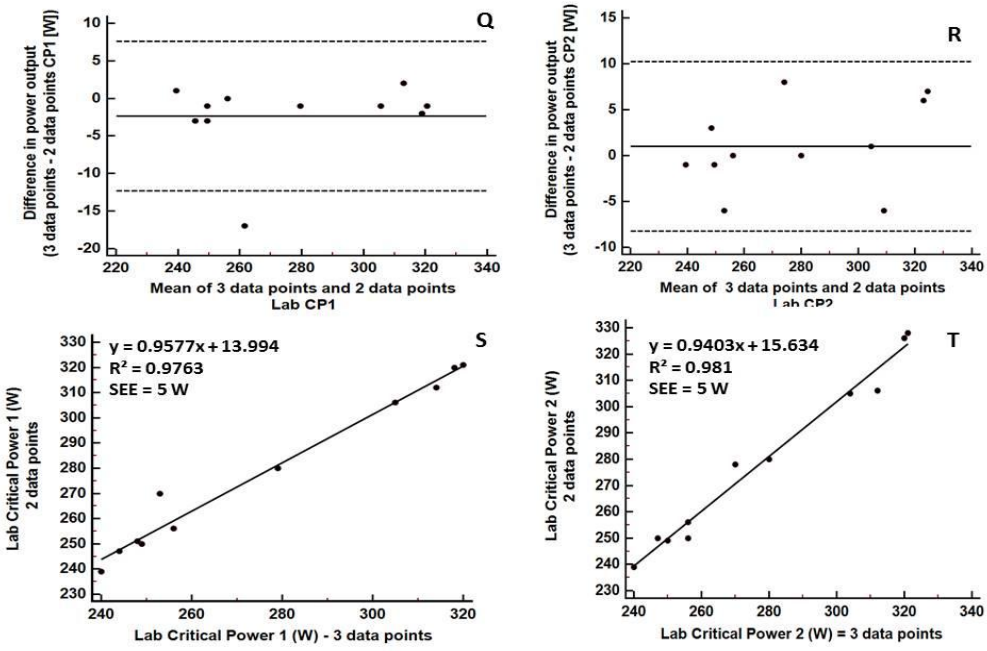

Figure 25. Bland-Altman plots of the limits of agreement (panel Q and R) and the relationship (panel S and T) between laboratory CP (W) using 3 data points and 2 data points determined CP. In panels Q and R the horizontal solid line represents the mean difference between 3 data points and 2 data points CP1 and CP2 values.

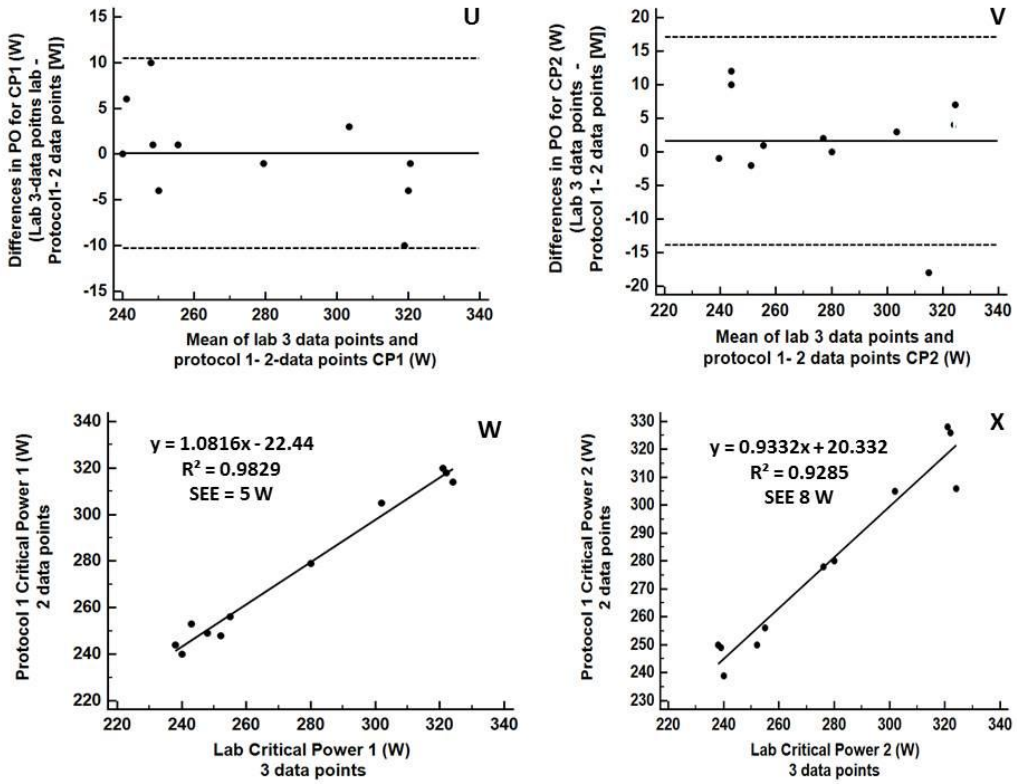

Figure 26. Bland-Altman plots of the limits of agreement (panel U and V) and the relationship (panel W and X) between laboratory CP using 3 data points and protocol 1 CP using 2 data points. In panels U and V the horizontal solid line represents the mean difference between laboratory CP using 3 data points and protocol 1 CP using 2 data points.

Table 23. 3 data points versus 2 data points analysis Study 4; protocol 1

Study 4 results	**Mean Difference (W)**	**SEE (W)**	**SEE CI**	**SEE (%)**	**LoA (W)**
CP1 (Prot. 1 - 3 data vs. 2 data points)	0 ± 1	1	1 - 2	0.4	-10 to 10
CP2 (Prot. 1 - 3 data vs. 2 data points)	-1 ± 10	10	7 –16	3.8	- 14 to 17
CP1 (Lab - 3 data points vs. 2 data points)	-2 ± 5	4	3 –7	1.7	- 12 to 8
CP2 (Lab - 3 data points vs. 2 data points)	1 ± 5	8	6 –13	2.7	- 8 to 10
CP1 (Lab 3 data points vs. 2 data points Prot.1)	0 ± 5	5	4- 9	2.1	- 10 to 10
CP2 (Lab 3 data points vs. 2 data points Prot.1)	2 ± 8	5	3 - 8	1.7	- 14 to 17

8.4 Discussion

The main findings of this study were low mean, non-significant differences between 3 data points CP values and 2 data points CP values. Therefore the hypothesis of non-significant differences between 2 data points and 3 data points CP values was accepted. Both, Study 2 and Study 4 were re-analysed comparing laboratory with laboratory CP values, field versus field CP values and laboratory versus field CP values. Low mean SEE values of 2.9% and 1.7% for the respective field CP values determined from 3 and 2 data points in Studies 2 and 4 and acceptable LoA [314] which are considerably higher than those established by Gonzalez-Haro et al. [12], demonstrate that CP field determination using only 2 data points can be considered valid. This furthermore applies to the comparison between laboratory determined 3 data points and 2 data points CP values. Mean SEE values of 1.1% (Study 2) and 2.2 % (Study 4) and acceptable LoA [314] illustrate the validity of this alternative, i.e. 2 data points laboratory CP determination method. Finally, when comparing 3 data points laboratory with 2 data points field CP values, an average SEE value of 4% was observed in Study 2 and 1.9% in Study 4.

Frequently used performance tests in cycling are TTs [22]. The typical error of measurement for recreational cyclists has been reported to be as high as 3% in these type of performance tests [8]. The SEE of 4% in Study 2 might raise some concerns when applying this protocol and a further investigation, ideally with a larger sample, is advised to gain further insight into the random error associated with the determination of field CP using 2 data points. As the systematic error associated with the PowerTap has been reported to be ~0.9% [46], the observed 4% therefore might contain a random error of ~ 3.1% caused by the cyclist [8]. Cyclists in Study 2 performed a single maximal effort of each fixed duration trial. Contrary, Study 4 utilised the mean of 3 relevant efforts, i.e. the mean of three 3 min, three 7 min and three 12 min max efforts. Study 4 consequently contains less error, i.e. a lower variability associated with a single maximal performance effort of either TT trial. A ‘bad’ test has less impact on outcomes if averaging the results of repeated trials. However the reported 4% SEE is below that reported by Gonzalez-Haro et al. [12].

Non-significant but greater differences between 2 data points and 3 data points CP values in the present analysis were expected. This in particular applies to the comparison between laboratory and field CP values as the data points originated from different data sets, i.e. laboratory and field exhaustive trials. Laboratory 2 data points CP results did not demonstrate a better study control when compared to respective field CP values. As discussed in the literature review the ‘’inevitably high’’ risk of an incorrect CP value by incorporating a “bad” test when using only 2 data points can therefore be refuted when using sport-specific trained athletes. Supported by the findings of Toubekeis et al. [103,266], coaches consequently may now have an additional 2 data points field and laboratory CP performance monitoring tool, containing a 3 min and a 12 min maximal effort.

8.5 Conclusion

Using data collected in Study 2 and Study 4 this final investigation demonstrates the feasibility of CP determination using 2 data points. Coaches wishing to physically test their athletes by integration of a regular testing schedule can now utilise this additional protocol. This might increase the attractiveness of CP further, enhancing its practical application.

CHAPTER 9: GENERAL DISCUSSION

9.1 Summary of main research findings

The work above describes five studies addressing two convergent research questions: one relating to the feasibility of field-based determination of CP, the other relating to the feasibility of reducing the total time required to determine CP via traditional methods. The principle methodology was to compare a modified method with either the conventional one (Studies 1, 2 and 3), or with a newly developed method (Studies 4 and 5).

Study 1 investigated the validity of the recently developed 3-min all-out cycling CP test. Theoretically the 3-min all-out test allows CP determination within a single test, and consequently it appears attractive to coaches wishing to continuously monitor their athletes. The 'conventional' single test however does require an incremental maximal exhaustive test. This was until recently either overlooked or ignored [57,90,102,312,324], but has now been investigated using an alternative novel 3-min all-out method [321,364] which together with a maximal incremental test can be performed within one day [365]. Analysis of the 3-min all-out CP test in Study 1 identified issues when using a different ergometer and/or a different mode (i.e. isokinetic, which is cadence independent) [23,304]. As the particular setting of the ergometer in the original research (i.e. linear mode, which is cadence dependent) is specific [312], the study investigated whether or not the proposed all-out CP test would provide similar results when using the SRM ergometer in isokinetic mode. The isokinetic mode provided CP values that were consistently higher (mean + 37 ± 19 W) than those derived from a conventional CP determination protocol. This suggested that differences in CP values were either caused by the ergometer mode/setting or by issues related to the specific subject group of trained cyclists utilised in Study 1, such as generally higher cadences and/or fitness levels, when compared to the original investigation. However there is now evidence that the test provides valid results in running [366]. This agreement has yet to be replicated in cycling and a most recent study by Dekerle et al. [307] lends some further support to the findings presented in this thesis.

Study 2 investigated whether it is possible to determine CP from field data. The field testing methodology of the study was built on previous findings in running [74,268,269] and swimming [196,367,368] where CV/CS had been successfully implemented in a field testing format. Both exercise modes (swimming and running) employed a fixed distance testing method. Findings in cycling by Hiyane et al. [267] and Quod et al. [11] lent further support to the development of this novel field testing protocol. Hiyane et al. [267] used a set of fixed distances which cyclists had to cover in the fastest possible time. In contrast, Quod et al. [11] employed a fixed time method of 1 min, 4 min and 10 min and recorded mean PO over these set durations. Both cycling studies generated exhaustive times between 1 and 10 minutes which as noted under heading 2.7.2 does not fully comply with some of the CP determination recommendations [182,208].

Therefore the findings of these studies are only conditionally transferable and yet have to be replicated with the recommended exhaustive times by Poole et al [164] and Hill [79,182] as the resultant CP values are more likely inflated, which may have distorted the physiological meaning of CP as prolonged sustainable exercise intensity. In Study 2 the use of a road racing bicycle equipped with a mobile power meter allowed data to be collected whilst cycling on an outdoor track. Field test CP1 values were consistently close to laboratory values. In 9 of 14 participants the difference between laboratory and field CP was $\leq$ 5 W, for 4 participants the difference was between 6 W and 8 W and for one participant it was 11 W. A similar picture emerged for CP2. In 6 of 14 participants the difference between laboratory and field CP was $\leq$ 5 W, for 4 participants the difference was between 6 W and 8 W and for 4 participants it was $\leq$ 13 W. However, whilst providing strong evidence for valid CP field testing, the method still required multi-day trials and therefore does not fulfil the requirement of being athlete-friendly. Moreover, athletes and coaches in favour of this approach still have to rely on the availability of a cycling track. Whilst environmental conditions of the outdoor track are not as controlled as in the laboratory, issues such as an undulated terrain or possible traffic were not present. In conclusion the study outcomes indicated that the novel field test protocol was justified for use on an outdoor track.

Traditionally, 24 h recovery periods between exhaustive exercise trials are used in the determination of CP. Study 3 asked whether a 3 h and a 30 min inter-trial recovery period could be used to accurately determine CP, thus increasing the usability of CP. In cycling only one study directly compared conventionally determined CP values with those determined using an alternative recovery period of 3 h [263]. The study identified a high agreement for both, CP and W' values. Whilst high levels of agreement between CP determined using a 24 h and a 30 min intra-exhaustive trial period were established in Study 3, low levels of agreement for respective W' values were identified. Mean differences for CP1/CP2 between the 24 h and the 30 min protocol were -2 ± 9 W (LoA; -9–6 W) and -2 ± 12 W (LoA; -12–7 W) respectively. Mean differences for W' between the 24 h and the 30 min protocol were 0.1 ± 3.5 kJ for W'1 (LoA; - 2.6–2.9 kJ) and 0.2 ± 3.9 kJ for W'2 (LoA; -2.8–3.2 kJ). Being significant these differences either suggest an incomplete restoration of W' or, as insinuated under 2.5 confirm an inherent difficulty in accurately and reliably determining this parameter [197]. The results of this study are particularly important as they are the first to show that a 30 min recovery period is sufficiently long enough to accurately determine CP.

Study 4 questioned whether it is possible to accurately determine CP from road testing data, i.e. from an environment that provides very limited test control. Participants were equipped with a power meter and PO values were extracted from training and racing files. The study design was based on previous findings of this thesis by adopting fixed exhaustive trial durations of 12, 7 and 3 min and by using a 30 min recovery period between trials. Furthermore, based upon the only research study to have determined CP from road racing data [11], two additional protocols were designed to question whether CP determination has to follow a specific testing protocol or whether it is possible to determine CP from single intentional or non-intentional maximal efforts over the validated durations of 12, 7 and 3 min. To investigate reliability, each protocol was performed three times, producing 3 CP values. Protocol 1 used the above described durations and a 30 min inter-trial recovery period. Protocol 2 utilised randomised but individual maximal efforts of the set durations and modelled CP once a set of efforts (i.e. one 12 min, one 7 min and one 3 min) was completed. Protocol 3 extracted highest intentional and unintentional maximal efforts from training and racing data. A good level of agreement between

laboratory CP1 and field CP1 values was found for all protocols. Protocol 1 provided LoA of 11–10 W, protocol 2 of -26–40 W and protocol 3 of -23–32 W. For all protocols a good level of agreement was also found between laboratory CP2 and field CP2. Protocol 1 provided LoA of -26–29 W, protocol 2 of -32–53 W and protocol 3 of -34–44 W. Low SEE values (≤ 3.0%) were only identified for protocol 1 and protocol 3 (CP1). Expressed in watts this led to a prediction error of 4 W and of 9 W for the laboratory CP value by the field CP value for protocol 1 and 3 respectively. Both protocols also provided CoV values below 3% with protocol 2 resulting in a CoV value of 6 %. When applying protocol 2 and 3, lower LoA values and higher prediction errors have to be acknowledged but in spite of this, both protocol 2 and 3 have the advantage of being more easily integrated into the 'training schedule' of riders. Each proposed CP field testing protocols can therefore be recommended to coaches and athletes as routine assessment. Study 4 also demonstrated that CP can be determined via a simple data extraction method using training and racing files. This, considering appropriate maximal efforts are included in regular training and of course competition, allows for 'unlimited' tracking opportunities of the performance parameter of CP. Being deemed as reliable (< 5%) the potential of this extraction method may also exceed the usability and practicability of other commonly measured index of endurance fitness.

Study 5 utilised some of the collected data and assessed whether it is possible to accurately determine CP using only 2 data points. Combinations of laboratory CP, field CP and laboratory-field CP values using 3 and 2 data points from studies 2 and 4 were analysed. The validity of using only 2 data points was considered with mean SEE values of 2.9% (Study 2) and 1.7% (Study 4) and acceptable LoA [314] for field CP values determined from 3 and 2 data points. Similarly low mean SEE values (1.1% in Study 2; 2.2 % in Study 4) and acceptable LoA [314] indicate the validity of 2 data points determined laboratory CP values. Low mean SEE (4% Study 2; 1.9% Study 4) and acceptable levels of agreement [314] were also identified for 2 data points field CP values when compared to 3 data points determined laboratory CP. CP determination protocols using only 2 TTs provide the greatest ease of field determination for athletes. The mathematical risk of an incorrect CP value using only 2 data points can consequently be refuted when using sport-specific trained athletes.

It is evident that the data and conclusions presented above rely on the statistics of agreement, and that questions regarding the validity of this statistic could theoretically undermine both. LoA has been the focus of some debate in the sport and exercise sciences. Hopkins [315] urges discarding the Bland Altman analysis and suggests that the plots can result in an incorrect conclusion about the validity of a measure. Hopkins advocates the use of the regression approach and judges it as being superior to the Bland-Altman analysis as the LoA plot tends to incorrectly suggest the presence of systematic bias in the relationship between two measures. When fitting a regression line to the plot, with a slope gradient significantly different from zero, Hopkins argues the existence of proportional bias. According to Hopkins [315] linear regression does not lead to incorrect predictions of the established by a new measurement method whilst providing the magnitude of prediction error as SEE value. Others such as Atkinson [278] and Batterham [369] contributed to the debate by acknowledging inherent 'errors' of the Bland-Altman analysis whilst still arguing for its place in research. Furthermore, Currell and Jeukendrup [305] pointed out that LoA are affected in the presence of heteroscedasticity, i.e. the measurement error becomes larger as the magnitude of the test score increases. A further review on this subject is beyond the scope of this research thesis. However, it is apparent that to date a uniform answer to the conundrum of the superiority of either method cannot be provided.

Validity in aforementioned results was consequently indicated using both, LoA and SEE values. Whilst Bland and Altman [314] stipulated clear recommendations about the acceptability of a novel measurement method (i.e. if all data points fall within established LoA), no such recommendations exists for values of SEE. Researchers can use a magnitude based analysis approach, i.e. a smallest meaningful difference as suggested by Butterham and Hopkins [370], to take an informed decision when comparing the criterion with the practical measure. For example Paton and Hopkins [299] identified a 1.7% (CL 1.2 – 2.6%) performance improvement significantly affecting the athlete's chances of winning a road TT in elite cycling. Paton and Hopkins when investigating seasonal performance chances in competitive cyclists also identified a 6.1% and a further 2.2% (CL ± 2.2%) between base and pre-competitive and between pre-competitive and competitive season. The discrepancy between those values and mean presented study SEE values are deemed to be

acceptable, considering that lower SEE values would have been expected from elite cyclists [300,371].

The reliability of a test is indicated by CoV and ICC values. Accepted as reliable, presented CoV and ICC values for Study 1 and 4 (protocol 1 and 3) are below the recommended 5% [301]. As it is the case with validity, it is not unreasonable to expect lower CoV values from a group of elite cyclists [300,371]. Raw mean typical error values in protocol 1 and 3 produced 5 W and 8 W respectively. A CP value of 300 W for example would contain a typical error of ~ 1.15% (protocol 1) and 2.7% (protocol 2). Furthermore, presented ICC values in study 4 were above the recommended benchmark (i.e. > 0.8) [304] and can be categorised as highly reliable (i.e. > 0.9). Findings in Study 4 further support the validity of the novel CP determination method by repeatedly producing accurate results. As discussed in the literature review, a test whilst being reliable is not valid if it repeatedly produces inaccurate results.

9.1.2 Relevant research

Sports science research should influence real-world activity and thereby enhance human performance [372]. Only research that leads to applied practice can ultimately enhance performance. Bishop [280] called for a new framework for researchers to design studies with a focus on how results can directly improve athletic performance, i.e. relevant studies. In his publication on an applied research model for the sport sciences, Bishop [280] stated a general consensus of poor translations between sports science research and practice. Moreover science according to Bishop [280] is too often restrained to the laboratory. For research outcomes to become effective two criteria have to be fulfilled. The first criteria requires researchers to effectively communicate with practitioners about research findings. The second criteria requires researchers to provide evidence that demonstrates the feasibility and effectiveness of the research findings in practice. Criticisms of the apparent low application of research findings into a real-world sporting context and of little impact on elite performance were raised by the House of Lords in 2012 [3]. Atkinson et al. [373] in response to these criticisms stipulated good practice criteria which should benchmark relevant research. It is identified below how each of these criteria were observed in the above thesis.

Relevant population

Atkinson et al. [373] emphasised the importance of selecting a relevant population. Research should clearly define this population as Olympic athletes or recreational athletes. This criterion of good practice was followed throughout all the above studies and each corresponding method section states 'recreational' cyclists as the relevant population.

Relevance of measures

Atkinson et al. [373] asked for performance outcomes to be directly relevant to the particular study population and for performance predicting correlational studies to be used only if it is impossible or difficult to directly measure or simulate the performance. This criterion of good practice was also followed throughout all studies. The significance of CP for cyclists is more discussed under heading 2.2 and its value as performance predictor has been demonstrated [63,66]. In fulfilment of good practice, CP was directly determined within each study and protocol.

Interpretation of statistical significance

Another important issue highlighted by Atkinson et al. [373] is that of the role of statistical significance and non-significance as sole evidence of the presence or absence of a meaningful effect. This was particularly important for the data analysis in this thesis, as a non-significant difference does not indicate agreement in method comparison studies. In short, a non-significant difference does not express the validity of novel testing protocols, nor does it express reliability. For example Bergstrom et al. [321] investigated the 3-min all-out test using a Monark cycling ergometer. The validity of the alternative testing protocol was stated in the absence of a significant difference between the 'conventional' 3-min all-out and the proposed alternative 3-min all-out test protocol. In contrast to this, all studies in this research thesis, whilst reporting significant and non-significant differences indicated validity through LoA and SEE values, which are meaningful in method comparison research [301,314] . SEE values, expressed as a percentage or in raw units, can include confidence limits, which further define the likely range of the true magnitude of the prediction error. This is important in making meaningful statements about the precision of a testing protocol. A brief discussion on the differences between LoA and SEE is

presented above. In a wider context response of this good practice point, reliability was indicated as CoV in all studies.

Selection of variables

Atkinson et al. [373] additionally advises the use only of those physiological variables which have a clear and strong relationship with performance outcomes. All studies presented in this thesis are based on PO or in more detail on CP, both of which have been demonstrated to have a strong relationship with cycling performance [55,63,66,309,374–376]. Physiological variables, such as HR, blood [La] and $\dot{V}O_{2max}$ were analysed to support test results only where appropriate.

Participants

Graphical presentation of individual data (including for example gender and training level), according to Atkinson et al. [373] should be considered and included as a possible moderator in the data analysis. Study 1 and 2 included 2 female participants. As the presented research in this thesis is concerned with the validity and reliability of an athlete-friendly CP field testing protocol, gender was initially not considered. Furthermore age was only considered as a risk factor and an age limit of 50 yrs applied to all studies. However, some journals appear to prefer a homogenous gender and/or age population as study group (the rejection of Studies 1 and 2 was partially based on the inclusion of the female cyclists or a 'too high' average age). Consequent studies only recruited male cyclists but kept the upper limit of 50 yrs throughout the thesis. To minimise variance in results, Study 4 selected only male cyclists with a defined volume of either > 250-300 km and/or > 10 hours of training peer week.

Control of dietary variables

A final criterion stated by Atkinson et al. [373] requires researchers to clearly rationalise strict dietary controls if these are to be implemented in an investigation. All MAP tests presented in this thesis followed the testing guidelines as set by the American College of Sports Medicine (ACSM) for $\dot{V}O_{2max}$ testing [377]. For all other testing procedures (i.e. TTE trials and TTs) recommendations were provided for cyclists to perform these tests fully rested and hydrated, a condition under which cyclists in the real world generally race.

The relevant research problems addressed in this thesis are concerned with enhancing a very limited body of research evidence attempting to translate a standardised laboratory test into the field [11,12] and to validate this field testing protocol. The presented research findings have potential applications well beyond the boundaries of this thesis. For example, the proposed testing or data extraction protocol can be used within any sporting context which uses power meters. In the future this might include any sporting equipment, capable of measuring PO, as the principle of strain gauges can potentially be built into any material, which is solid enough to incorporate such force measurement devices. Whilst hand-cycling has already adopted power meters, when for example investigating wheelchair marathon performances, CV as a suitable performance parameter is influenced by confounding issues such as wind and a more undulated terrain. As previously mentioned PO is independent of these external conditions and potentially offers a more appropriate testing variable when designing field testing protocols [294]. Developing sporting equipment which is capable of measuring force produced by the athlete, such as paddles for kayakers, rowers and canoeists in the future might provide an opportunity to apply the presented thesis research finding of CP field testing into a wider sporting context. This could inform coaches and athletes appropriately and disseminate research findings whilst potentially further enhancing the application of the results into real-world sport.

However, the scientific body of literature on CP cycling field testing yet has to develop. Quod et al. [11] provided some first support for CP field testing. The testing method employed by these researchers is similar to protocol 3 in Study 4 in that best efforts over set durations were extracted from racing data and in that CP was determined using the power-1/time model. Furthermore Quod et al. [11] also used three data points for determination of CP. To date no further scientific support has been provided to demonstrate the validity of CP field testing. Furthermore no research has been published which investigates the reliability of CP field testing. Whilst Study 4 provides some support for the reliability and reproducibility of CP field testing in particular when using protocol 1 and protocol 3, the duration of the study over a 5-week period was not sufficiently long to demonstrate that CP determined in the field is sensitive to changes in performance capacity. This has only been demonstrated for laboratory CP [62,68,85,155]. The presented research findings further provide strong

support that field CP can be tested in a more athlete-friendly way in form of a single testing session, not requiring a MAP test. This contrasts with conventional CP laboratory testing which currently still requires a MAP test.

The findings of the presented thesis are relevant for future applied research and real-world testing of athletes. In summary, findings and derived recommendations are:

I. CP can be determined in the field within a single session of maximal 1.5 hrs duration, allowing a more regular monitoring of an athlete's performance capacity.
II. CP can be determined from track and road data using maximal efforts of 12 min, 7 min and 3 min with a 30 min recovery period in between maximal efforts.
III. Freewheeling during maximal efforts has to be avoided as power drops to zero. This might impact on mean TT effort PO values and consequently can distort CP results.
IV. It is not advisable to analyse the resultant value of W' and to draw any conclusions from this value using the suggested CP field testing method.
V. It is possible to determine field CP from maximal efforts of 3 min and 12 min.
VI. It is possible to determine field CP using a data extraction method of defined mean maximal efforts (i.e. 3 min, 7 min and 12 min) from training and racing files.

Using the suggested field CP protocol coaches and athletes do not have to rely as much on:

I. The availability of a sports science laboratory
II. The expertise of a sport scientist
III. Or expose themselves/their athletes as much to extensive, time consuming and training interruptive testing protocols.

It is the opinion of the author that the findings presented in the above research thesis have the potential to impact directly on real-world cycling training and outcomes. The findings might in time help develop similar protocols in sports in which power output over time generally, and critical power specifically, are key factors.

9.2 General research limitations

The main purpose of this thesis was to research the viability of an athlete-friendly CP field testing protocol. This required the accurate and valid determination of CP for a robust investigation of this research topic. One prominent limitation of the research outcomes presented in this thesis is that it only applies for the protocols and methods used. As outlined under heading 2.6, the choice of CP model and the duration of exhaustive times mathematically provide specific values of CP and therefore are only applicable to those models and durations. Aforementioned results are therefore limited to the work-time and power-1/time CP model using exhaustive laboratory and field times between 2 – 15 min.

Another possible limitation was the characteristics of the participation group. Whilst every attempt was made to recruit experienced cyclists, none of them was performing at elite level and individual differences in fitness levels are evident within each study. This potentially resulted in a greater random error caused by the participating cyclists rather than the test. As cyclists had to self-select a pacing strategy, they were instructed to perform one familiarisation trial for all field efforts [178]. Both relevant field studies had to rely on the participants' compliance to perform these familiarisation trials. If not adhered to, this likely resulted in less reliable CP values, i.e. values for CoV and ICC in Study 4 could have been higher and lower respectively than those presented in the result section (Table 20). Not having supervised familiarisation trials consequently presents an additional limitation to the presented research. To minimise possible learning effects it might have been beneficial to have included such familiarisation trial for each required laboratory TTE effort [65]. Practical considerations such as number of visits to the laboratory as well as recommendations of Hopkins et al. [178] however justified the decision of

performing only one respective TTE effort for the determination of laboratory CP. Finally Study 2 would have benefited from at least one repeated CP field test to establish track CoV values. However, this would have extended the overall duration of the study substantially and might have resulted in a lower participation number, since laboratory and field CP was determined each over a duration of 3 days. Similarly Study 3 might have benefited from repeated trials but practical considerations led to testing each respective CP not repeatedly.

One of the principal aims of the current thesis was to investigate whether it is possible to test CP using a more athlete-friendly method. In order to achieve such aim, participants in Study 3 undertook 13 separate exhaustive laboratory exercise trials. This required highly motivated individuals to give a maximal effort for each exhaustive test in order to obtain accurate CP values. The minimum number of individual tests in the presented studies was 9 with a maximum of 28 individual tests. It was not possible nor was it practical (i.e. athlete-friendly) to perform additional trials even when values of SEE resulted in ≥ 5 W [86,87], thereby the absolute accuracy of CP might be limited.

Studies 2 and 4 compared laboratory constant work-rate efforts (TTE) with field set duration efforts (TT). Whilst both efforts are of a maximal exhaustive nature, the resultant power profiles differ. Cyclists during constant work-rate efforts have to perform 'against' a constant resistance, that is PO is fixed and stable. Using an electronically controlled ergometer, a change in cadence does not lead to a change in PO, as this is held relatively stable by the ergometer. During field testing, cyclists at the start of a fixed duration effort typically produced high instantaneous PO values which as a function of elapsed time decreased. The level of decrease in PO is dependent on the level of accumulated fatigue. The difference in resultant PO curves between laboratory and field tests present a practical limitation in this study. Ideally same type PO efforts should have been compared. However whilst TTE trials are recognised in CP laboratory testing, these cannot be identically replicated in the field.

Study 4 was generally limited by the number of available PowerTaps. Such resource limitations are not unusual in sports science research, especially at doctoral level. Running the study multiple times would have increased the probability of introducing seasonal performance changes [363]. Whilst attempting to recruit athletes who have their road bicycles equipped with a PowerTap, this was only successful on two occasions.

9.3 Future directions

The findings presented in this research thesis can be used as a basis to influence the design of CP field research studies and to inspire future research.

A main focus on translations of relevant performance laboratory tests into the field can be recommended. Whilst providing a 'rough' template, individual study designs can be used in future studies to investigate such translations and to enrich the scientific body of literature with validated field performance tests.

A focus on field CP changes throughout a competitive cycling season and as a function of training status will also be required. With an increasing number of cyclists using power meters the potential for large scale data collections and analysis now exists. This would provide important information about the magnitude of change in CP throughout the year, and could moreover provide a range of reference values for the categorisation of cyclists with respect to their CP values.

As competitive road cycling comprises events ranging a few kilometres up to thousands of kilometres, typical CP values of athletes for each of these races will vary. Whilst playing a more significant role in events lasting 2 to ~ 30 min., CP is still relevant for longer duration events [55]. According to Joyner and Coyle [49] the outcome of all Olympic endurance events is decided at intensities above 85% $\dot{V}O_{2max}$. This highlights the role of the very heavy and severe domain for training, both being encompassed by CP. The categorization of race specific CP values could shed some further light into the significance of CP within explicit events. A study could be suggested which collects real-world CP racing data and to correlate resultant CP

values with the performance of these events. For example Smith et al. [63] found a strong relation between CP and TT performances of 17 km and 40 km in competitive cyclists. However the study related CP with TT durations and studies which relate CP with race power can be suggested.

A further research area can be suggested for the use of field CP in talent identification in adults as well as in underage athletes. CP has previously been investigated in children and adolescents [102,104,254,378]. This could provide further insight in athletic maturation using CP.

Whilst CP has been well researched in road cycling within the constraints of a laboratory, there are no such studies available for mountain or cyclocross cycling. CP however is relevant in both of these events [55]. With limited published research available on CP track cycling, events like the 3000 m (female) and 4000 m (male and female) individual and team pursuits, track cycling would offer ideal research conditions, such as a controlled environment whilst collecting real-world data.

Another key area for future research can be suggested by an investigation of CP based training zones. FTP, for which power based training zones already exist [16] spans a tolerable duration of ~ 60 min and as such does not identify a particular intensity domain demarcation point. CP spans a narrower training zone [113,132] and consequently as a valid marker between the very heavy and severe intensity domain (Table 3) could be used as baseline value from which other training zones could be defined. Therefore training prescription could theoretically be based on CP without the reliance of additional physiological markers.

Finally, the improved protocol using 30 min inter-trial recovery durations and/or using only 2 TTs or TTE trials may enhance the utility of CP determination in research and clinical settings. As suggested by Whipp and Ward [379], CP offers a more appropriate quantitative index for the interpretation of improved exercise tolerance, which in a clinical setting is commonly tested with a single unfamiliarised TTE trial.

In short, the overall significantly increased ease of CP related data collection, using either field testing data or training/performance data offers abundant opportunities for applied sport, race, clinical or gender specific research investigations but also opportunities to standardise testing procedures and to translate CP determination into a clinical setting.

9.4 Final Conclusions

Since its original investigation by Monod and Scherrer [51] in 1965 the CP model and its underpinning power-duration relationship has received extensive research attention. This is not surprising giving that the power-duration relationship both predicts and describes exercise tolerance in not just humans but also some animal species [107,109,110,380].

The findings of this research thesis extend our ability to perform CP testing in a more athlete-friendly way and to perform CP testing in the field. Results have shown that CP can be tested using a 30 min inter-trial recovery period and that it can also be tested in the field, i.e. on the track and on the road when using trained, recreational cyclists.

Being non-invasive, given the relative ease of testing and simple calculation procedure, CP may now be recognised by coaches and athletes as being advantageous over alternative endurance performance index marker. Testing cyclists in their 'natural training and racing environment' enhances the ecological validity of such tests. In short, CP might now become a regularly tested or simply 'modelled' aerobic performance index for coaches and athletes who work with power, making this a day to day assessment tool for the sport of cycling. Furthermore the aforementioned results have been demonstrated to be valid and reliable in recreational athletes. Accordingly the reliability of the presented CP field testing options in elite cyclists, who are more experienced, might be even higher [22]

The research above presents a substantial bridge between the sports science laboratory and the real-world of recreational and elite competitive cycling.

REFERENCES

1. BBC Sport - London 2012: Which sports won and lost at GB's home Olympics? Available at: http://www.bbc.co.uk/sport/0/olympics/19226042. Accessed 14 February 2014.

2. Grous, A. *The british cycling economy*. 1–17 (2011). Availbe at: http://www.britishcycling.org.uk. Accessed 14 February 2014.

3. Winston, Lord. Sport and exercise science and medicine : building on the Olympic legacy to improve the nation 's health. *Sel. Comm. Sci. Technol.* (2012).

4. Department of Health. Start active, stay active: a report on physical activity from the four home countries'. *Chief Med. Off.* (2011).

5. Palmer, G. S., Dennis, S. C., Noakes, T. D. & Hawley, J. A. Assessment of the reproducibility of performance testing on an air-braked cycle ergometer. *Int. J. Sports Med.* **17,** 293–8 (1996).

6. MacDougall, J.Duncan; Wenger HA, Green, H. *Physiological Testing of the High-performance Athlete*. (Human Kinetics Publishers, 1991).

7. Jeukendrup, A., Saris, W. H., Brouns, F. & Kester, A. D. A new validated endurance performance test. *Med. Sci. Sports Exerc.* **28,** 266–70 (1996).

8. Paton, C. D. & Hopkins, W. G. Tests of Cycling Performance. *Sport. Med.* **31,** 489–496 (2001).

9. Peveler, W. W. The Accuracy of Simulated Indoor Time Trials Utilizing a CompuTrainer and GPS Data. *J. Strength Cond. Res.* (2013). doi:10.1519/JSC.0b013e318280ce76

10. Bertucci, W., Taiar, R. & Grappe, F. Differences between sprint tests under laboratory and actual cycling conditions. *J. Sports Med. Phys. Fitness* **45,** 277–83 (2005).

11. Quod, M., Martin, D. & Martin, J. The power profile predicts road cycling MMP. *Int J Sport. Med* **31,** 397–401 (2010).

12. González-Haro, C., Galilea, P. A., Drobnic, F. & Escanero, J. F. Validation of a field test to determine the maximal aerobic power in triathletes and endurance cyclists. *Br. J. Sports Med.* **41,** 174–9 (2007).

13. Allen, H. & Coggan, A. Training and Racing With a Power Meter Journal. *2006.* Available at: http://www.trainingandracingwithapowermeter.com/. Accessed 21 February 2014.

14. Glaister, M. Multiple Sprint Work. *Sport. Med.* **35,** 757–777 (2005).

15. Powermeter im Vergleich - die aktuellen Wattmesser - ECL. Available at: http://www.cycling-league.eu/themenseiten/powermeter/136-alle-powermeter-im-%C3%BCberblick-und-vergleich. Accessed 22 February 2014.

16. Hunter, A. & Coggan, A. *Training and Racing with a Power Meter*. p. 5–20, 41–65 (Velo Press, 2010).

17. Jobson, S. a, Passfield, L., Atkinson, G., Barton, G. & Scarf, P. The analysis and utilisation of cycling training data. *Sports Med.* **39,** 833–44 (2009).

18. Jones, A. M., Wilkerson, D. P., DiMenna, F., Fulford, J. & Poole, D. C. Muscle metabolic responses to exercise above and below the "critical power" assessed using 31P-MRS. *Am. J. Physiol. Regul. Integr. Comp. Physiol.* **294,** R585–93 (2008).

19. Jones, A. M., Vanhatalo, A., Burnley, M., Morton, R. H. & Poole, D. C. Critical power: implications for determination of V˙O2max and exercise tolerance. *Med. Sci. Sports Exerc.* **42,** 1876–90 (2010).

20. Coyle, E., Feltner, M. & Kautz, S. Physiological and biomechanical factors associated with elite endurance cycling performance. *Med Sci Sport. Exerc.* **23,** 93–107 (1991).

21. Passfield, L., Jobson, S. A., Atkinson, G., Barton, G. & Scarf, P. Longitudinal changes in weighted indices of work-rate but not average power output of cyclists during a racing season. *J Sport. Sci* **27,** S102–03 (2009).

22. Lamberts, R. & Swart, J. Measurement error associated with performance testing in well-trained cyclists; application to the precision of monitoring changes in training status. *Int. Sport. Med. J.* **10,** 33–44 (2009).

23. Meeusen, R. *et al.* Prevention, diagnosis, and treatment of the overtraining syndrome: joint consensus statement of the European College of Sport Science and the American College of Sports Medicine. *Med. Sci. Sports Exerc.* **45,** 186–205 (2013).

24. Bertucci, W., Grappe, F. & Groslambert, A. Laboratory versus outdoor cycling conditions: differences in pedaling biomechanics. *J. Appl. Biomech.* **23,** 87–92 (2007).

25. Carter, J. B., Banister, E. W. & Blaber, A. P. Effect of Endurance Exercise on Autonomic Control of Heart Rate. *Sport. Med.* **33,** 33–46 (2003).

26. Thomas, K., Stone, M., St Clair Gibson, A., Thompson, K. & Ansley, L. The effect of an even-pacing strategy on exercise tolerance in well-trained cyclists. *Eur. J. Appl. Physiol.* **113,** 3001–10 (2013).

27. Atkinson, G. & Brunskill, A. Pacing strategies during a cycling time trial with simulated headwinds and tailwinds. *Ergonomics* **43,** 1449–60 (2000).

28. Balmer, J., Davison, R. C., Coleman, D. A. & Bird, S. R. The validity of power output recorded during exercise performance tests using a Kingcycle air-braked cycle ergometer when compared with an SRM powermeter. *Int. J. Sports Med.* **21,** 195–9 (2000).

29. Balmer, J., Bird, S., Davison, R. C. R., Doherty, M. & Smith, P. Mechanically braked Wingate powers: agreement between SRM, corrected and conventional methods of measurement. *J. Sports Sci.* **22,** 661–7 (2004).

30. Bertucci, W., Duc, S., Villerius, V. & Grappe, F. Validity and reliability of the Axiom PowerTrain cycle ergometer when compared with an SRM powermeter. *Int. J. Sports Med.* **26,** 59–65 (2004).

31. Stapelfeldt, B., Mornieux, G., Oberheim, R., Belli, A. & Gollhofer, A. Development and evaluation of a new bicycle instrument for measurements of pedal forces and power output in cycling. *Int. J. Sports Med.* **28,** 326–32 (2007).

32. Van der Woude, L. H. V *et al.* Power output and metabolic cost of synchronous and asynchronous submaximal and peak level hand cycling on a motor driven treadmill in able-bodied male subjects. *Med. Eng. Phys.* **30,** 574–80 (2008).

33. Millet, G. P., Tronche, C., Fuster, N., Bentley, D. J. & Candau, R. Validity and reliability of the Polar S710 mobile cycling powermeter. *Int. J. Sports Med.* **24,** 156–61 (2003).

34. Bertucci, W. Analysis of the agreement between the Fortius cycling ergometer and the PowerTap powermeter PO during time trials of 6 and 30 min. *Comput. Methods Biomech. Biomed. Engin.* **15 Suppl 1,** 212–4 (2012).

35. Siedlik, J. A. *et al.* Methods Comparison: Assessing Agreement of physiological parameters obtained from exercise on two different cycle ergometers. *J. Strength Cond. Res.* (2013).

36. Chiementin, X., Crequy, S. & Bertucci, W. Validity and reliability of the G-Cog device for kinematic measurements. *Int. J. Sports Med.* **34,** 945–9 (2013).

37. Duc, S., Villerius, V., Bertucci, W. & Grappe, F. Validity and reproducibility of the ErgomoPro power meter compared with the SRM and Powertap power meters. *Int. J. Sports Physiol. Perform.* **2,** 270–81 (2007).

38. Jones, S. M. & Passfield, L. The dynamic calibration of bicycle power measuring cranks. *Eng. Sport* 265–274 (1998).

39. Martin, J. & Milliken, D. Validation of a mathematical model for road cycling power. *J. Appl. Biomech.* **14,** 276–291 (1998).

40. Lawton, E. W., Martin, D. T. & Lee, H. Validation of SRM power cranks using dynamic calibration. in *5th IOC World Congr. Sport Sci.* (1999).

41. Abbiss, C. R., Quod, M. J., Levin, G., Martin, D. T. & Laursen, P. B. Accuracy of the Velotron ergometer and SRM power meter. *Int. J. Sports Med.* **30,** 107–12 (2009).

42. Gardner, A. S. *et al.* Accuracy of SRM and Power Tap Power Monitoring Systems for Bicycling. *Med. Sci. Sport. Exerc.* **36,** 1252–1258 (2004).

43. Martin, J. ., Douglas, L., Cobb, J. ., McFadden, K. . & Coggan, A. . Validation of a mathematical model for road cycling power. *J. Appl. Biomech.* **14,** 276–291 (1998).

44. Bertucci, W., Duc, S., Villerius, V., Pernin, J. N. & Grappe, F. Validity and reliability of the PowerTap mobile cycling powermeter when compared with the SRM Device. *Int. J. Sports Med.* **26,** 868–873 (2005).

45. Why Power? Available at: http://www.srm.de/srm-training-system/why-power. Accessed 03.01.2014

46. Paton, C. D. & Hopkins, W. G. Ergometer error and biological variation in power output in a performance test with three cycle ergometers. *Int. J. Sports Med.* **27,** 444–7 (2006).

47. CycleOps PowerTap SL+ (Mountain Bike) | Singletracks Mountain Bike Blog. Available at: http://www.singletracks.com/blog/mtb-gear/cycleops-powertap-sl-mountain-bike. Accessed 05.01.2014.

48. Hill, A. V. *Muscular Movement in Man : the Factors governing Speed and Recovery from Fatigue.* p 104 (McGraw-Hill Book Company, Inc., 1927).

49. Joyner, M. J. & Coyle, E. F. Endurance exercise performance: the physiology of champions. *J. Physiol.* **586,** 35–44 (2008).

50. Skiba, P. F. & Tilbury-Davis, D. PhysFarm Training Systems. Available at: http://physfarm.com. Accessed 21.01.2014

51. Monod, H. & Scherrer, J. The work capacity of a synergic muscular group. *Ergonomics* **8,** 329–338 (1965).

52. Moritani, T., Nagata, A., Devries, H. A. & Muro, M. Critical power as a measure of physical work capacity and anaerobic threshold. *Ergonomics* **24,** 339–350 (1981).

53. Miura, A., Sato, H., Whipp, B. J. & Fukuba, Y. The effect of glycogen depletion on the curvature constant parameter of the power-duration curve for cycle ergometry. *Ergonomics* **43,** 133–41 (2000).

54. Hill, D. W. & Smith, J. C. A comparison of methods of estimating anaerobic work capacity. *Ergonomics* **36,** 1495–500 (1993).

55. Vanhatalo, A., Jones, A. M. & Burnley, M. Application of Critical Power in Sport What Is the Critical Power Concept ? *Int. J. Sports Physiol. Perform.* **6,** 128–136 (2011).

56. Vanhatalo, A. & Jones, A. M. Influence of Creatine Supplementation on the Parameters of the "All-Out Critical Power Test."*J. Exerc. Sci. Fit.* **7,** 9–17 (2009).

57. Parker Simpson, L., Jones, A. M., Vanhatalo, A. & Wilkerson, D. P. Influence of initial metabolic rate on the power-duration relationship for all-out exercise. *Eur. J. Appl. Physiol.* **112,** 2467–73 (2012).

58. Jones, A. M. & Carter, H. The effect of endurance training on parameters of aerobic fitness. *Sports Med.* **29,** 373–86 (2000).

59. Smith, C. G. M. & Jones, A. M. The relationship between critical velocity, maximal lactate steady-state velocity and lactate turnpoint velocity in runners. *Eur. J. Appl. Physiol.* **85,** 19–26 (2001).

60. Dekerle, J., Baron, B., Dupont, L., Vanvelcenaher, J. & Pelayo, P. Maximal lactate steady state, respiratory compensation threshold and critical power. *Eur. J. Appl. Physiol.* **89,** 281–8 (2003).

61. McLellan, T. & Cheung, K. A comparative evaluation of the individual anaerobic threshold and the critical power. *Med. Sci. Sports Exerc.* **24,** 543–550 (1992).

62. Jenkins, D. Endurance training enhances critical power. *Med Sci Sport. Exerc.* **24,** 1283–1289 (1992).

63. Smith, J. C., Dangelmaier, B. S. & Hill, D. W. Critical power is related to cycling time trial performance. *Int. J. Sports Med.* **20,** 374–8 (1999).

64. De Lucas, R. D., de Souza, K. M., Costa, V. P., Grossl, T. & Guglielmo, L. G. A. Time to exhaustion at and above critical power in trained cyclists: The relationship between heavy and severe intensity domains. *Sci. Sports* **28,** e9–e14 (2013).

65. Poole, D. C., Ward, S. A., Gardner, G. W. & Whipp, B. J. Metabolic and respiratory profile of the upper limit for prolonged exercise in man. *Ergonomics* **31,** 1265–79 (1988).

66. Housh, D. & Terry, J. The accuracy of the critical power test for predicting time to exhaustion during cycle ergometry. *Ergonomics* **32,** 997–1004 (1989).

67. Poole, D. C., Ward, S. A. & Whipp, B. J. The effects of training on the metabolic and respiratory profile of high-intensity cycle ergometer exercise. *Eur. J. Appl. Physiol. Occup. Physiol.* **59,** 421–9 (1990).

68. Jenkins, D. & Quigley, B. The influence of high-intensity exercise training on the Wlim-Tlim relationship. *Med. Sci. Sports Exerc.* **25,** 275–82 (1993).

69. Stickland, M. K., Petersen, S. R. & Dressendorfer, R. H. Critical aerobic power during simulated 20 km bicycle racing. *Sport. Med. Train. Rehabil.* **9,** 289–301 (2000).

70. Billat, L. V, Koralsztein, J. P. & Morton, R. H. Time in human endurance models. From empirical models to physiological models. *Sports Med.* **27,** 359–79 (1999).

71. Smith, D. J., Norris, S. R. & Hogg, J. M. Performance evaluation of swimmers: scientific tools. *Sports Med.* **32,** 539–54 (2002).

72. Gaesser, G. A. & Wilson, L. A. Effects of continuous and interval training on the parameters of the power-endurance time relationship for high-intensity exercise. *Int. J. Sports Med.* **9,** 417–421 (1988).

73. Nebelsick-Gullett, L. J., Housh, T. J., Johnson, G. O. & Bauge, S. M. A comparison between methods of measuring anaerobic work capacity. *Ergonomics* **31,** 1413–9 (1988).

74. Denadai, B. S., Gomide, E. B. G., Greco, C. C. & Dendai BS, Gomide EBG, G. C. The Relationship Between Onset of Blood Lacate Accumulation, Critical velocity, and Maximal Lactate Steady State in Soccer Players. *J. Strength Cond. Res.* **19,** 364–368 (2005).

75. Kolbe, T., Dennis, S., Selley, E., Noakes, T. D. & Lambert, M. I. The relationship between critical power and running performance. *J. Sport.* **5,** 265–269 (1995).

76. Gavin, T. P. *et al.* Comparison of a field-based test to estimate functional threshold power and power output at lactate threshold. *J. Strength Cond. Res.* **26,** 416–21 (2012).

77. García-Manso, J. M., Martín-González, J. M., Dávila, N. & Arriaza, E. Middle and long distance athletics races viewed from the perspective of complexity. *J. Theor. Biol.* **233,** 191–198 (2005).

78. Louis Passfield, Sarah Coakley, S. J. Cycling performance is described by the power-law not critical power. in *Eur. Conf. Sport. Sci.* Abstract ID 2218 (2013).

79. Hill, D. W. The critical power concept. A review. *Sports Med.* **16,** 237–54 (1993).

80. Morton, R. H. The critical power and related whole-body bioenergetic models. *Eur. J. Appl. Physiol.* **96,** 339–54 (2006).

81. Pettitt, R. W. Using scatterplots to teach the critical power concept. *Adv. Physiol. Educ.* **36,** 134–152 (2012).

82. Copp, S. W., Hirai, D. M., Musch, T. I. & Poole, D. C. Critical speed in the rat: implications for hindlimb muscle blood flow distribution and fibre recruitment. *J. Physiol.* **588,** 5077–87 (2010).

83. Bassett, D. R. & Howley, E. T. Limiting factors for maximum oxygen uptake and determinants of endurance performance. *Med. Sci. Sports Exerc.* **32,** 70–84 (2000).

84. Dekerle, J., Mucci, P. & Carter, H. Influence of moderate hypoxia on tolerance to high-intensity exercise. *Eur. J. Appl. Physiol.* **112,** 327–35 (2012).

85. Vanhatalo, A., Fulford, J., DiMenna, F. J. & Jones, A. M. Influence of hyperoxia on muscle metabolic responses and the power-duration relationship during severe-intensity exercise in humans: a 31P magnetic resonance spectroscopy study. *Exp. Physiol.* **95,** 528–40 (2010).

86. Ferguson, C. *et al.* Effects of prior very-heavy intensity exercise on indices of aerobic function and high-intensity exercise tolerance. *J. Appl. Physiol.* **103,** 812–22 (2007).

87. Ferguson, C. *et al.* Effect of recovery duration from prior exhaustive exercise on the parameters of the power-duration relationship. *J. Appl. Physiol.* **108,** 866–74 (2010).

88. Chidnok, W. *et al.* Muscle metabolic determinants of exercise tolerance following exhaustion: relationship to the "critical power."*J. Appl. Physiol.* 243–250 (2013).

89. Skiba, P. F., Chidnok, W., Vanhatalo, A. & Jones, A. M. Modeling the expenditure and reconstitution of work capacity above critical power. *Med. Sci. Sports Exerc.* **44,** 1526–32 (2012).

90. Burnley, M., Doust, J. H. & Vanhatalo, A. A 3-min all-out test to determine peak oxygen uptake and the maximal steady state. *Med. Sci. Sports Exerc.* **38,** 1995–2003 (2006).

91. Murgatroyd, S. R. & Wylde, L. A. The power-duration relationship of high-intensity exercise: from mathematical parameters to physiological mechanisms. *J. Physiol.* **589,** 2443–5 (2011).

92. Vogt, S. *et al.* Power output during stage racing in professional road cycling. *Med. Sci. Sports Exerc.* **38,** 147–51 (2006).

93. Mujika, I. & Padilla, S. Physiological and performance characteristics of male professional road cyclists. *Sports Med.* **31,** 479–87 (2001).

94. Atkinson, G., Peacock, O., St Clair Gibson, A. & Tucker, R. Distribution of power output during cycling: impact and mechanisms. *Sports Med.* **37,** 647–67 (2007).

95. Hill, D. W., Alain, C. & Kennedy, M. D. Modeling the relationship between velocity and time to fatigue in rowing. *Med. Sci. Sports Exerc.* **35,** 2098–105 (2003).

96. Clingeleffer, A., Naughton, L. R. & Davoren, B. The use of critical power as a determinant for establishing the onset of blood lactate accumulation. *Eur. J. Appl. Physiol. Occup. Physiol.* **68,** 182–187 (1994).

97. Clingeleffer, A., Naughton, L. M. & Davoren, B. Critical power may be determined from two tests in elite kayakers. *Eur. J. Appl. Physiol. Occup. Physiol.* **68,** 36–40 (1994).

98. Wang, T. Y., Lee, M. . & Chang, K. H. High-intensity interval and moderate-continous training have similar effects on canoeing ergometer performance. in *2010 Anu. Meet. Am. Coll. Sport. Med. Balitomore, Maryl.* (2011).

99. Zagatto, A., Miranda, M. & Gobatto, C. Critical power concept adapted for the specific table tennis test: comparisons between exhaustion criteria, mathematical modeling, and correlation with gas exchange. *Int. J. Sports Med.* **8,** 503–510 (2011).

100. Fukuda, D. H., Smith, A. E., Kendall, K. L., Cramer, J. T. & Stout, J. R. The determination of critical rest interval from the intermittent critical velocity test in club-level collegiate hockey and rugby players. *J. Strength Cond. Res.* **25,** 889–95 (2011).

101. Arabi, H., Vandewalle, H., Kapitaniak, B. & Monod, H. Evaluation of wheelchair users in the field and in laboratory: Feasibility of progressive tests and critical velocity tests. *Int. J. Ind. Ergon.* **24,** 483–491 (1999).

102. Barker, A. R., Bond, B., Toman, C., Williams, C. a & Armstrong, N. Critical power in adolescents: physiological bases and assessment using all-out exercise. *Eur. J. Appl. Physiol.* **112,** 1359–70 (2012).

103. Toubekis, A. G., Vasilaki, A., Douda, H., Gourgoulis, V. & Tokmakidis, S. Physiological responses during interval training at relative to critical velocity intensity in young swimmers. *J. Sci. Med. Sport* **14,** 363–8 (2011).

104. Zacca, R. *et al.* Critical velocity, anaerobic distance capacity, maximal instantaneous velocity and aerobic inertia in sprint and endurance young swimmers. *Eur. J. Appl. Physiol.* **110,** 121–31 (2010).

105. Overend, T. J., Cunningham, D. A., Paterson, D. H. & Smith, W. D. F. Physiological responses of young and elderly men to prolonged exercise at critical power. *Eur. J. Appl. Physiol. Occup. Physiol.* **64,** 187–193 (1992).

106. Murgatroyd, S. R., Ferguson, C., Ward, S. A., Whipp, B. J. & Rossiter, H. B. Pulmonary O2 uptake kinetics as a determinant of high-intensity exercise tolerance in humans. *J. Appl. Physiol.* **110,** 1598–606 (2011).

107. Lauderdale, M. A. & Hinchcliff, K. W. Hyperbolic relationship between time-to-fatigue and workload. *Equine Vet. J.* **31,** 586–590 (2010).

108. Full, R. J. Locomotion without lungs: energetics and performance of a lungless salamander. *Am J Physiol Regul. Integr. Comp Physiol* **251,** R775–780 (1986).

109. Full, R. J. & Herreid, C. F. Aerobic response to exercise of the fastest land crab. *Am. J. Physiol.* **244,** R530–6 (1983).

110. Billat, V. L., Mouisel, E., Roblot, N. & Melki, J. Inter- and intrastrain variation in mouse critical running speed. *J. Appl. Physiol.* **98,** 1258–63 (2005).

111. Plaut, I. Critical swimming speed: its ecological relevance. *Comp. Biochem. Physiol. A. Mol. Integr. Physiol.* **131,** 41–50 (2001).

112. Massé, A. J., Thomas, J. K. & Janz, D. M. Reduced swim performance and aerobic capacity in adult zebrafish exposed to waterborne selenite. *Comp. Biochem. Physiol. C. Toxicol. Pharmacol.* **157,** 266–71 (2013).

113. Pringle, J. S. M. & Jones, A. M. Maximal lactate steady state, critical power and EMG during cycling. *Eur. J. Appl. Physiol.* **88,** 214–26 (2002).

114. Beneke, R. & Duvillard, S. P. Determination of maximal lactate steady state response in selected sports events. *Med. Sci. Sports Exerc.* **28,** 241–246 (1996).

115. Billat, V. L., Sirvent, P., Py, G., Koralsztein, J.-P. & Mercier, J. The Concept of Maximal Lactate Steady State. *Sport. Med.* **33,** 407–426 (2003).

116. Brooks, G. a. Intra- and extra-cellular lactate shuttles. *Med. Sci. Sports Exerc.* **32,** 790–9 (2000).

117. Antonutto, G. & Di Prampero, P. E. The concept of lactate threshold. A short review. *J. Sports Med. Phys. Fitness* **35,** 6–12 (1995).

118. Hoogeveen, A. R., Hoogsteen, J. & Schep, G. The maximal lactate steady state in elite endurance athletes. *Jpn. J. Physiol.* **47,** 481–5 (1997).

119. Leti, T., Mendelson, M., Laplaud, D. & Flore, P. Prediction of maximal lactate steady state in runners with an incremental test on the field. *J. Sports Sci.* **30,** 609–16 (2012).

120. Laplaud, D., Guinot, M., Favre-Juvin, A. & Flore, P. Maximal lactate steady state determination with a single incremental test exercise. *Eur. J. Appl. Physiol.* **96,** 446–52 (2006).

121. Beneke, R., Heck, H., Hebestreit, H. & Leithäuser, R. M. Predicting maximal lactate steady state in children and adults. *Pediatr. Exerc. Sci.* **21,** 493–505 (2009).

122. Harnish, C. R., Swensen, T. C. & Pate, R. R. Methods for estimating the maximal lactate steady state in trained cyclists. *Med. Sci. Sports Exerc.* **33,** 1052–5 (2001).

123. Kuphal, K. E., Potteiger, J. A., Frey, B. B. & Hise, M. P. Validation of a single-day maximal lactate steady state assessment protocol. *J. Sports Med. Phys. Fitness* **44,** 132–40 (2004).

124. Mamen, A. & Tillaar, R. Van Den. Estimating the Maximal Lactate Steady State Power from an Incremental Test Using Lactate Pro ® LP1710. *Int. J. Appl. Sport. Sci.* **21,** 74–85 (2009).

125. Beneke, R., Hütler, M. & Leithäuser, R. M. Maximal lactate-steady-state independent of performance. *Med. Sci. Sports Exerc.* **32,** 1135–9 (2000).

126. Caritá, R. A. C., Greco, C. C. & Denadai, B. S. Maximal lactate steady state and critical power in well-trained cyclists. *Rev. Bras. Med. do Esporte* **15,** 370–373 (2009).

127. Jenkins, D. G. & Quigley, B. M. Blood lactate in trained cyclists during cycle ergometry at critical power. *Eur J Appl Physiol* **61,** 278–283 (1990).

128. Hill, D. W. & Ferguson, C. S. A physiological description of critical velocity. *Eur. J. Appl. Physiol. Occup. Physiol.* **79,** 290–3 (1999).

129. Pepper, M. L., Housh, T. J. & Johnson, G. O. The accuracy of the critical velocity test for predicting time to exhaustion during treadmill running. *Int. J. Sports Med.* **13,** 121–4 (1992).

130. Wakayoshi, K. *et al.* Does critical swimming velocity represent exercise intensity at maximal lactate steady state? *Eur. J. Appl. Physiol. Occup. Physiol.* **66,** 90–5 (1993).

131. Sid-Ali, B., Vandewalle, H., Chaïr, K., Moreaux, A. & Monod, H. Lactate steady state velocity and distance-exhaustion time relationship in running. *Arch. Physiol. Biochem.* **99,** 297–301 (2008).

132. Brickley, G., Doust, J. & Williams, C. a. Physiological responses during exercise to exhaustion at critical power. *Eur. J. Appl. Physiol.* **88,** 146–51 (2002).

133. Dekerle, J. *et al.* Critical swimming speed does not represent the speed at maximal lactate steady state. *Int. J. Sports Med.* **26,** 524–30 (2005).

134. Carter, H. The physiological basis of exercise training zones. *J. Cycl. Coach.* (2009).

135. Whipp, B. J. & Wasserman, K. Oxygen uptake kinetics for various intensities of constant-load work. *J Appl Physiol* **33,** 351–356 (1972).

136. Whipp, B. J. Physiological mechanisms dissociating pulmonary CO2 and O2 exchange dynamics during exercise in humans. *Exp. Physiol.* **92,** 347–55 (2007).

137. Carter, H., Pringle, J., Jones, A. & Doust, J. Oxygen uptake kinetics during treadmill running across exercise intensity domains. *Eur. J. Appl. Physiol.* **86,** 347–354 (2002).

138. Jones, A. & Poole, D. in *Oxyg. uptake Kinet. Sport. Exerc. Med.* (Jones, A. M. & Poole, D. C.) 3–38 (Routledge, London, 2005).

139. Wilkerson, D. P., Koppo, K., Barstow, T. J. & Jones, A. M. Effect of work rate on the functional "gain" of Phase II pulmonary O2 uptake response to exercise. *Respir. Physiol. Neurobiol.* **142,** 211–23 (2004).

140. Gaesser, G. & Poole, D. The slow component of oxygen uptake kinetics in humans. *Exerc. Sport Sci. Rev.* **24,** 35–71 (1996).

141. Whipp, B. J., Ward, S. a. & Rossiter, H. B. Pulmonary O2 Uptake during Exercise: Conflating Muscular and Cardiovascular Responses. *Med. Sci. Sport. Exerc.* **37,** 1574–1585 (2005).

142. Poole, D. C. Resolving the determinants of high-intensity exercise performance. *Exp. Physiol.* **94,** 197–8 (2009).

143. Poole, D. C. & Jones, A. M. Oxygen uptake kinetics. *Compr. Physiol.* **2,** 933–96 (2012).

144. Czuba, M., Zając, A. & Cholewa, J. Lactate threshold (d-max method) and maximal lactate steady state in cyclists. *J. Hum. Kinet.* **21,** 49–56 (2009).

145. Beaver, W. L., Wasserman, K. & Whipp, B. J. A new method for detecting anaerobic threshold by gas exchange. *J. Appl. Physiol.* **60,** 2020–7 (1986).

146. Beneke, R., Leithäuser, R. M. & Ochentel, O. Blood lactate diagnostics in exercise testing and training. *Int. J. Sports Physiol. Perform.* **6,** 8–24 (2011).

147. Wasserman, K., Beaver, W. L. & Whipp, B. J. Gas exchange theory and the lactic acidosis (anaerobic) threshold. *Circulation* **81,** II14–30 (1990).

148. Wasserman, K. *Principles of Exercise Testing and Interpretation: Including Pathophysiology and Clinical Applications*. 592 (Lippincott Williams and Wilkins, 2004).

149. Diamond, L. B., Casaburi, R., Wasserman, K. & Whipp, B. J. Kinetics of gas exchange and ventilation in transitions from rest or prior exercise. *J. Appl. Physiol.* **43,** 704–8 (1977).

150. Coyle, E. F. Substrate utilisation during exercise in active people. *Am. J. Clin. Nutr.* **61,** 968S–979S (1995).

151. González-Alonso, J., Mora-Rodríguez, R., Below, P. R. & Coyle, E. F. Dehydration markedly impairs cardiovascular function in hyperthermic endurance athletes during exercise. *J. Appl. Physiol.* **82,** 1229–36 (1997).

152. St Clair Gibson, A., Schabort, E. J. & Noakes, T. D. Reduced neuromuscular activity and force generation during prolonged cycling. *Am. J. Physiol. Regul. Integr. Comp. Physiol.* **281,** R187–96 (2001).

153. Bigland-Ritchie, B. & Woods, J. J. Changes in muscle contractile properties and neural control during human muscular fatigue. *Muscle Nerve* **7,** 691–9 (1984).

154. Camic, C. L. *et al.* An EMG frequency-based test for estimating the neuromuscular fatigue threshold during cycle ergometry. *Eur. J. Appl. Physiol.* **108,** 337–45 (2010).

155. Jones, A. M., Vanhatalo, A. & Doust, J. *Kinanthropometry and Exercise Physiology Laboratory Manual: Exercise physiology*. 271–306 (Taylor & Francis, 2009).

156. Whipp, B. J. & Ward, S. A. Physiological determinants of pulmonary gas exchange kinetics during exercise. *Med. Sci. Sports Exerc.* **22,** 62–71 (1990).

157. Whipp, B. The slow component of O2 uptake kinetics during heavy exercise. *Med. Sci. Sports Exerc.* **26,** 1319–26 (1994).

158. Henson, L. C., Poole, D. C. & Whipp, B. J. Fitness as a determinant of oxygen uptake response to constant-load exercise. *Eur. J. Appl. Physiol. Occup. Physiol.* **59,** 21–8 (1989).

159. Whipp, B. J. The slow component of O2 uptake kinetics during heavy exercise. *Med. Sci. Sports Exerc.* **26,** 1319–26 (1994).

160. MacDougall, J. D., Reddan, W. G., Layton, C. R. & Dempsey, J. A. Effects of metabolic hyperthermia on performance during heavy prolonged exercise. *J. Appl. Physiol.* **36,** 538–44 (1974).

161. Carter, H., Pringle, J. S. M., Boobis, L., Jones, A. M. & Doust, J. H. Muscle Glycogen Depletion Alters Oxygen Uptake Kinetics during Heavy Exercise. *Med. Sci. Sport. Exerc.* **36,** 965–972 (2004).

162. Whipp, B. J. in *Physiol. Pathophysiol. Exerc. Toler.* (Steinacker, J. M. & Ward, S. A.) 83–89 (Springer US, 1996). doi:10.1007/978-1-4615-5887-3

163. Ozyener, F., Rossiter, H. B., Ward, S. a & Whipp, B. J. Influence of exercise intensity on the on- and off-transient kinetics of pulmonary oxygen uptake in humans. *J. Physiol.* **533,** 891–902 (2001).

164. Poole, D. C., Ward, S. A., Gardner, G. W. & Whipp, B. J. Metabolic and respiratory profile of the upper limit for prolonged exercise in man. *Ergonomics* **31,** 1265–79 (1988).

165. Endo, M. Y. *et al.* Thigh muscle activation distribution and pulmonary VO2 kinetics during moderate, heavy, and very heavy intensity cycling exercise in humans. *Am. J. Physiol. Regul. Integr. Comp. Physiol.* **293,** R812–20 (2007).

166. Krustrup, P., Söderlund, K., Mohr, M. & Bangsbo, J. The slow component of oxygen uptake during intense, sub-maximal exercise in man is associated with additional fibre recruitment. *Pflugers Arch.* **447,** 855–66 (2004).

167. Amann, M. Central and peripheral fatigue: interaction during cycling exercise in humans. *Med. Sci. Sports Exerc.* **43,** 2039–45 (2011).

168. Zoladz, J. A., Gladden, L. B., Hogan, M. C., Nieckarz, Z. & Grassi, B. Progressive recruitment of muscle fibers is not necessary for the slow component of VO2 kinetics. *J. Appl. Physiol.* **105,** 575–80 (2008).

169. Hepple, R. T., Howlett, R. A., Kindig, C. A., Stary, C. M. & Hogan, M. C. The O2 cost of the tension-time integral in isolated single myocytes during fatigue. *Am. J. Physiol. Regul. Integr. Comp. Physiol.* **298,** R983–8 (2010).

170. Sih, B., Ng, L. & Stuhmiller, J. Generalization of a model based on biophysical concepts of muscle activation, fatigue and recovery that explains exercise performance. *Int. J. Sports Med.* **33,** 258–67 (2012).

171. Westerblad, H., Allen, D. G. & Lännergren, J. The Major Cause ? Muscle Fatigue : Lactic Acid or Inorganic Phosphate the Major Cause ? *News Physiol. Sci.* **17,** 17–21 (2002).

172. Poole, D. C., Ward, S. A. & Whipp, B. J. The effects of training on the metabolic and respiratory profile of high-intensity cycle ergometer exercise. *Eur. J. Appl. Physiol. Occup. Physiol.* **59,** 421–429 (1990).

173. Hill, D. W. & Smith, J. C. Determination of Critical Power by Pulmonary Gas Exchange. *Can. J. Appl. Physiol.* **24,** 74–86 (1999).

174. Hill, D. & Poole, DC, S. J. The relationship between power and the time to achieve VO2max. *Med Sci Sport. Exerc.* 709–714 (2002).

175. Vautier, J. F., Vandewalle, H., Arabi, H. & Monod, H. Critical power as an endurance index. *Appl. Ergon.* **26,** 117–21 (1995).

176. McLellan, T. M., Cheung, S. S. & Jacobs, I. Variability of time to exhaustion during submaximal exercise. *Can. J. Appl. Physiol.* **20,** 39–51 (1995).

177. Carter, H. *et al.* Effect of Prior Exercise above and below Critical Power on Exercise to Exhaustion. *Med. Sci. Sport. Exerc.* **37,** 775–781 (2005).

178. Hopkins, W. G., Schabort, E. J. & Hawley, J. A. Reliability of power in physical performance tests. *Sports Med.* **31,** 211–34 (2001).

179. Bull, A. J., Housh, T. J., Johnson, G. O. & Rana, S. R. Physiological responses at five estimates of critical velocity. *Eur. J. Appl. Physiol.* **102,** 711–20 (2008).

180. Bull, A. J., Housh, T. J., Johnson, G. O. & Perry, S. R. Effect of mathematical modeling on the estimation of critical power. *Med. Sci. Sports Exerc.* **32,** 526–30 (2000).

181. Housh, T. J., Cramer, J. T., Bull, A. J., Johnson, G. O. & Housh, D. J. The effect of mathematical modeling on critical velocity. *Eur. J. Appl. Physiol.* **84,** 469–475 (2001).

182. Hill, D. W. The relationship between power and time to fatigue in cycle ergometer exercise. *Int. J. Sports Med.* **25,** 357–61 (2004).

183. Billat, V., Binsse, V., Petit, B. & Koralsztein, J. P. High level runners are able to maintain a VO2 steady-state below VO2max in an all-out run over their critical velocity. *Arch Physiol Biochem.* **106,** 38–45 (1998).

184. Roston, W. L. *et al.* Oxygen uptake kinetics and lactate concentration during exercise in humans. *Am. Rev. Respir. Dis.* **135,** 1080–4 (1987).

185. Brickley, G. *et al.* Muscle metabolism during constant- and alternating-intensity exercise around critical power. *Int. J. Sports Med.* **28,** 300–5 (2007).

186. Jeukendrup, A. E., Craig, N. P. & Hawley, J. A. The bioenergetics of World Class Cycling. *J. Sci. Med. Sport* **3,** 414–33 (2000).

187. Armstrong, R. B. & Laughlin, M. H. Blood flows within and among rat muscles as a function of time during high speed treadmill exercise. *J. Physiol.* **344,** 189–208 (1983).

188. Armstrong, R. B. & Laughlin, M. H. Exercise blood flow patterns within and among rat muscles after training. *Am J Physiol Hear. Circ Physiol* **246,** H59–68 (1984).

189. Laughlin, M. H. & Armstrong, R. B. Muscular blood flow distribution patterns as a function of running speed in rats. *Am J Physiol Hear. Circ Physiol* **243,** H296–306 (1982).

190. Heglund, N. C. & Cavagna, G. A. Mechanical work, oxygen consumption, and efficiency in isolated frog and rat muscle. *Am J Physiol Cell Physiol* **253,** C22–29 (1987).

191. Gollnick, P. D., Piehl, K. & Saltin, B. Selective glycogen depletion pattern in human muscle fibres after exercise of varying intensity and at varying pedalling rates. *J. Physiol.* **241,** 45–57 (1974).

192. Vøllestad, N. K. & Blom, P. C. Effect of varying exercise intensity on glycogen depletion in human muscle fibres. *Acta Physiol. Scand.* **125,** 395–405 (1985).

193. Barclay, C. J. Mechanical efficiency and fatigue of fast and slow muscles of the mouse. *J. Physiol.* **497 (Pt 3,** 781–94 (1996).

194. Henneman, E., Somjen, G. & Carpenter, D. O. Excitability and inhibitability of motoneurons of different sizes. *J. Neurophysiol.* **28,** 599–620 (1965).

195. Grimby, L. & Hannerz, J. Firing rate and recruitment order of toe extensor motor units in different modes of voluntary conraction. *J. Physiol.* **264,** 865–879 (1977).

196. Dekerle, J. The use of critical velocity in swimming. A place for Critical Stroke Rate. *Port. J. Sport Sci.* **2,** 201–205 (2006).

197. Gaesser, G. A., Carnevale, T. J., Garfinkel, A., Walter, D. O. & Womack, C. J. Estimation of critical power with nonlinear and linear models. *Med. Sci. Sports Exerc.* **27,** 1430–8 (1995).

198. Ebert, T. R., Martin, D. T., Stephens, B. & Withers, R. T. Power output during a professional men's road-cycling tour. *Int. J. Sports Physiol. Perform.* **1,** 324–35 (2006).

199. Coats, E. M. *et al.* Intensity-dependent tolerance to exercise after attaining V(O2) max in humans. *J. Appl. Physiol.* **95,** 483–90 (2003).

200. Fukuba, Y. *et al.* The curvature constant parameter of the power-duration curve for varied-power exercise. *Med. Sci. Sports Exerc.* **35,** 1413–8 (2003).

201. Fukuda, D. H. *et al.* The effects of creatine loading and gender on anaerobic running capacity. **24,** 1826–1833 (2010).

202. Smith, J. C., Stephens, D. P., Hall, E. L., Jackson, a W. & Earnest, C. P. Effect of oral creatine ingestion on parameters of the work rate-time relationship and time to exhaustion in high-intensity cycling. *Eur. J. Appl. Physiol. Occup. Physiol.* **77,** 360–5 (1998).

203. Miura, A., Kino, F., Kajitani, S., Sato, H. & Fukuba, Y. The effect of oral creatine supplementation on the curvature constant parameter of the power-duration curve for cycle ergometry in humans. *Jpn. J. Physiol.* **49,** 169–74 (1999).

204. Heubert, R. A. P. *et al.* Effect of a previous sprint on the parameters of the work-time to exhaustion relationship in high intensity cycling. *Int. J. Sports Med.* **26,** 583–92 (2005).

205. Sawyer, B. J. *et al.* Strength Training Increases Endurance Time to Exhaustion During High-Intensity Exercise Despite no Change in Critical Power. *J. Strength Cond. Res.* (2013).

206. Clark, I. E., West, B., Reynolds, S., Murray, S. R. & Pettitt, R. W. Applying the Critical Velocity Model for an Off-Season Interval Training Program. *J. Strength Cond. Res.* (2013).

207. Morton, R. H. & Hodgson, D. J. The relationship between power output and endurance: a brief review. *Eur. J. Appl. Physiol. Occup. Physiol.* **73,** 491–502 (1996).

208. DiPrampero, P. E. The concept of critical velocity: a brief analysis. *Eur. J. Appl. Physiol. Occup. Physiol.* **80,** 162–4 (1999).

209. Jones, A. M., Koppo, K. & Burnley, M. Effects of prior exercise on metabolic and gas exchange responses to exercise. *Sports Med.* **33,** 949–71 (2003).

210. Forbes, S. C., Raymer, G. H., Kowalchuk, J. M., Thompson, R. T. & Marsh, G. D. Effects of recovery time on phosphocreatine kinetics during repeated bouts of heavy-intensity exercise. *Eur. J. Appl. Physiol.* **103,** 665–75 (2008).

211. Forbes, S. C., Paganini, A. T., Slade, J. M., Towse, T. F. & Meyer, R. a. Phosphocreatine recovery kinetics following low- and high-intensity exercise in human triceps surae and rat posterior hindlimb muscles. *Am. J. Physiol. Regul. Integr. Comp. Physiol.* **296,** R161–70 (2009).

212. Rossiter, H. B. *et al.* Dynamic asymmetry of phosphocreatine concentration and O(2) uptake between the on- and off-transients of moderate- and high-intensity exercise in humans. *J. Physiol.* **541,** 991–1002 (2002).

213. Jones, A. M., Wilkerson, D. P., Burnley, M. & Koppo, K. Prior heavy exercise enhances performance during subsequent perimaximal exercise. *Med. Sci. Sports Exerc.* **35,** 2085–92 (2003).

214. Burnley, M. & Jones, A. M. Oxygen uptake kinetics as a determinant of sports performance. *Eur. J. Sport Sci.* **7,** 63–79 (2007).

215. Jones, A. M. *et al.* Slow component of VO2 kinetics: mechanistic bases and practical applications. *Med. Sci. Sports Exerc.* **43,** 2046–62 (2011).

216. Bogdanis, G. C., Nevill, M. E., Boobis, L. H. & Lakomy, H. K. Contribution of phosphocreatine and aerobic metabolism to energy supply during repeated sprint exercise. *J. Appl. Physiol.* **80,** 876–84 (1996).

217. Bogdanis, G. C., Nevill, M. E., Boobis, L. H., Lakomy, H. K. & Nevill, A. M. Recovery of power output and muscle metabolites following 30 s of maximal sprint cycling in man. *J. Physiol.* **482 (Pt 2,** 467–80 (1995).

218. Fitts, R. H. Cellular mechanisms of muscle fatigue. *Physiol. Rev.* **74,** 49–94 (1994).

219. Johnson, M. A., Mills, D. E., Brown, P. I. & Sharpe, G. R. Prior Upper Body Exercise Reduces Cycling Work Capacity but Not Critical Power. *Med. Sci. Sports Exerc.* (2013).

220. Fukuba, Y. & Whipp, B. J. A metabolic limit on the ability to make up for lost time in endurance events. *J. Appl. Physiol.* **87,** 853–61 (1999).

221. Gastin, P. B. Energy system interaction and relative contribution during maximal exercise. *Sports Med.* **31,** 725–41 (2001).

222. Hill, A. V. The Physiological Basis of Athletic Records. *Nature* **116,** 544–548 (1925).

223. Dekerle, J., Brickley, G., Hammond, A. J. P., Pringle, J. S. M. & Carter, H. Validity of the two-parameter model in estimating the anaerobic work capacity. *Eur. J. Appl. Physiol.* **96,** 257–64 (2006).

224. Vandewalle, H., Vautier, J. F., Kachouri, M., Lechevalier, J. M. & Monod, H. Work-exhaustion time relationships and the critical power concept. A critical review. *J. Sports Med. Phys. Fitness* **37,** 89–102 (1997).

225. Amann, M. *et al.* Peripheral fatigue limits endurance exercise via a sensory feedback-mediated reduction in spinal motoneuronal output. *J. Appl. Physiol.* **115,** 355–64 (2013).

226. Jones, A. The physiology of the world record holder for the women's marathon. *Int. J. Sport. Sci. Coach.* **1,** 101–116 (2006).

227. Green, S. A definition and systems view of anaerobic capacity. *Eur. J. Appl. Physiol. Occup. Physiol.* **69,** 168–73 (1994).

228. St Clair Gibson, A. & Noakes, T. D. Evidence for complex system integration and dynamic neural regulation of skeletal muscle recruitment during exercise in humans. *Br. J. Sports Med.* **38,** 797–806 (2004).

229. Amann, M. & Dempsey, J. A. Locomotor muscle fatigue modifies central motor drive in healthy humans and imposes a limitation to exercise performance. *J. Physiol.* **586,** 161–73 (2008).

230. Morton, R. H. A 3-parameter critical power model. *Ergonomics* **39,** 611–9 (1996).

231. Chatagnon, M., Pouilly, J.-P., Thomas, V. & Busso, T. Comparison between maximal power in the power-endurance relationship and maximal instantaneous power. *Eur. J. Appl. Physiol.* **94,** 711–7 (2005).

232. Wilkie, D. in *Exerc. Bioenerg. gas Exch.* (Carretelli, P. & Whipp, B. J.) 75–80 (Elsevie, Amsterdam, 1980).

233. Chatagnon, M. & Busso, T. Modelling of aerobic and anaerobic energy production during exhaustive exercise on a cycle ergometer. *Eur. J. Appl. Physiol.* **97,** 755–60 (2006).

234. Busso, T., Gimenez, P. & Chatagnon, M. A comparison of modelling procedures used to estimate the power-exhaustion time relationship. *Eur. J. Appl. Physiol.* **108,** 257–63 (2010).

235. Hill, D. W., Vingren, J. L., Nakamura, F. Y. & Kokobun, E. Relationship between speed and time in running. *Int. J. Sports Med.* **32,** 519–22 (2011).

236. Morton, R. H. & Billat, L. V. The critical power model for intermittent exercise. *Eur. J. Appl. Physiol.* **91,** 303–307 (2004).

237. Bosquet, L., Delhors, P. R., Duchene, a, Dupont, G. & Leger, L. Anaerobic running capacity determined from a 3-parameter systems model: relationship with other anaerobic indices and with running performance in the 800 m-run. *Int. J. Sports Med.* **28,** 495–500 (2007).

238. Enoka, R. M. & Duchateau, J. Muscle fatigue: what, why and how it influences muscle function. *J. Physiol.* **586,** 11–23 (2008).

239. Sawyer, B. J., Morton, R. H., Womack, C. J. & Gaesser, G. a. VO2max May Not Be Reached During Exercise to Exhaustion above Critical Power. *Med. Sci. Sports Exerc.* **44,** 1533–1538 (2012).

240. Bulbulian, R., Jeong, J. & Murphy, M. Comparison of anaerobic components of the Wingate and Critical Power tests in males and females. *Med. Sci. Sports Exerc.* **28,** 1336–41 (1996).

241. Hughson, R. L., Orok, C. J. & Staudt, L. E. A high velocity treadmill running test to assess endurance running potential. *Int. J. Sports Med.* **5,** 23–5 (1984).

242. Coast, J. & Welch, H. Linear increase in optimal pedal rate with increased power output in cycle ergometry. *Eur. J. Appl. Physiol.* **53,** 339–342 (1985).

243. Coast, J. R., Cox, R. H. & Welch, H. G. Optimal pedalling rate in prolonged bouts of cycle ergometry. *Med. Sci. Sports Exerc.* **18,** 225–30 (1986).

244. Van Soest, A. J. & Casius, L. J. R. Which factors determine the optimal pedal rate in sprint cycling. *Med. Sci. Sport. Exerc.* **32,** 1927–1934 (1927).

245. Ahlquist, L. E., Bassett, D. R., Sufit, R., Nagle, F. J. & Thomas, D. P. The effect of pedaling frequency on glycogen depletion rates in type I and type II quadriceps muscle fibers during submaximal cycling exercise. *Eur. J. Appl. Physiol. Occup. Physiol.* **65,** 360–364 (1992).

246. Sjoogard, G. in *Int. Ser. Biomech.* (Asmussen, E. & Jorgensen, K.) 93–96 (University Park Press, 1978).

247. Carnevale, T. J. & Gaesser, G. A. Effects of pedaling speed on the power-duration relationship for high-intensity exercise. *Med. Sci. Sports Exerc.* **23,** 242–6 (1991).

248. Hill, D. W., Smith, J. C., Leuschel, J. L., Chasteen, S. D. & Miller, S. A. Effect of pedal cadence on parameters of the hyperbolic power-time relationship. *Int. J. Sports Med.* **16,** 82–7 (1995).

249. Lepers, R., Hausswirth, C., Maffiuletti, N., Brisswalter, J. & van Hoecke, J. Evidence of neuromuscular fatigue after prolonged cycling exercise. *Med. Sci. Sports Exerc.* **32,** 1880–6 (2000).

250. McNaughton, L. & Thomas, D. Effects of differing pedalling speeds on the power-duration relationship of high intensity cycle ergometry. *Int. J. Sports Med.* **17,** 287–92 (1996).

251. Barker, T., Poole, D. C., Noble, M. L. & Barstow, T. J. Human critical power-oxygen uptake relationship at different pedalling frequencies. *Exp. Physiol.* **91,** 621–32 (2006).

252. Green, S., Bishop, D. & Jenkins, D. Effect of end-point cadence on the maximal work-time relationship. *Eur. J. Appl. Physiol. Occup. Physiol.* **71,** 559–561 (1995).

253. Poole, D. C. Letter to the editor in chief. *Med Sci Sport. Exerc.* **18,** 703–704 (1986).

254. Dekerle, J., Williams, C., McGawley, K. & Carter, H. Critical power is not attained at the end of an isokinetic 90-second all-out test in children. *J. Sports Sci.* **27,** 379–85 (2009).

255. Ginn, E. in *Sport Specif. Guidel. Physiol. Assess. Elit. Athl.* (Draper, J. & Telford, R.) 53–66 (Australian Sport Commission, 1988).

256. Hinckson, E. a. & Hopkins, W. G. Reliability of Time to Exhaustion Analyzed with Critical-Power and Log-Log Modeling. *Med. Sci. Sport. Exerc.* **37,** 696–701 (2005).

257. Le Chevalier, J. M., Vandewalle, H., Thépaut-Mathieu, C., Stein, J. F. & Caplan, L. Local critical power is an index of local endurance. *Eur. J. Appl. Physiol.* **81,** 120–7 (2000).

258. Pringle, J. S. M. *et al.* Critical Power And Aerodynamic Drag Accurately Predict Road Time-trial Performance In British Champion Cyclists. *Med. Sci. Sport. Exerc.* **43,** 160–161 (2011).

259. Bishop D, Jenkins D.G, H. A. The Critical Power Function depends on the Duration of the predictive Exercise Test chosen. *Int J Sport. Med.* **19,** 125–129 (1998).

260. Housh, D. J., Housh, T. J. & Bauge, S. M. A methodological consideration for the determination of critical power and anaerobic work capacity. *Res. Q. Exerc. Sport* **61,** 406–9 (1990).

261. Burnley, M., Davison, G. & Baker, J. R. Effects of priming exercise on VO2 kinetics and the power-duration relationship. *Med. Sci. Sports Exerc.* **43,** 2171–9 (2011).

262. Kendall, K. L., Smith, A. E., Fukuda, D. H., Dwyer, T. R. & Stout, J. R. Critical velocity: a predictor of 2000-m rowing ergometer performance in NCAA D1 female collegiate rowers. *J. Sports Sci.* **29,** 945–50 (2011).

263. Bishop, D. & Jenkins, D. G. The influence of recovery duration between periods of exercise on the critical power function. *Eur. J. Appl. Physiol. Occup. Physiol.* **72,** 115–20 (1995).

264. Galbraith, A., Hopker, J. & Lelliott, S. A Single-Visit Field Test of Critical Speed. *Int. J. Sports Physiol. Perform.* (2014).

265. deVries, H. A. *et al.* A method for estimating physical working capacity at the fatigue threshold (PWCFT). *Ergonomics* **30,** 1195–204 (1987).

266. Toubekis, A. G., Tsami, A. P. & Tokmakidis, S. P. Critical velocity and lactate threshold in young swimmers. *Int. J. Sports Med.* **27,** 117–23 (2006).

267. Hiyane, W. C., Simões, H. G., Sílvia, C. & Campbell, G. Critical velocity as a noninvasive method to estimate the lactate minimum velocity on cycling *. *Rev. Bras. Med. do Esporte* **12,** 340–344 (2006).

268. Kranenburg, K. Comparison of critical speed determined from track running and treadmill tests in elite runners. *Med. Sci. Sports Exerc.* **28,** 614–618 (1996).

269. Galbraith. A Novel Field Test to Determine Critical Speed. *J. Sports Med. Doping Stud.* **01,** 1–4 (2011).

270. Dekerle, J., Sidney, M., Hespel, J. M. & Pelayo, P. Validity and reliability of critical speed, critical stroke rate, and anaerobic capacity in relation to front crawl swimming performances. *Int. J. Sports Med.* **23,** 93–8 (2002).

271. Alberty, M. & Sidney, M. Reproducibility of performance in three types of training test in swimming. *Int. J. Sports Med.* **27,** 623–628 (2006).

272. Atkinson, G. & Nevill, A. M. Mathematical constants that may vary? *Med Sci Sport. Exerc.* **37,** 1822 (2005).

273. Smith, J. C. & Hill, D. W. Stability of parameter estimates derived from the power/time relationship. *Can. J. Appl. Physiol.* **18,** 43–7 (1993).

274. Laursen, P. B., Shing, C. M. & Jenkins, D. G. Reproducibility of the Cycling Time to Exhaustion at in Highly Trained Cyclists. (2003). doi:10.1139/h03-046

275. Laursen, P. B., Francis, G. T., Abbiss, C. R., Newton, M. J. & Nosaka, K. Reliability of time-to-exhaustion versus time-trial running tests in runners. *Med. Sci. Sports Exerc.* **39,** 1374–9 (2007).

276. Hampson, D. B., St Clair Gibson, A., Lambert, M. I. & Noakes, T. D. The influence of sensory cues on the perception of exertion during exercise and central regulation of exercise performance. *Sports Med.* **31,** 935–52 (2001).

277. Jeukendrup, A. E. & Currell, K. Should time trial performance be predicted from three serial time-to-exhaustion tests? *Med. Sci. Sports Exerc.* **37,** 1820; author reply 1821 (2005).

278. Atkinson, G. & Nevill, A. M. Statistical Method for Assessing Measurement Error (Reliability) in Variables Relevant to Sports Medicine. *Sport. Med.* **26,** 217–238 (1998).

279. Laursen, P. B., Shing, C. M., Peake, J. M., Coombes, J. S. & Jenkins, D. G. Interval training program optimization in highly trained endurance cyclists. *Med. Sci. Sports Exerc.* **34,** 1801–7 (2002).

280. Bishop, D. An Applied Research Model for the Sport Sciences. *Sport. Med.* **38,** 253–263 (2008).

281. Fregly, B. J., Zajac, F. E. & Dairaghi, C. A. Bicycle drive system dynamics: theory and experimental validation. *J. Biomech. Eng.* **122,** 446–52 (2000).

282. Cole, M. *Encyclopaedia of Social Science Research Methods*. (Sage, 2004).

283. Lucía, A., Hoyos, J., Pérez, M. & Chicharro, J. L. Heart rate and performance parameters in elite cyclists: a longitudinal study. *Med. Sci. Sports Exerc.* **32,** 1777–82 (2000).

284. Lucia, A., Hoyos, J., Santalla, A., Earnest, C. & Chicharro, J. L. Tour de France versus Vuelta a España: which is harder? *Med. Sci. Sports Exerc.* **35,** 872–8 (2003).

285. Fernández-García, B., Pérez-Landaluce, J., Rodríguez-Alonso, M. & Terrados, N. Intensity of exercise during road race pro-cycling competition. *Med. Sci. Sports Exerc.* **32,** 1002–6 (2000).

286. Casa, D. J. *et al.* Influence of hydration on physiological function and performance during trail running in the heat. *J. Athl. Train.* **45,** 147–56

287. Gardner, S., Martin, J. C., Martin, D. T., Barras, M. & Jenkins, D. G. Maximal torque- and power-pedaling rate relationships for elite sprint cyclists in laboratory and field tests. *Eur. J. Appl. Physiol.* **101,** 287–92 (2007).

288. Jobson, S. A. *et al.* The ecological validity of laboratory cycling: Does body size explain the difference between laboratory- and field-based cycling performance? *J. Sports Sci.* **25,** 3–9 (2007).

289. Smith, M. F., Davison, R. C., Balmer, J. & Bird, S. R. Reliability of mean power recorded during indoor and outdoor self-paced 40 km cycling time-trials. *Int. J. Sports Med.* **22,** 270–4 (2001).

290. Padilla, S., Mujika, I., Cuesta, G., Polo, J. M. & Chatard, J. C. Validity of a velodrome test for competitive road cyclists. *Eur. J. Appl. Physiol. Occup. Physiol.* **73,** 446–51 (1996).

291. Jobson, S. A., Nevill, A. M., George, S. R., Jeukendrup, A. E. & Passfield, L. Influence of body position when considering the ecological validity of laboratory time-trial cycling performance. *J. Sports Sci.* **26,** 1269–78 (2008).

292. Bertucci, W., Betik, A. C., Duc, S. & Grappe, F. Gross Efficiency and Cycling Economy Are Higher in the Field as Compared with on an Axiom Stationary Ergometer. *J. Appl. Biomech.* **29,** 636–644 (2012).

293. Arkesteijn, M., Hopker, J., Jobson, S. A. & Passfield, L. The effect of turbo trainer cycling on pedalling technique and cycling efficiency. *Int. J. Sports Med.* **34,** 520–5 (2013).

294. Nimmerichter, A., Williams, C., Bachl, N. & Eston, R. Evaluation of a field test to assess performance in elite cyclists. *Int. J. Sports Med.* **31,** 160–6 (2010).

295. Driller, M. The reliability of a 30-minute performance test on a Lode cycle ergometer. *J. Sci. Cycl.* **1,** 21–27 (2012).

296. Vogt, S. *et al.* Cycling power output produced during flat and mountain stages in the Giro d'Italia: a case study. *J. Sports Sci.* **25,** 1299–305 (2007).

297. Ebert, T. R. *et al.* Power output during women's World Cup road cycle racing. *Eur. J. Appl. Physiol.* **95,** 529–36 (2005).

298. Paton, C. D. & Hopkins, W. G. Seasonal changes in power of competitive cyclists: implications for monitoring performance. *J. Sci. Med. Sport* **8,** 375–81 (2005).

299. Paton, C. D. & Hopkins, W. G. Variation in performance of elite cyclists from race to race. *Eur. J. Sport Sci.* **6,** 25–31 (2006).

300. Lamberts, R. P. *et al.* Measurement error associated with performance testing in well- trained cyclists : Application to the precision of monitoring changes. *Int. Sport. J.* **10,** 33–44 (2009).

301. Hopkins, W. G. A new view on statistics. *Internet Soc. Sport Sci.* (2000). Available at: http://www.sportsci.org/resource/stats. Accessed 12.03.2014

302. Atkinson G., N. A. M. & Edwards, B. What is an acceptable amount of measurement error? The application of meaningful "analytical goals" to the reliability analysis of sports science measurements made on a ratio scale. *J. Sports Sci.* **117,** 8 (1999).

303. Hopkins, W. G. Measures of reliability in sports medicine and science. *Sports Med.* **30,** 1–15 (2000).

304. Vincent, W. J. & Weir, J. *Statistics in Kinesiology-4th Edition*. (Human Kinetics, 2012).

305. Currell, K. & Jeukendrup, A. E. Validity, reliability and sensitivity of measures of sporting performance. *Sports Med.* **38,** 297–316 (2008).

306. Gastin, P. B., Costill, D. L., Lawson, D. L., Krzeminski, K. & McConell, G. K. Accumulated oxygen deficit during supramaximal all-out and constant intensity exercise. *Med. Sci. Sports Exerc.* **27,** 255–63 (1995).

307. Dekerle, J., Barstow, T. J., Regan, L. & Carter, H. The critical power concept in all-out isokinetic exercise. *J. Sci. Med. Sport* 2–6 (2013).

308. Argyris, T. G. & Savvas, T. P. Metabolic responses at various intensities relative to critical swimming velocity. *J. Strength Cond. Res.* (2013).

309. Atkinson, G., Davison, R., Jeukendrup, A. & Passfield, L. Science and cycling: current knowledge and future directions for research. *J. Sports Sci.* **21,** 767–87 (2003).

310. Dahmen, T. Simulation and Optimization of Race-Bike Training on Real Tracks. **1042,** 22–23

311. Dahmen, T., Byshko, R., Saupe, D., Roeder, M. & Mantler, S. Validation of a model and a simulator for raod cycling on real tracks. *Sport. Eng.* **14,** 95–100 (2011).

312. Vanhatalo, A., Doust, J. H. J. & Burnley, M. Determination of critical power using a 3-min all-out cycling test. *Med. Sci. Sports Exerc.* **39,** 548–555 (2007).

313. Dekerle J, Pelayo P, Sidney M, B. G. Challenges of using critical swimming velocity. From Scientists to coaches. *Port. J. Sport Sci.* **6,** 296–299 (2006).

314. Bland, J. M. & Altman, D. G. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* **1,** 307–10 (1986).

315. Hopkins, W. Bias in Bland-Altman but not regression validity analyses. *Sportscience* **8,** 42–48 (2004).

316. Brickley, G., Dekerle, J., Hammond, A. J., Pringle, J. & Carter, H. Assessment of maximal aerobic power and critical power in a single 90-s isokinetic all-out cycling test. *Int. J. Sports Med.* **28,** 414–9 (2007).

317. Gastin, P. B. & Lawson, D. L. Variable resistance all-out test to generate accumulated oxygen deficit and predict anaerobic capacity. *Eur. J. Appl. Physiol. Occup. Physiol.* **69,** 331–336 (1994).

318. Cheng, C.-F., Yang, Y.-S., Lin, H.-M., Lee, C.-L. & Wang, C.-Y. Determination of critical power in trained rowers using a three-minute all-out rowing test. *Eur. J. Appl. Physiol.* **112,** 1251–60 (2012).

319. Francis, J. T., Quinn, T. J., Amann, M. & LaRoche, D. P. Defining intensity domains from the end power of a 3-min all-out cycling test. *Med. Sci. Sports Exerc.* **42,** 1769–75 (2010).

320. Vanhatalo, A., Doust, J. H. & Burnley, M. Robustness of a 3 min all-out cycling test to manipulations of power profile and cadence in humans. *Exp. Physiol.* **93,** 383–90 (2008).

321. Bergstrom, H. & Housh, T. A new single work bout test to estimate critical power and anaerobic work capacity. *J. Strength Cond. Res.* **26,** 656–63 (2012).

322. Blair, S., Painter, P., Pate, R., Smith, K. & Taylor, C. *Resource manual for Guidelines for exercise testing and prescription*. *pp. 414-420* (Lippincott Williams and Wilkins, 1988).

323. Atkinson, G. & Nevill, A. M. Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. *Sports Med.* **26,** 217–38 (1998).

324. Johnson, T. M., Sexton, P. J., Placek, A. M., Murray, S. R. & Pettitt, R. W. Reliability analysis of the 3-min all-out exercise test for cycle ergometry. *Med. Sci. Sports Exerc.* **43,** 2375–80 (2011).

325. Withers, R. T., Ploeg, G. & Finn, J. P. Oxygen deficits incurred during 45, 60, 75 and 90-s maximal cycling on an air-braked ergometer. *Eur. J. Appl. Physiol. Occup. Physiol.* **67,** 185–191 (1993).

326. Jenkins, D. Blood lactate in trained cyclists during cycle ergometry at critical power. *Eur. J. Appl. Physiol.* **61,** 278–283 (1990).

327. Wakayoshi, K. *et al.* Determination and validity of critical velocity as an index of swimming performance in the competitive swimmer. *Eur. J. Appl. Physiol. Occup. Physiol.* **64,** 153–7 (1992).

328. MacDougall, J. D. *et al.* Muscle performance and enzymatic adaptations to sprint interval training. *J. Appl. Physiol.* **84,** 2138–42 (1998).

329. Denadai, B. S. & Denadai, M. L. Effects of caffeine on time to exhaustion in exercise performed below and above the anaerobic threshold. *Braz. J. Med. Biol. Res.* **31,** 581–5 (1998).

330. Wilber, R. L. & Moffatt, R. J. Influence of carbohydrate ingestion on blood glucose and performance in runners. *Int. J. Sport Nutr.* **2,** 317–27 (1992).

331. Krebs, P. S.; Powers, S. K. Reliability of Laboratory Endurance Tests. *Med Sci Sport. Exerc.* **21,** S 10 (1989).

332. Billat, V., Renoux, J. C., Pinoteau, J., Petit, B. & Koralsztein, J. P. Reproducibility of running time to exhaustion at VO2max in subelite runners. *Med. Sci. Sports Exerc.* **26,** 254–7 (1994).

333. Hickey, M. S., Costill, D. L., McConell, G. K., Widrick, J. J. & Tanaka, H. Day to day variation in time trial cycling performance. *Int. J. Sports Med.* **13,** 467–70 (1992).

334. Schabort, E. J., Hawley, J. A., Hopkins, W. G. & Blum, H. High reliability of performance of well-trained rowers on a rowing ergometer. *J. Sports Sci.* **17,** 627–32 (1999).

335. Schabort, E. J., Hopkins, W. G. & Hawley, J. A. Reproducibility of self-paced treadmill performance of trained endurance runners. *Int. J. Sports Med.* **19,** 48–51 (1998).

336. Swain, D. P. A model for optimizing cycling performance by varying power on hills and in wind. *Med. Sci. Sports Exerc.* **29,** 1104–1108 (1997).

337. Zavorsky, G. S. *et al.* Laboratory 20-km cycle time trial reproducibility. *Int. J. Sports Med.* **28,** 743–8 (2007).

338. Sjödin, B. Lactate dehydrogenase in human skeletal muscle. *Acta Physiol. Scand. Suppl.* **436,** 1–32 (1976).

339. Tesch, P., Sjödin, B. & Karlsson, J. Relationship between lactate accumulation, LDH activity, LDH isozyme and fibre type distribution in human skeletal muscle. *Acta Physiol. Scand.* **103,** 40–6 (1978).

340. Jenkins, D., Kretek, K. & Bishop, D. The Duration of Predicting Trials Influences Time to Fatigue at Critical Power. *J. Sci. Med. Sport* **1,** 213–218 (1998).

341. Jenkins, D. G. & Quigley, B. M. The y-intercept of the critical power function as a measure of anaerobic work capacity. *Ergonomics* **34,** 13–22 (1991).

342. Bailey, S. J., Vanhatalo, A., Wilkerson, D. P., Dimenna, F. J. & Jones, A. M. Optimizing the "priming" effect: influence of prior exercise intensity and recovery duration on O2 uptake kinetics and severe-intensity exercise tolerance. *J. Appl. Physiol.* **107,** 1743–56 (2009).

343. Burnley, M., Doust, J. & Jones, A. Time required for the restoration of normal heavy exercise VO2 kinetics following prior heavy exercise. *J. Appl. Physiol.* **101,** 1320–1327 (2006).

344. Weltman, A., Stamford, B. A. & Fulco, C. Recovery from maximal effort exercise: lactate disappearance and subsequent performance. *J Appl Physiol* **47,** 677–682 (1979).

345. Nielsen, O. B., de Paoli, F. & Overgaard, K. Protective effects of lactic acid on force production in rat skeletal muscle. *J. Physiol.* **536,** 161–6 (2001).

346. Pate, E., Bhimani, M., Franks-Skiba, K. & Cooke, R. Reduced effect of pH on skinned rabbit psoas muscle mechanics at high temperatures: implications for fatigue. *J. Physiol.* **486,** 689–694 (1995).

347. Westerblad, H., Bruton, J. D. & Lännergren, J. The effect of intracellular pH on contractile function of intact, single fibres of mouse muscle declines with increasing temperature. *J. Physiol.* **500 (Pt 1,** 193–204 (1997).

348. Wiseman, R. W., Beck, T. W. & Chase, P. B. Effect of intracellular pH on force development depends on temperature in intact skeletal muscle from mouse. *Am J Physiol Cell Physiol* **271,** C878–886 (1996).

349. Wilkerson, D. P., Koppo, K., Barstow, T. J. & Jones, A. M. Effect of prior multiple-sprint exercise on pulmonary O2 uptake kinetics following the onset of perimaximal exercise. *J. Appl. Physiol.* **97,** 1227–36 (2004).

350. Vanhatalo, A. & Jones, A. M. Influence of prior sprint exercise on the parameters of the "all-out critical power test" in men. *Exp. Physiol.* **94,** 255–63 (2009).

351. Burnley, M., Doust, J. H. & Jones, A. M. Effects of Prior Warm-up Regime on Severe-Intensity Cycling Performance. *Med. Sci. Sport. Exerc.* **37,** 838–845 (2005).

352. Burnley, M., Doust, J. H. & Jones, A. M. Effects of prior heavy exercise, prior sprint exercise and passive warming on oxygen uptake kinetics during heavy exercise in humans. *Eur. J. Appl. Physiol.* **87,** 424–32 (2002).

353. Koppo, K. & Bouckaert, J. The decrease in the VO2 slow component induced by prior exercise does not affect the time to exhaustion. *Int. J. Sports Med.* **23,** 262–267 (2002).

354. Bangsbo, J., Madsen, K., Kiens, B. & Richter, E. A. Effect of muscle acidity on muscle metabolism and fatigue during intense exercise in man. *J. Physiol.* **495 (Pt 2,** 587–96 (1996).

355. Vandewalle, H., Kapitaniak, B., Grün, S., Raveneau, S. & Monod, H. Comparison between a 30-s all-out test and a time-work test on a cycle ergometer. *Eur. J. Appl. Physiol. Occup. Physiol.* **58,** 375–81 (1989).

356. Karsten, B., Jobson, S. a, Hopker, J., Jimenez, a & Beedie, C. High Agreement between Laboratory and Field Estimates of Critical Power in Cycling. *Int. J. Sports Med.* **35,** 298–303 (2014).

357. Kyle, C. R. Reduction of Wind Resistance and Power Output of Racing Cyclists and Runners Travelling in Groups. *Ergonomics* **22,** 387–397 (1979).

358. Mognoni, P. & di Prampero, P. E. Gear, inertial work and road slopes as determinants of biomechanics in cycling. *Eur. J. Appl. Physiol.* **90,** 372–6 (2003).

359. DiPrampero, P. E., Cortili, G., Mognoni, P. & Saibene, F. Equation of motion of a cyclist. *J. Appl. Physiol.* **47,** 201–6 (1979).

360. Jobson, S. A., Woodside, J., Passfield, L. & Nevill, A. M. Allometric scaling of uphill cycling performance. *Int. J. Sports Med.* **29,** 753–7 (2008).

361. Racinais, S., Hue, O. & Blonc, S. Time-of-day effects on anaerobic muscular power in a moderately warm environment. *Chronobiol. Int.* **21,** 485–95 (2004).

362. Padilla, S., Mujika, I., Cuesta, G. & Goiriena, J. J. Level ground and uphill cycling ability in professional road cycling. *Med. Sci. Sports Exerc.* **31,** 878–85 (1999).

363. Nimmerichter, A., Eston, R. G., Bachl, N. & Williams, C. Longitudinal monitoring of power output and heart rate profiles in elite cyclists. *J. Sports Sci.* **29,** 831–40 (2011).

364. Clark, I. E., Murray, S. R. & Pettitt, R. W. Alternative procedures for the 3-min all-out exercise test. *J. Strength Cond. Res.* (2012).

365. Constantini, K., Sabapathy, S. & Cross, T. J. A single-session testing protocol to determine critical power and W′. *Eur. J. Appl. Physiol.* (2014).

366. Pettitt, R. W., Jamnick, N. & Clark, I. E. 3-min all-out exercise test for running. *Int. J. Sports Med.* **33,** 426–31 (2012).

367. Takahashi, S., Wakayoshi, K., Hayashi, A., Sakaguchi, Y. & Kitagawa, K. A method for determining critical swimming velocity. *Int. J. Sports Med.* **30,** 119–123 (2009).

368. Toussaint, H., Wakayoshi, K. & Hollander, A. Simulated front crawl swimming performance related to critical speed and critical power. *Med Sci Sport. Exerc.* **30,** 144–151 (1998).

369. Batterham, A. Commentary on: Bias in Bland-Altman but not regression validity analyses. *Sportscience* **8,** 47–49 (2004).

370. Batterham, A. M. & Hopkins, W. G. Making meaningful inferences about magnitudes. *Int. J. Sports Physiol. Perform.* **1,** 50–7 (2006).

371. Zavorsky, G. S. *et al.* Laboratory 20-km cycle time trial reproducibility. *Int. J. Sports Med.* **28,** 743–8 (2007).

372. Bishop, D., Burnett, A., Farrow, D., Gabbett, T. & Newton, R. Sports-science roundtable: does sports-science research influence practice? *Int. J. Sports Physiol. Perform.* **1,** 161–8 (2006).

373. Atkinson, G., Batterham, A. M. & Hopkins, W. G. Sports performance research under the spotlight. *Int. J. Sports Med.* **33,** 949 (2012).

374. Tanaka H, Bassett DR Jr, Swensen TC, S. R. Aerobic and anaerobic power characteristics of competitive cyclists in the United States. *Int. J. Sports Med.* **14,** 334–338 (1993).

375. Craig, N. P. & Norton, K. I. Characteristics of Track Cycling. *Sport. Med.* **31,** 457–468 (2001).

376. Martin, D. T. *et al.* Physiological Characteristics of Nationally Competitive Female Road Cyclists and Demands of Competition. *Sport. Med.* **31,** 469–477 (2001).

377. Medicine, A. C. of S. *ACSM's Guidelines for Exercise Testing and Prescription*. 480 (Lippincott Williams & Wilkins, 2013).

378. Greco, C., Bianco, A., Gomide, E. & Denadai, B. Validity of the critical speed to determine blood lactate response and aerobic performance in swimmers aged 10–15 years. *Sci. Sports* **17,** 306–308 (2002).

379. Whipp, B. J. & Ward, S. A. Quantifying intervention-related improvements in exercise tolerance. *Eur. Respir. J.* **33,** 1254–60 (2009).

380. Barstow, T. J., Jones, A. M., Nguyen, P. H. & Casaburi, R. Influence of muscle fiber type and pedal frequency on oxygen uptake kinetics of heavy exercise. *J. Appl. Physiol.* **81,** 1642–50 (1996).

381. Oxford Dictionaries. at <http://www.oxforddictionaries.com/>

APPENDICES

Appendix I: Example participation information letter

the UNIVERSITY of GREENWICH

MEDWAY SCHOOL OF SCIENCE

SCHOOL OF LIFE AND SPORTS SCIENCE

Title:

Reproducibility and Validity of the 3-min All-Out Cycling Test

Researcher:

Bettina Karsten: Tel: 0208 331 7927, mobile 07974126956, e-mail: kb20@gre.ac.uk

---

PARTICIPANT INFORMATION

A research study is being conducted at Greenwich University by Bettina Karsten (Lecturer in Exercise Physiology). Prof Alfonso Jimenez from the University of Greenwich and Dr Simon Jobson and Dr James Hopker from the University of Kent are co-researchers.

Important

You are free to take part or not in this study. You can withdraw from your participation at any time without any reason given or consequences.

What will be expected of you?

As a participant in this study you will be asked to visit the laboratory for an initial consultation and then be tested on 7 different occasions.

These 7 occasions are divided into one baseline assessment, three 3min All-Out cycling trials (including one 10 sec and one 30 sec all out test) and three trials at a constant work load.

For all of the above sessions you are required to wear your usual cycling clothes and cycling shoes/trainers.

All testing is required to be completed within 3 weeks starting at the endurance fitness assessment.

What does the initial consultation and testing involve?

- Measurement of weight, height, blood pressure
- Fitness assessment for endurance (VO_{2max} test)

Are there any risks?

You will probably experience physical tiredness and discomfort during the exercise tests.

It is unlikely that you will suffer any injuries, although injuries to muscles and ligaments can happen. You will not be expected to continue with the exercise programme if an injury occurs.

What are the benefits to you?

You will receive individual feedback about your results for all tests. These results can be used by yourself and/or your coach for training purposes.

How the results of the study will be used

Your data will be mathematically analysed together with all the other participants' data, and the findings from this analysis will be communicated to other researchers and scientists. Communication of the findings will be in the form of reports in scientific journals, articles in newsletters, and presentation at a conference.

Confidentiality

All data and personal information will be stored securely within University of Greenwich premises in accordance with the terms of the Data Protection Act 1998 and the University's own data protection requirements, and will be accessed only by Bettina Karsten. After completion of the study, all data will be made anonymous (i.e. all personal information associated with your data will be removed). Your data will be anonymous in any written reports, articles, and presentations of the results of the study.

Deciding whether to participate

If you would like to participate, please return the consent form, health history questionnaire to me in the envelope provided. If you have any questions, please contact me on the telephone number or email address above.

Once again, thank you for volunteering!

Version I/ 10/02/2010

Appendix II: Example consent form

UNIVERSITY *of* GREENWICH
RESEARCH ETHICS COMMITTEE
CONSENT FORM

Title of study:_____

- I have read the information sheet about this study
- I have had an opportunity to ask questions and discuss this study
- I have received satisfactory answers to all my questions
- I have received enough information about this study
- I understand that I am / the participant is free to withdraw from this study:
    - At any time (until such date as this will no longer be possible, which I have been told)
    - Without giving a reason for withdrawing
    - (If I am / the participant is, or intends to become, a student at the University of Greenwich) without affecting my / the participant's future with the University
    - Without affecting any medical or nursing care I / the participant may be receiving.
- I understand that my research data may be used for a further project in anonymous form, but I am able to opt out of this if I so wish, by ticking her
- I agree to take part in this study

Signed (participant)	Date
Name in block letters	
Signed (parent / guardian / other) (if under 18)	Date
Name in block letters	
Signature of researcher	Date
This project is supervised by:	
Researcher's contact details (including telephone number and e-mail address):	