Validity of a simulation-based performance testing model in cycling: a study protocol

JASPER SUTTMEYER[1], ANNA DMITRIEVA[1], YVES REUTER[1], JAN VENZKE[1], PETRA PLATEN[1], ROBIN SCHÄFER[1,2]

[1] Department of Sports Medicine and Sports Nutrition, Faculty of Sports Science, Ruhr University Bochum
[2] Division of Physiotherapy, Department of Applied Health Sciences, Hochschule für Gesundheit, Bochum

*** Correspondence:** jasper.suttmeyer@rub.de

Cite as: Suttmeyer, J., Dmitrieva, A., Reuter, Y., Venzke, J., Platen, P., Schäfer, R. (2023). Validity of a simulation-based performance testing model in cycling: a study protocol. *SportRxiv*.

Abstract

Background: Metabolic simulations as described by Mader (2003) can be used to model the physiological response (e.g. blood lactate, phosphocreatine, pH, aerobic and anaerobic energy contribution) to exercise. While some parameters of the model were derived from the literature and are assumed to be constant, the individual performance markers $\dot{V}O_2max$ (maximal oxygen consumption) and ċLamax (maximal lactate production rate) can be used as input to create individual performance predictions. Further, the MLSS (maximal lactate steady-state) can be estimated via the model. In practice, this model is already used to infer those performance markers from observed testing data. However, a thorough evaluation of this approach is still missing.

Objective: To assess the concurrent validity of simulation-based performance testing ($_{sim}$) in cycling compared to experimental estimates from performance testing ($_{exp}$).

Study Design: Agreement Study

Participants: Recreational cyclists and triathletes

Methods: Five exercise tests will be conducted on at least 5 days. On day 1, body composition, a 15-second sprint ($ċLamax_{exp}$) and a ramp test until exhaustion ($\dot{V}O2max_{exp}$) will be conducted. On day 2 and 3, an individualized test protocol and a standard graded exercise test are conducted to observe model-based performance markers ($\dot{V}O_2max_{sim}$, $ċLamax_{sim}$, $MLSS_{sim}$) with each procedure. From day 4 on, multiple 30-minute constant workload tests are performed to measure $MLSS_{exp}$. Agreement analyses will be conducted via Bland-Altman analyses using *a priori* defined Limits of Agreement.

Registration: This study protocol will be pregistered via the Open Science Framework (OSF) upon an ethical vote.

Ethics: A vote by the ethics committee of the Ruhr University Bochum is still pending.

Key Words: maximal oxygen consumption, maximal lactate production rate, anaerobic threshold, endurance testing

Background

Measurement of blood lactate concentration in endurance-related performance testing has been implemented in sport-scientific practice since the early 1980s[1]. The German researchers who laid the foundation of modern lactate measurement were also able to describe a theoretical framework of human energy metabolism using mathematical equations. The resulting model is a differential equation system implemented as a computer simulation. The model is vastly driven by two individual performance parameters: the maximal glycolytic ($\dot{c}Lamax^{a}$) and aerobic ($\dot{V}O_2max$) flux rates. The physiological response to exercise (e.g. blood lactate concentration, phosphocreatine stores) can be determined based on these two parameters, the individual body composition, and further parameters where constant values are assumed based on previous research[3]. Further, the equilibrium of lactate production and removal – the maximal lactate steady-state (MLSS) – can be calculated[4]. Consequently, the simulation model can explain real-life phenomena and give modern lactate testing a theoretical foundation.

Despite that the model has already been implemented in sport-scientific endurance testing the validity of the model for this purpose has not been evaluated sufficiently. Therefore, this study aims to validate the use of simulation-based performance testing in cycling. Experimental ($_{exp}$) derived performance parameters will be compared to those from the simulation-based test ($_{sim}$). $\dot{V}O_2max_{sim}$, $\dot{c}Lamax_{sim}$, and $MLSS_{sim}$ will be derived from a typical graded exercise test protocol as well as an alternative individualized protocol. Those values will be compared with the gold standard measurements of $\dot{V}O_2max_{exp}$, $\dot{c}Lamax_{exp,}$ and $MLSS_{exp}$, respectively.

Methods

Participants

Recreational cyclists and triathletes with a weekly training volume above 6h will be recruited for this study. Sport-specific experience is a prerequisite to participate in the study for three reasons. First, to reduce potential training-specific adaptations due to multiple tests during the time course of the study. Second, to ensure a sufficiently developed endurance performance capability to collect multiple data points at different intensities. Third, to achieve a representative sample to which the results can be generalized.

Recruitment for the study begins in March 2023. Sports science students with sufficient background in endurance sports as well as athletes from local clubs will be recruited for the study. The first examinations will begin in April 2023.

Participants will be informed by the investigators about the goal of the study and potentially arising risks before they participate in this study. Furthermore, written information and a declaration of consent to be signed by the participants are provided. Afterwards, the participants will receive a detailed evaluation of their outcomes. Data of all participants will be anonymized.

Study Design

The study encompasses 5 – 6 testing days with multiple tests on the cycle ergometer. The body composition is measured on the first day using bioelectrical impedance analysis (BIA). Afterward, two performance tests are conducted: a 15-second sprint test to measure $\dot{c}Lamax_{exp}$ followed by a ramp

[a] It has recently been debated that ċLamax is a more appropriate name for the maximal lactate production rate (instead of V̇Lamax). This study protocol acknowledges this recent debate and complies with the newly recommended abbreviation.[2]

test until exhaustion to measure $\dot{V}O_2max_{exp}$. A break of at least 45 minutes between both tests will be ensured. On the second and third day, a step test protocol and an individualized protocol are performed to obtain simulation-based parameters for each test ($\dot{V}O_2max_{sim}$, $\dot{c}Lamax_{sim}$, $MLSS_{sim}$). During the remaining visits to the laboratory, multiple 30-minute constant workload tests are performed to measure $MLSS_{exp}$. The performance testing takes place at two different locations, the Faculty of Sports Sciences (Ruhr University Bochum) and a commercial institution (Dr. Ralf Lindschulten, Hannover).

Pilot studies

The study was designed based on a previously conducted pilot study using a simplified approach[5]. As a main outcome the here proposed study design includes additional tests to assess the validity of the applied model. Further, this study protocol was tested in 3 new participants and consequently refined.

Equipment

Performance testing is conducted on a Cyclus 2 ergometer (TE 2 %, 8 Hz, RBM electronic automation GmbH, Leipzig, Germany). The Cyclus 2 system is considered the gold standard for measuring and controlling mechanical power [6]. Participants are required to bring their own bike which is mounted onto the Cyclus 2. Configurations of the Cyclus 2 are performed according to the specifics of the participant's bike (crank length and size of frontal chain ring).

Blood lactate samples are taken from the earlobe with a 20 µl-capillary tube mixed in a reaction vessel. Measurement of lactate concentration is performed enzymatic-amperometrically using chip-sensor technology (Biosen C-Line Sport, EKF-Diagnostik Germany). Gas exchange is measured using a spirometric device (MetaMax 3B, Cortex Germany). All systems are calibrated before the beginning of each performance test.

Body composition is measured using BIA (InBody 770, JP Global Markets GmbH, Germany).

Simulation

The simulation of the energy metabolism from Mader has been developed through the course of multiple publications [3,4,7]. Calculations in this study are based on Mader (2003)[3] and summarized in the following:

1. *Calculation of phosphates concentrations and free enthalpy*

$$[ADP] = \frac{S[A]}{\left(M_3 \cdot \frac{(S[C] - [PCr])}{M_2} \cdot [H^+] \cdot [PCr]\right) + \left(M_2 \cdot [H^+] \cdot \frac{[PCR]}{S[C] - [PCr]}\right) + 1}$$

$$[ATP] = M_2 \cdot [H^+] \cdot [PCr] \cdot \frac{[ADP]}{S[C] - [PCr]}$$

$$[AMP] = M_3 \cdot \frac{[ADP]^2}{[ATP]}$$

$$\Delta G_{ATP.cyt} = \Delta G_0 - R \cdot T \cdot ln(M_2 \cdot [H^+] \cdot \frac{[PCR]}{(S[C] - [PCr])^2})$$

2. Activation curves

$$\dot{V}O_{2.a} = \frac{\dot{V}O_{2.max}}{1 + \frac{ks_1}{[ADP] \cdot \Delta\Delta G_{ADP}}}$$

$$vLa_{ss.ph} = \frac{1}{1 + \frac{[H^+]^3}{ks_3}} \cdot \frac{\dot{c}La_{max}}{1 + \frac{ks_2}{[ADP] \cdot [AMP]}}$$

3. Differential equation system

$$\frac{d[GP]}{dt} = b_{VO2} \cdot \dot{V}O_{2.a} + b_{vLa} \cdot vLa_{ss.ph} - b_{pow} \cdot E$$

$$\frac{d[VO2]}{dt} = k_{VO2} \cdot (\dot{V}O_{2.ss} - \dot{V}O_{2.a})$$

$$\frac{d[La]_m}{dt} = -K_1 \cdot ([La]_m - [La]_b) + \frac{1}{La_{space.muscle}} \cdot (vLa_{ss.ph} - vLa_{ox.m})$$

$$\frac{d[La]_b}{dt} = V_{rel} \cdot (K_1 \cdot ([La]_m - [La]_b) - vLa_{ox.b})$$

Table 1: Constants used in the computer simulation [3]

Constant	Value	Unit	Description
M3	1.05	[mol^{-1}]	Adenylate kinase equilibrium constant
M2	1.66 x 10^9	[mol^{-1}]	Hydrogen ion independent creatine kinase constant
S[A]	7	[$mmol/kg_m$]	Sum of adenylate phosphates
ks_1	0.035	[$mmol/kg_m$]	50%-activation constant of oxidative phosphorylation related to $\dot{V}O_2max$
ks_2	0.00035	[$mmol/kg_m$]	50%-activation constant of glycolysis related to $\dot{c}$Lamax
ks_3	10^{-20}	[mol^{-1}]	50%-inhibition rate constant of glycolysis due to non-competitive inhibition of $[H]^+$
ΔG_0	-30.6	[kJ/mol]	Standard free energy of the [ATP]/[PCr] system
R	2 x 10^{-3} x 4.184	[kJ/mol]	Gas constant
T	310	[K]	Absolute Temperature
$La_{space.muscle}$	0.75	AU	Average relative muscle H_2O space

The simulation was implemented in the programming language *Python*. The differential equation system is solved with the package *scipy.integrate* with the function *solve_ivp* and the method *RK45* [8]. Open-source software for the simulation will be published separately.

Testing Protocol

Day 1 – Body composition, sprint and ramp test

First, the body composition is measured using BIA. The collected data is necessary to properly calibrate the simulation for the individual participant. It is necessary to approximate the lactate distribution space (La_{space}) and the active muscle mass (AMM). La_{space} is considered to be 69% of the total body water[4]. As lactate is measured with the dimension *concentration* the overall distribution space is needed to calculate the amount of substance. AMM is a similarly important parameter because the simulation calculates the metabolic behavior on the level of one-kilogram muscle and on this basis the exchange of metabolites to the whole body. Consequently, the maximal flux rates and the ATP-turnover must be scaled towards one kilogram of AMM. AMM is considered to be 65% of the overall muscle mass for cycling.

Sprint test

Second, the sprint test to approximate ċLamax is performed on the Cyclus 2. Based on current scientific knowledge the sprint test is highly reliable [9]. Nevertheless, the validity of the test to determine ċLamax has yet to be demonstrated. The estimation of the anaerobic-alactic contribution to the ATP-synthesis, the inhibitory effect of the intramuscular pH-value on glycolysis, the ongoing activation of glycolysis after the termination of the sprint and the potential influence of the oxidative metabolism on ATP-synthesis and lactate clearance challenge the validity of the test. Recommended is a short maximal sprint (15 seconds), after which multiple lactate samples are taken. From the highest lactate concentration after the sprint the ċLamax is approximated.

The sprint test is conducted based on the defined standards[9]. The participant's own bike is mounted on the Cyclus 2. Before the execution of the sprint, a standardized ten-minute warm-up at low mechanical power is performed. During the warm-up, three linear increments up to 500 W for ten seconds are performed. After the termination of the warm-up, a passive phase of five minutes follows in which the participant is instructed to move as little as possible. Before the sprint follows three capillary blood samples are drawn from the earlobe ($[La]_{pre}$). The following all-out sprint is conducted in the isokinetic maximal force mode of the Cyclus 2. The initial force is set to 2 N x kg^{-1} and the maximal cadence is regulated to not exceed values above 130 U x min^{-1} [9,10]. Throughout the test the participant is required to remain seated, may not change gearing, and is verbally motivated by the investigator. After the termination of the sprint blood lactate samples are drawn every minute for the following ten minutes during which the participant is instructed to move as little as possible.

Ramp test

$\dot{V}O_2max$ is measured using a ramp protocol. The test starts with a warm-up protocol at a low intensity. After the warm-up, the ramp is started to reach exhaustion within 7 – 10 minutes[11]. The initial mechanical power is set to 50 W and increases 1 W per second. The test is terminated when volitional exhaustion is attained, or cadence drops below 40 rpm. It is the goal to maximally activate the oxidative phosphorylation before fatigue from other systems exhausts the participant. Exhaustion will be evaluated using the following criteria:

1.) Levelling off in the measured oxygen consumption. Defined as an increase below 150 ml/min during the last 30 seconds of the test.
2.) Respiratory exchange ratio is above 1.1 at the end of the test.
3.) Maximal lactate concentrations after the test exceed 10 mmol/L.
4.) Rating of perceived exertion is above 19 (BORG-Scale).

During the test, the oxygen consumption and carbon dioxide production are continually measured using spirometry. Collected data are evaluated using the Breath-by-Breath-Method. $\dot{V}O_2max$ is defined as the highest value of a 30-second rolling average.

Day 2 – individualized test protocol

In practice, individualized test protocols are commonly used for simulation-based performance testing. Multiple constant intervals with varying duration and intensity are conducted. Based on anecdotal evidence the individualized test protocol is thought to be superior and deliver more accurate results in collaboration with the simulation than the classical step test protocol. Theoretical rationales supporting this claim are outlined parallel to the following description of the protocol. A scientific comparison of both methods has never been performed to the author's knowledge.

In this study, the individualized test protocol consists of four intervals which can be distinguished in 3 phases. The interval lengths are the same for all participants and the mechanical power for each participant will be estimated based on the individual endurance performance and physiological data throughout the test. Spirometric data, lactate measurement and rate of perceived exhaustion (BORG, 6-20) will be collected throughout the test.

Phase 1

The test is preceded by the measurement of the basal metabolic rate via a spirometric cart. Phase 1 is structured like a classical graded exercise test (3-minute-stages) with the aim of obtaining a six-minute interval with constant intensity with a post lactate concentration between 2.5 – 3 mmol/L. A starting intensity is chosen by the investigator. After two minutes a first blood sample is drawn and directly analyzed. If the lactate concentration is below 1.8 mmol/L another three-minute stage is entered with increased intensity. This scheme continues until the lactate concentration exceeds 1.8 mmol/L. Once this has occurred the intensity is kept constant, and the duration of the stage is extended to six minutes. After the end of the six-minute-stage, capillary blood samples are drawn at rest at minutes 0' and 1' and afterward in two-minute intervals until the lactate concentration decreases.

The structure of the first phase is justified by simulation results. Two hypothetical athletes who only differ in $\dot{V}O_2max$ were simulated to conduct six-minute trials of varying intensities - see Figure 2. The post-interval lactate values are shown on the y-axis. The results show low discrimination between both athletes in regions of low lactate. Thus, lactate concentration well below 2.5 mmol/L may not be suitable to derive information about underlying performance parameters.

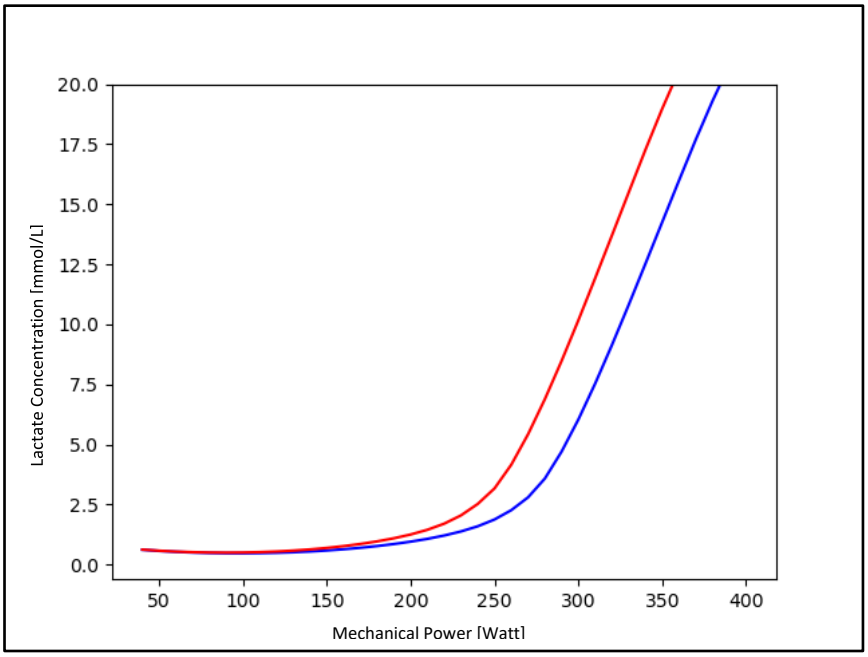

Fig. 2: Lactate concentration as a function of mechanical power for a six-minute interval. Red: $\dot{V}O_2max$ = 4125 ml/min (55 ml/min/kg), blue: 4500 ml/min (60 ml/min/kg). All other parameters are identical.

Phase 2

Phase 2 consists of 2 intervals which aim to reach post lactate concentrations between 3.5 – 4.5 mmol/L and 4.5 – 6 mmol/L. Hence, intensity will be subjectively estimated by the investigators. Interval duration is 6 minutes. A longer duration than in the classical step test protocol was chosen because a longer duration results in higher lactate concentrations, which leads to better estimation of the physiological parameters – see Figure 2. Both intervals are preceded with active recovery until lactate concentration decreases below 3.5 mmol/L (at least 2 minutes). After each interval, blood samples are drawn at minutes 0', 1', and afterward every two minutes until the lactate concentration starts to decrease and thus ensuring that the maximal concentration has been exceeded. After the passive rest of at least 2 minutes, the next interval started with active recovery.

Phase 3

Phase 3 consists of one interval which aims to elicit exhaustion within 3 – 5 minutes. A duration of at least three minutes is necessary to reach oxygen consumption values near $\dot{V}O_2max$ [12]. To increase test economy, the maximal duration is aimed to be 5 minutes. The interval is preceded with active recovery until lactate concentration decreases below 3.5 mmol/L. The subjective assessment of the athlete and previous test data are used to determine the individual intensity. Blood samples are drawn as described in phase 2.

Generally, it is to be stressed that the attainment of lactate concentrations exactly within the predefined confines is unlikely and not necessary to fulfill the requirements of the protocol. One should rather think of this as a rough orientation to obtain data points that are evenly distributed over the entire metabolic profile.

Day 3 – Graded exercise test

Due to the lack of a study demonstrating the superiority of the individualized test protocol in comparison to a classical step test both protocols are implemented in this study. The graded exercise test starts with a mechanical power of 80 W and increases by 40 W every five minutes. At the end of

each stage, a capillary blood sample is drawn, heart rate is measured, and the ratio of perceived exertion (BORG-Scale) is recorded. The test is terminated at the exhaustion of the participant. Afterwards, three capillary blood samples at 0', 1', and 3' are drawn.

Day 4-6 – MLSS Trials

The MLSS is defined as the intensity at which the lactate concentration in the blood during a 30-minute test with constant mechanical power does not rise further than 1 mmol/L between the 10th and 30th minute of the test ($\Delta[La]_{MLSS}$)[13]. Concerning the practical validation of the calculated MLSS based on the simulated performance parameters at least two MLSS-Trials will be performed. Before the test, a 14-minute warm-up is performed at low intensity including two two-minute intervals closely below the intensity of the 30-minute test. Throughout the entire constant load test the participants' gas exchange and heart rate are measured. Capillary blood samples are drawn at intervals of five minutes and after the test at 0', 1', and 3'.

Analysis of performance parameters

ċLamax$_{exp}$

The arithmetic mean of the three lactate concentrations drawn before the sprint is considered $[La]_{pre}$ (mmol/L). Lactate concentrations after the sprint test are interpolated using the Bateman function[9]. The maximal value of the interpolated function is considered $[La]_{postmax}$ (mmol/L). t_{alac} (s) is calculated using the following formula:

$$t_{alac} = 0.0909 \times t_{sprint} + 2.045$$

Finally, ċLamax$_{exp}$ (mmol/L/s) is calculated:

$$\dot{c}\mathrm{Lamax}_{exp} = \frac{[La]_{postmax} - [La]_{pre}}{t_{sprint} - t_{alac}}$$

V̇O$_2$max$_{sim}$, ċLamax$_{sim}$ and MLSS$_{sim}$

Simulation-based performance parameters were derived from the individualized and the classic test protocol. For both tests, individual body composition and the equations and constants as outlined in the section "Simulation" were used. The parameters V̇O$_2$max and ċLamax as input were perturbed to find the best combination of parameters that minimizes the residuals of simulated versus observed lactate values. To do so, non-linear least squares estimation via the Levenberg-Marquart algorithm[14] as implemented in the lmfit[15] function from the scipy package[8] in the programming language Python was used. Initial values for fitting are derived from the respective test and the fit will be constrained to physiological plausible values. For the individualized protocol, the residuals were calculated based on a lactate-power curve, where the lines represent the lactate concentration after the time spent in the trial as function of power (see Figure 3). For the classical test protocol, the residuals are calculated based on a single lactate-time curve.

In order to illustrate the analysis, data of a participant from a pilot study is shown. The dots in figure 3 show the maximal lactate concentration after each interval. Colored lines illustrate the simulated lactate concentration as a function of mechanical power. After varying V̇O$_2$max $_{sim}$ and ċLamax$_{sim}$, the best fit was obtained by $\chi 2$ and the experimentally obtained V̇O$_2$peak. The simulated performance parameters for this participant show 65 ml/min/kg (V̇O$_2$max$_{sim}$), 0.70 mmol/L/s (ċLamax$_{sim}$), and 255 Watts (MLSS$_{sim}$). $\chi 2$ is 3.45. The analysis of the step test protocol is similar. To illustrate this process,

data of the same participant are visualized in figure 4. Calculated performance parameters are 64.0 ml/min/kg ($\dot{V}O_2max_{sim}$), 0.60 mmol/L/s ($\dot{c}Lamax_{sim}$), and 254 Watts ($MLSS_{sim}$).

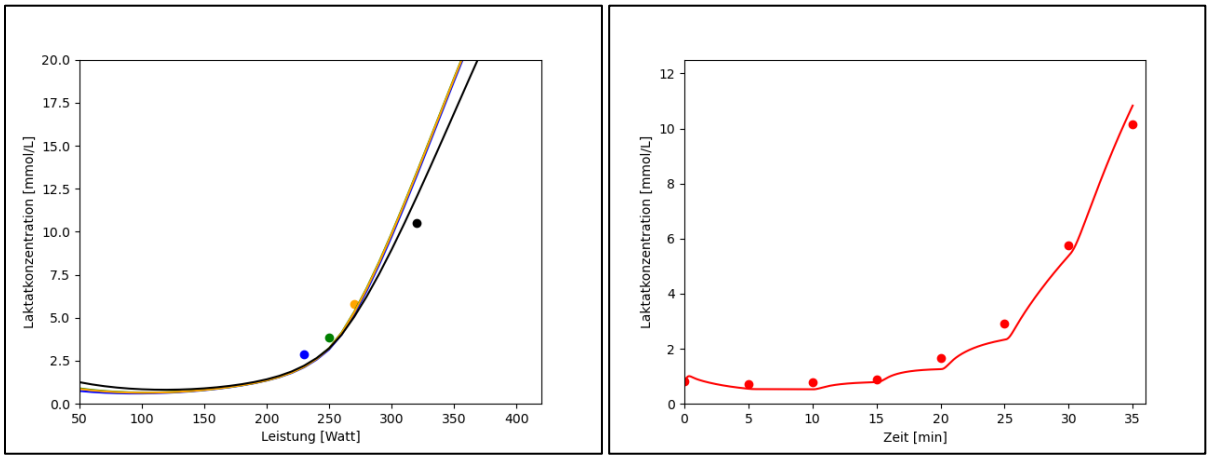

Fig. 3: Exemplary analysis of the individual test protocol (left) and the graded exercise test protocol (right). Left: Blue, green and yellow show the six-minute intervals; black is the third phase. Dots illustrate measured values. Lines show the outcomes of the simulation. Right: Dots show measured values. The line illustrates the outcome of the simulation.

$MLSS_{exp}$

In previous studies the MLSS was only determined to an accuracy of 10 W[16]. It is unlikely to truly measure the MLSS during a 30-minute interval due to the precision of measurement (lactate concentration and power control) and the sensitivity of the lactate kinetics at intensities around the MLSS. Nevertheless, rounding the power at the MLSS to 10 W to achieve a more practical measurement interferes with the validation of the calculated MLSS from the simulation. $MLSS_{sim}$ is calculated on an absolute scale, whereas rounded $MLSS_{exp}$ is ordinally scaled. In order to calculate $MLSS_{exp}$ on an absolute scale $MLSS_{exp}$ is linearly interpolated based on the two 30-minute constant load tests closest to MLSS. The following explains the mathematical approach:

$$MLSS_{exp} = (1 - \Delta[\mathrm{La}]_{MLSS.low}) \div \frac{\Delta[\mathrm{La}]_{MLSS.high} - \Delta[\mathrm{La}]_{MLSS.low}}{p1_{trial} - p2_{trial}} + p1_{trial}$$

1. MLSS Trial ($p1_{trial}$ = 218 W)
   $\Delta[La]_{MLSS}$ = 0.6 mmol/L

2. MLSS-Trial ($p2_{trial}$ 228 W)
   $\Delta[La]_{MLSS}$ = 1.7 mmol/L

$$MLSS_{exp} = (1 - 0.6) \div \frac{1.7 - 0.6}{228 - 218} + 218 = 221.6\,W$$

The intensity of the first MLSS-Trial is based on the $MLSS_{sim}$. If $MLSS_{sim}$ from the individualized and the graded exercise test differs the average is calculated and defined as the intensity for the first MLSS-trial. If the blood lactate concentration rises further than 1 mmol/L during the last 20 minutes of the test the mechanical power is increased for the next trial and vice versa. The test battery is terminated, if one MLSS trial above and below the MLSS has been performed. $MLSS_{exp}$ is consequently calculated based on the two MLSS-trials closest to MLSS.

Statistical Analysis

Agreement analyses via Bland-Altman analyses[17] and the concordance correlation coefficient (CCC)[18] will be conducted for MLSS, $\dot{V}O_2max$ and ċLamax by comparing their experimental versus simulation-derived versions. The primary outcome is the MLSS. Based on our power analysis, a hypothesis test for agreement on an *a priori* delta value of ± 20W will be conducted, i.e., confidence intervals (1-2*alpha) of the Limits of Agreement (LoA) not exceeding this value yield a significant test result. Significance was set at alpha = 0.05. Confidence intervals around point estimates are calculated for the biases, LoAs and CCCs. Estimation error of the parameters $\dot{V}O_2max$ and ċLamax will be derived by the outlined procedure and evaluated in each protocol.

Further parameters are investigated in adjunct qualification work:

Master-Thesis: $\dot{V}O_2max_{exp}$ and $\dot{V}O_2max_{sim}$ from the classical step test are compared. For the sake of validity, a Bland-Altman-Comparison is conducted. The limits of agreement are defined *a priori* as ± 300 ml/min based on the results of a previously published paper allegedly using a similar method and the reliability of $\dot{V}O_2max$ measurements[19,20].

Master-Thesis: $\dot{V}O_2max_{exp}$ and $\dot{V}O_2max_{sim}$ from the lactate over time curve from the MLSS-Trials are compared. Validity is assessed by performing a Bland-Altman-Analysis. The limits of agreement are defined *a priori* as ± 300 ml/min based on the reliability of $\dot{V}O_2max$ measurements[19,20].

Bachelor-Thesis: $ċLamax_{exp}$ and $ċLamax_{calc}$ based on the maximal lactate concentration after the ramp test are compared [7]. In order to assess the validity a Bland-Altmann-Analysis is performed. Based on the reliability of $ċLamax_{exp}$ measurements *a priori* limits of agreement are defined as ± 0.15 mmol/L/s [9].

An overview of the planned analyses is provided in table 2.

Power analysis

A power analysis was conducted for sample sizes up to 25 participants which was in the range of our personal resources for this study. Primary outcome is the difference of $MLSS_{sim}$ to the practically determined $MLSS_{exp}$ (gold standard). A previously published study involved 19 participants to measure agreement between several methods compared to the gold standard[16]. This sample size was high compared to other studies evaluating agreement compared to the MLSS. Our power calculations are based on a hypothesis test, where the confidence intervals (1- alpha/2) of the Limits of Agreement (LoA) lie between a pre-defined threshold (delta) according to[21]. A mean difference (bias) of 1 W and a standard deviation of differences of 7 W was assumed based on a previously published paper[16]. Setting alpha = 0.05, agreement level = 0.9 and delta = 20 W for cycling, a sample size of n = 20 yields a power of 79.4 %. The calculation was performed in R (R Core Team, 2021) using the SimplyAgree-package[22]. Additionally, a sample size calculation for correlation was performed. Assuming a high correlation between measurements of Intra-Class-Correlation = 0.95 aiming for a precision of the 95 % confidence interval of ± 0.05 yields a sample size of n = 16 using the web-based sample size calculator[23]. We concluded that a sample size of n = 20 for cycling will be sufficiently informative to draw conclusions on agreement between measurements.

Table 2: Overview of research questions

Question	Hypothesis	Sampling plan	Analysis Plan	Rationale for deciding the sensitivity of the test for confirming or disconfirming the hypothesis	Interpretation given different outcomes	Theory that could be shown wrong by the outcomes
Can a simulation based performance testing accurately calculate the power at the maximal lactate steady state (MLSS)?	The MLSS is accurately estimated by simulation performance testing.	A power analysis was performed based on the mean bias and standard deviation of another study using a different approach to predict the same variable. It was concluded the n = 20 will be sufficiently informative to draw a conclusion.	1.) Bioelectrical Impedance Analysis 2.) Classical Step Test (Simulation Input) 3.) Individualized Step Test (Simulation Input) 4.) Multiple Constant Load Trials (MLSS)	Previously published studies and reliability of gold standard measurements.	Limits of agreement outside the a priori defined limits.	Use of metabolic simulations in performance testing settings.
Can a simulation based performance test using a classical step test accurately calculate $\dot{V}O_2max$?	The $\dot{V}O_2max$ is accurately estimated by simulation performance testing using a classical step test.	Sampling plan was chosen according to the primary outcome parameter of the study (MLSS).	1.) Bioelectrical Impedance Analysis 2.) Classical Step Test (Simulation Input) 3.) Ramp Test ($\dot{V}O_2max$)	Previously published studies and reliability of gold standard measurements.	Limits of agreement outside the a priori defined limits.	Use of metabolic simulations in performance testing settings.
Can a simulation based performance test using multiple constant load trials accurately calculate $\dot{V}O_2max$?	The $\dot{V}O_2max$ is accurately estimated by simulation performance testing multiple constant load trials.	Sampling plan was chosen according to the primary outcome parameter of the study (MLSS).	1.) Bioelectrical Impedance Analysis 2.) Multiple Constant Load Trials (Model Input) 3.) Ramp Test ($\dot{V}O_2max$)	Reliability of gold standard measurements.	Limits of agreement outside the a priori defined limits.	Use of metabolic simulations in performance testing settings.
Can a simulation based performance test using the post lactate concentration kinetics from a ramp test accurately calculate ċLamax?	The ċLamax is accurately estimated by simulation performance testing from the post lactate kinetics of a ramp test.	Sampling plan was chosen according to the primary outcome parameter of the study (MLSS).	1.) Bioelectrical Impedance Analysis 2.) Ramp Test (Model Input) 3.) Sprint Test (ċLamax)	Reliability of gold standard measurements.	Limits of agreement outside the a priori defined limits.	Use of metabolic simulations in performance testing settings.

Prospect

In addition to the conducted analysis further analyses will be performed in future studies. A rough outline of these analyses can be formulated as follows:

1.) Comparison between simulated performance parameters from individualized and classical graded exercise test protocols.
2.) Simulation of a ramp test and calculation of performance parameters based on the post-exercise lactate kinetics.
3.) Simulation of a sprint based on the power-time curve converted into ATP-turnover as an input parameter to calculate ċLamax based on the post-exercise lactate kinetics.
4.) Calculation of ċLamax based on the ATP-Turnover without measuring lactate.

Prospective analyses of the collected data are not included in table 2.

Data availability

Data and code will be publicly available in our OSF repository: https://osf.io/6xf4n/

Contributions

Conceptualization JS, RS, Methodology JS, RS, AD, YR, Statistical Analysis JS, RS, Visualization JS, Software JS, RS, Investigation JS, AD, YR, Writing – Original Draft JS, Writing – Review & Editing Everyone, Supervision RS, JV, Resources PP

Funding

No external funding was received for this study.

Conflict of interest

The authors declare no conflict of interest.

Literature

1. Wackerhage H, Gehlert S, Schulz H, Weber S, Ring-Dimitriou S, Heine O. Lactate Thresholds and the Simulation of Human Energy Metabolism: Contributions by the Cologne Sports Medicine Group in the 1970s and 1980s. *Front Physiol*. 2022;13:899670. doi:10.3389/fphys.2022.899670

2. Nolte S, Quittmann O, Meden V. Simulation of Steady-State Energy Metabolism in Cycling and Running. *SportRxiv*. Published online 2022.

3. Mader A. Glycolysis and oxidative phosphorylation as a function of cytosolic phosphorylation state and power output of the muscle cell. *Eur J Appl Physiol*. 2003;88(4):317-338. doi:10.1007/s00421-002-0676-3

4. Mader A, Heck H. A Theory of the Metabolic Origin of "Anaerobic Threshold." *Int J Sports Med*. 1986;07(S 1):S45-S65. doi:10.1055/s-2008-1025802

5. Suttmeyer J, Theodoropoulos M, Guillaume A, Schäfer R. Estimation of endurance performance markers using a metabolic model in cycling: a pilot study. *SportRxiv*. Published online March 7, 2023. doi:10.51224/SRXIV.269

6. Granier C, Hausswirth C, Dorel S, Le Meur Y. Validity and Reliability of the Stages Cycling Power Meter. *Journal of Strength and Conditioning Research*. 2020;34(12):3554-3559. doi:10.1519/JSC.0000000000002189

7. Heck H, Bartmus U, Grabow V. *Laktat: Stoffwechselgrundlagen, Leistungsdiagnostik, Trainingssteuerung*. Springer; 2022.

8. Virtanen P, Gommers R, Oliphant TE, et al. SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat Methods*. 2020;17(3):261-272. doi:10.1038/s41592-019-0686-2

9. Quittmann OJ, Schwarz YM, Mester J, Foitschik T, Abel T, Strüder HK. Maximal Lactate Accumulation Rate in All-out Exercise Differs between Cycling and Running. *Int J Sports Med*. 2021;42(04):314-322. doi:10.1055/a-1273-7589

10. Adam J, Oehmichen M, Oehmichen E, et al. Reliability of the calculated maximal lactate steady state in amateur cyclists. *Biol Sport*. 2014;32(2):97-102. doi:10.5604/20831862.1134311

11. McConnell TR. Practical Considerations in the Testing of VO2 max in Runners: *Sports Medicine*. 1988;5(1):57-68. doi:10.2165/00007256-198805010-00005

12. di Prampero P, Osgnach C. Metabolic Power in Team Sports - Part 1: An Update. *Int J Sports Med*. 2018;39(08):581-587. doi:10.1055/a-0592-7660

13. Heck H. *Laktat in der Leistungsdiagnostik*. Hofmann; 1990.

14. Marquardt DW. An algorithm for least-squares estimation of nonlinear parameters. *Journal of the society for Industrial and Applied Mathematics*. 1963;11(2):431-441. doi:10.1137/0111030

15. Newville M, Stensitzki T, Allen DB, Ingargiola A. LMFIT: Non-Linear Least-Square Minimization and Curve-Fitting for Python. Published online September 21, 2014. doi:10.5281/ZENODO.11813

16. Wahl P, Manunzio C, Vogt F, et al. Accuracy of a Modified Lactate Minimum Test and Reverse Lactate Threshold Test to Determine Maximal Lactate Steady State. *Journal of Strength and Conditioning Research*. 2017;31(12):3489-3496. doi:10.1519/JSC.0000000000001770

17. Giavarina D. Understanding Bland Altman analysis. *Biochem Med*. 2015;25(2):141-151. doi:10.11613/BM.2015.015

18. Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *Journal of Chiropractic Medicine*. 2016;15(2):155-163. doi:10.1016/j.jcm.2016.02.012

19. Podlogar T, Cirnski S, Bokal Š, Kogoj T. Utility of INSCYD athletic performance software to determine Maximal Lactate Steady State and Maximal Oxygen Uptake in cyclists. *J Sci Cycling*. 2022;11(1):30-38. doi:10.28985/1322.jsc.06

20. Vickers R. MEASUREMENT ERROR IN MAXIMAL OXYGEN UPTAKE TESTS. Published online 2003.

21. Lu MJ, Zhong WH, Liu YX, Miao HZ, Li YC, Ji MH. Sample Size for Assessing Agreement between Two Methods of Measurement by Bland−Altman Method. *The International Journal of Biostatistics*. 2016;12(2). doi:10.1515/ijb-2015-0039

22. Caldwell AR. SimplyAgree: An R package and jamovi Module forSimplifying Agreement and Reliability Analyses. *JOSS*. 2022;7(71):4148. doi:10.21105/joss.04148

23. Arifin W. Sample Size Calculator (web). Accessed January 17, 2023. https://wnarifin.github.io/ssc_web.html