A Novel Method to Predict Carbohydrate and Energy Expenditure during Endurance Exercise Using Measures of Training Load

Authors: Jeffrey A. Rothschild[1,2*], Stuart Hofmeyr[2], Shaun J. McLaren[3,4], and Ed Maunder[2]

Affiliations:
[1] High Performance Sport New Zealand (HPSNZ), Auckland, New Zealand
[2] Sports Performance Research Institute New Zealand (SPRINZ), Auckland University of Technology, Auckland, New Zealand
[3] Newcastle Falcons Rugby Club, Newcastle upon Tyne, United Kingdom
[4] Department of Sport and Exercise Sciences, Manchester Metropolitan University Institute of Sport, Manchester, UK

*Corresponding author - Jeffrey Rothschild – Jeffrey.Rothschild@aut.ac.nz

Abstract

Background: Sports nutrition guidelines recommend carbohydrate (CHO) intake be individualized to the athlete and modulated according to changes in training load (TL). However, there are limited methods to assess CHO utilization during training sessions.

Objectives: To 1) quantify bivariate relationships between both CHO and overall energy expenditure (EE) during exercise and commonly-used, non-invasive measures of TL across sessions of varying duration and intensity, and 2) build and evaluate prediction models to estimate CHO utilization and EE with the same TL measures and easily-quantified individual factors.

Methods: This study was undertaken in two parts — a primary study, where participants performed four different laboratory-based cycle training sessions, and a validation study where different participants performed a single laboratory-based training session using one of three exercise modalities (cycling, running, or kayaking). The primary study included 15 cyclists (5 f; $\dot{V}O_2$max, 52 ± 7 mL·kg^{-1}·min^{-1}), the validation study included 21 cyclists (7 f; $\dot{V}O_2$max 53.5 ± 11.0 mL·kg^{-1}·min^{-1}), 20 runners (6 f; $\dot{V}O_2$max 57.5 ± 7.2 mL·kg^{-1}·min^{-1}), and 17 kayakers (4 f; $\dot{V}O_2$max 46.2 ± 4.1 mL·kg^{-1}·min^{-1}). Training sessions were quantified using six TL metrics: two using heart rate, three using power, and one using perceived exertion. CHO use and EE were determined separately for aerobic (gas exchange) and anaerobic (net lactate accumulation, body mass, and O_2 lactate equivalent method) energy systems and summed. Repeated-measures correlations were used to examine relationships between TL and both CHO utilization and EE. General estimating equations were used to model CHO utilization and EE, using TL alongside measures of fitness and sex. Models were built in the primary study and tested in the validation study. Model performance is reported as the coefficient of determination (R^2) and mean absolute error (MAE), with measures of calibration used for model evaluation and for sport-specific model re-calibration.

Results: Very-large to near-perfect within-subject correlations (r = 0.76–0.96) were evident between all TL metrics and both CHO utilization and EE. In the primary study, all models explained a large amount of variance (R^2 = 0.77–0.96) and displayed good accuracy (MAE; CHO = 16–21 g [10–14%], EE = 53–82 kcal [7–11%]). In the validation study the MAE ranged from 17–50 g [15–45%] for CHO models and 53–178 kcal [9–30%] for EE models. The calibrated MAE ranged from 8–20 g [7–18%] for CHO models and 36–72 kcal [6–12%] for EE models.

Conclusion: At the individual level, there are strong linear relationships between all measures of TL and both CHO utilization and EE during cycling. When combined with other measures of

fitness, EE and CHO utilization during cycling can be estimated accurately. These models can be applied in running and kayaking when used with a calibration adjustment.

Key Points

Sports nutrition guidelines recommend carbohydrate intake be individualized to the athlete and modulated according to changes in training load, but there are limited methods to assess CHO utilization during training sessions.

We examined bivariate relationships between both carbohydrate and energy expenditure during exercise and commonly used measures of training load, and then built prediction models to estimate carbohydrate and energy expenditure during exercise using training load and easily obtained measures of fitness.

We observed very-large to near-perfect within-subject correlations between both carbohydrate and energy expenditure and all measures of training load. Prediction models displayed good accuracy in cycling, and have potential application in running or kayaking when used with a calibration adjustment.

Introduction

Contemporary sports nutrition guidelines recommend carbohydrate intake be individualized to the athlete and their event, and modulated according to changes in exercise volume [1]. However, there are limited methods of assessing carbohydrate utilization during a given workout, leaving athletes and practitioners unclear as to how much carbohydrate or energy should be repleted. Indeed, there have been recent calls for a better understanding of the fuel costs and associated carbohydrate requirements of various training sessions commonly undertaken by athletes [2]. Based on the close relationship between mechanical work output and metabolic energy expenditure [3], it is plausible that readily-available measures of exercise quantification (i.e., training load) could be used to model and predict carbohydrate utilization during exercise, particularly when combined with other measures obtained from traditional laboratory testing. However, historically there have been several challenges in studying this.

Carbohydrate utilization is estimated using indirect calorimetry, but this method is not valid during high-intensity intermittent exercise due to shifting acid-base balance and excess (non-oxidative) CO_2 excretion through hyperpnea [4]. Changes in muscle glycogen are often used to estimate carbohydrate utilization, but this requires an invasive muscle biopsy with medical supervision and does not provide information on whole-body carbohydrate use. Furthermore, the level of muscle glycogen depletion is also specific at a subcellular and cellular level [5, 6], and therefore may vary with repeated sampling from the same individual, given the observed variability in muscle fiber type distribution along a muscle's length [7]. Another approach is to calculate the contribution of the three energy systems (aerobic, anaerobic alactic, and anaerobic

lactic) during exercise based on measurements of oxygen uptake, the fast component of excess post-exercise oxygen uptake, and net changes in blood lactate concentration [8]. This method has been used across a range of sports including cycling [8], boxing [9], running [10], and rowing [11], resulting in an estimate of kilojoules (kJ) produced by each system. However, this does not consider the substrate (i.e., fat or carbohydrate) used for energy production or differences in efficiency with each substrate [12]. In combination, the traditional gas exchange measurement and the three-system approach to energy contributions could be used to estimate the total carbohydrate and energy cost of exercise at any intensity, but to our knowledge this has yet to be reported.

In the daily training environment, athletes and coaches routinely capture multiple training load indices, which can be measured and classified as either internal and/or external, based on the measurable aspects occurring internally or externally to the athlete [13]. Internal load reflects the relative physiological strain and disturbance in homeostasis of the metabolic processes in response to an external load, which is characterized by objective measures such as distance, power, or speed [14]. Due to the wide availability of cycling power meters, total work done during exercise (TWD) is a common measure of external training load for cyclists. Furthermore, very large correlations (r = 0.96–0.97) have been reported across multiple measures of internal and external training load in cyclists during racing and training [15], suggesting other metrics such as session rating of perceived exertion (sRPE) [16], Lucia training impulse (LuTRIMP) [17], and training stress score (TSS) [18] can also provide relevant information to athletes and coaches. However, there is no gold standard measure of training load [13], and measures of external and

internal load are not always consistent. For example, TSS may overemphasize intensity compared with TWD, LuTRIMP, and sRPE [19, 20], but without a standard for comparison it is unclear which measure may be over- or under-emphasizing intensity.

Given the strong theoretical and mechanistic links between measures of training load and both carbohydrate and total expenditure during exercise, it seems plausible that a strong association should exist between these phenomena, to the extent that practitioners may be able to estimate expenditure from training load metrics with reasonable precision. To our knowledge, this has yet to be explored. Therefore, to investigate the relationship between measures of training load and carbohydrate/energy expenditure, we used a novel method to estimate carbohydrate utilization from aerobic and anaerobic sources during moderate-intensity steady-state exercise and high-intensity interval training. The primary aims of our study were twofold. The first was to quantify the bivariate relationships between both carbohydrate utilization and overall energy expenditure during exercise and commonly-used measures of training load across sessions of varying duration and intensity. This answers the question "can training load be used as a proxy measure of carbohydrate/energy expenditure to quantify the correlation between an individual's intake and exercise expenditure?", as we have recently proposed [21]. The second aim was to model and predict carbohydrate and energy expenditure during exercise using measures of training load, alongside measures of cardiorespiratory fitness, dietary intake, and sex. This answers the question "can individuals estimate their carbohydrate/energy expenditure based on commonly available, non-invasive measures?". In addition to using internal cross-validation of the prediction models, we also tested a separate set of athletes to quantify how well the model

predictions would translate to a different set of athletes performing a different type of workout, across different exercise modalities (cycling, running, and kayak). Finally, the study design also allowed us to examine day-to-day variability in heart rate (HR), rating of perceived exertion (RPE), carbohydrate oxidation, and oxygen consumption ($\dot{V}O_2$) during low-intensity cycling.

Methods

This study was undertaken in two parts — a primary study in cycling where participants performed four different laboratory-based training sessions, and a validation study where participants performed a single laboratory-based training session using one of three exercise modalities (cycling, running, or kayak). Conceptually, this study includes a cross-sectional observational study (primary study) as well as a prediction model development and validation study (primary and validation studies). Accordingly, we adhered to the STROBE [22] and TRIPOD+AI [23] reporting guidelines where applicable.

Participants. The primary study included 15 participants (10 male, 5 female), the validation study included 58 participants (41 male, 17 female). Sample size calculations are described in detail under *Statistical Analysis*. Participant characteristics are shown in Table 1. The study was open to all healthy males or females aged 18-55 regularly performing ≥ 3 h per week of training in the modality used for testing (cycle, run, or kayak). All interested participants that met the criteria were enrolled, and everyone enrolled completed all sessions. Study protocols and materials were approved by the Auckland University of Technology Ethics Committee (23/143 and 23/258).

Table 1. Participant characteristics

Characteristic	Cycle - primary N = 15[1]	Cycle - validation N = 21[1]	Kayak - validation N = 17[1]	Run - validation N = 20[1]
Sex				
F	5 (33%)	7 (33%)	4 (24%)	6 (30%)
M	10 (67%)	14 (67%)	13 (76%)	14 (70%)
Age	34.2 (9.7)	31.6 (10.7)	29.0 (11.3)	28.0 (7.6)
Mass (kg)	74.8 (9.8)	73.7 (11.3)	77.9 (10.1)	70.0 (13.3)
BMI	23.2 (1.8)	23.2 (2.3)	24.2 (1.9)	22.0 (2.2)
Training h/week	8.8 (3.1)	10.7 (6.6)	13.2 (4.1)	9.6 (5.1)
$\dot{V}O_2max$ (ml/kg/min)	51.9 (7.2)	53.5 (11.0)	46.2 (4.1)	57.5 (7.2)
PPO (W)	355 (54)	341 (76)	167 (34)	351 (68)
VT_1 (W)	167 (35)	175 (48)	80 (19)	249 (43)
VT_2 (W)	246 (42)	240 (58)	112 (22)	289 (52)
Dietary CHO (g/kg)	3.7 (0.9)	4.1 (1.8)	3.6 (1.1)	4.5 (1.8)
Dietary fat (g/kg)	1.5 (0.4)	1.3 (0.5)	1.3 (0.3)	1.5 (0.8)
Dietary protein (g/kg)	1.8 (0.5)	1.6 (0.6)	1.8 (0.4)	2.1 (1.0)
Dietary kcal (per kg)	35.7 (7.6)	34.1 (11.6)	33.4 (6.6)	40.9 (14.2)

Dietary intake represents mean values from four days for each participant for the primary study and one day for each participant in the validation study. CHO = carbohydrate, PPO = Peak Power Output, VT = Ventilatory Threshold.

[1]n (%); Mean (SD)

Primary study

Participants reported to the laboratory on five occasions, with 1–7 days between sessions and without performing high-intensity sessions on consecutive days. Participants refrained from intense exercise and alcohol 24 h before each visit and avoided caffeine 16 h before each visit. Exercise was permitted the day before each session, with the duration and sRPE recorded using the Borg CR100® scale [24]. No exercise was allowed on the day of any laboratory visit. Participants were asked to maintain their normal dietary habits and recorded their intake for one day prior to each of visits 2–4 using a smartphone-based application which features foods from

Australia and New Zealand (Easy Diet Diary, https://xyris.com.au/products/easy-diet-diary). All trials were conducted under standard laboratory conditions (18–20°C, 40–65% relative humidity), with participants fan-cooled during exercise.

Visit 1. Participants reported to the laboratory in an overnight-fasted state. After obtaining written informed consent and completing a health screening, a graded exercise test was performed to determine ventilatory thresholds and maximal oxygen consumption ($\dot{V}O_{2max}$). Participants cycled on an electronically-braked cycle ergometer (Excalibur Sport, Lode BV, Groningen, The Netherlands), with expired gas collected and analyzed using a computerized metabolic system with mixing chamber (TrueOne2400, ParvoMedics, Sandy, UT, USA). The test began at 95 W, and power output increased by 35 W every 3 min until identification of the second ventilatory threshold (VT_2), where the ventilatory equivalent for oxygen ($\dot{V}E.\dot{V}O_2^{-1}$) and carbon dioxide ($\dot{V}E.\dot{V}CO_2^{-1}$) increased alongside a reduction in $PetCO_2$ [17]. Participants then cycled for 10 min at 100 W, followed by a step test starting at 150 W and increasing 30 W per minute to task failure to obtain $\dot{V}O_{2max}$. Thirty seconds following the test a 0.3 µL blood sample was collected from the left index fingertip and analyzed immediately using a portable blood lactate analyzer (Lactate Pro 2, Carlton, Australia). The first ventilatory threshold (VT_1) was identified as the work rate at which $\dot{V}E.\dot{V}O_2^{-1}$ began to increase in the absence of changes in $\dot{V}E.\dot{V}CO_2^{-1}$. Peak power ($W_{max}$) was determined by the workload in the last completed stage plus the workload relative to the time spent in the last incomplete stage [power of completed stage + (30*(seconds at uncompleted stage/60)], and $\dot{V}O_{2max}$ and peak fat oxidation (PFO) were recorded as the highest 15-s value from a moving average, calculated from a 1-s interpolation of breath-by-breath data.

Visits 2-5. In a randomized and counter-balanced order, participants completed four different training sessions as follows: 1) 90 min continuous cycling at 90% of VT_1 power (low-intensity training (LIT) long), 2) 30 min continuous cycling at 90% of VT_1 power (LIT-short), 3) 15 min continuous cycling at 90% of VT_1 power, followed by two sets of 5 x 3 min intervals with 2 min recovery between intervals (RBI) and 8 min recovery between sets (RBS, high-intensity interval training long, HIIT-long), and 4) 15 min continuous cycling at 90% of VT_1 power, followed by two sets of 10 x 30-s intervals with 30-s RBI and 8 min RBS (HIIT-short). All sessions were performed on the Lode cycle ergometer, with intervals performed using the cadence-dependent linear mode set to produce a workload of 110% VT_2 power at their preferred cadence. Participants were instructed to produce their maximal power output across intervals. All recovery intervals were active recovery at 30% W_{max}. A standardized snack (Frooze balls, 27 g CHO, 8 g protein, 19 g fat, Revive foods, New Zealand) was provided for participants to consume 45 min prior to starting exercise, and *ad libitum* water intake was allowed before and during the training session.

Music was played during all sessions except the graded exercise test. This was because many cyclists listen to music while training, particularly during extended duration indoor training sessions. Each participant selected their own playlist from a commercial streaming platform, which was repeated for each subsequent visit. Music was standardized based on individual preferences rather than playing the same music for all participants because of the influence of preferred vs. non-preferred music on RPE during exercise [25]. An sRPE value was recorded 10–

15 min following exercise using the Borg CR100® scale, which offers additional precision compared with the CR10 scale [26]. Participants were familiarized with the CR100 scale in advance of the trials and given the scale for use at home two weeks prior to the first testing session.

Expired gas was measured during the last six of every 15 min period during LIT sessions, and from minute 9 onwards during the HIIT sessions, with the exception of a 2-min break during minutes 6-7 of the recovery period between the two sets of intervals. Blood lactate concentration was measured 30-s before and 30-s after each interval set during the HIIT sessions. A schematic overview of the sessions for the primary and validation arms is shown in Figure 1.
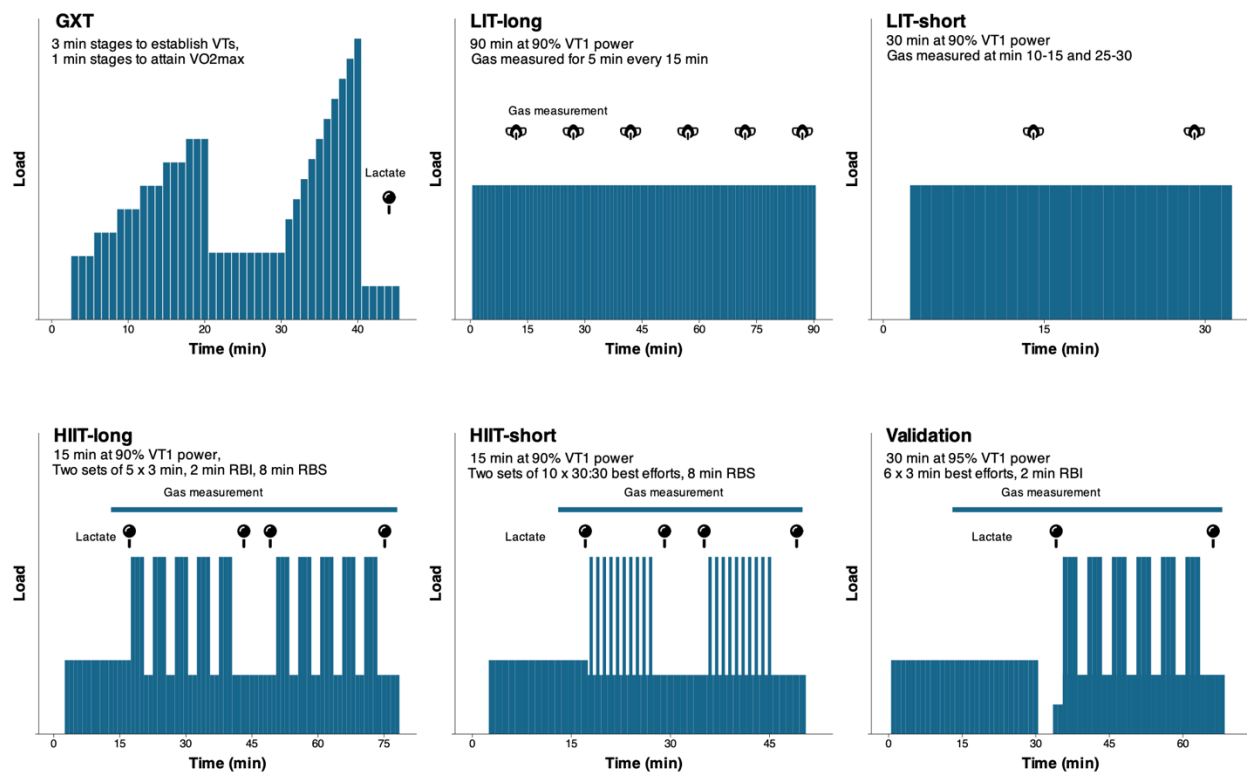
Figure 1. Schematic overview of the testing sessions. GXT: Graded exercise test, HIIT: High intensity interval training, LIT: Low intensity training, RBS: Rest between sets, VT1: first ventilatory threshold

Validation study

To validate the prediction equations established in the primary study and to assess their generalizability to other exercise modalities, 58 additional participants were recruited to perform a graded exercise test and a single exercise session using either a cycling ergometer, motorized treadmill (h/p/cosmos, Nussdorf, Germany), or kayak ergometer (Dansprint, Hvidovre, Denmark). The graded exercise test was performed for cyclists as described above, whereas running tests began at 10 km/h and increased in 1 km/h increments, and kayaking tests began at 40 W (females) and 60 W (males) and increased in 15 W (females) and 20 W (males) increments. Participants returned to the laboratory 2–7 days later to perform a mixed-intensity exercise session consisting of 30 min continuous exercise at 95% of VT_1 power, followed by 5 min recovery (5 min active recovery at 100 W for cyclists, or 3 min passive recovery followed by 2 min walking at 4 km/h for runners or 20 W on the kayak ergometer), and 6 x 3 min intervals with 2 min rest between intervals where participants were encouraged to give their best effort across the six intervals (Figure 1). The cycling intervals were performed as described for the initial HIIT sessions, using the cadence-dependent linear mode. Treadmill intervals were set at 107% VT_2 speed based on pilot testing. The intensity for kayak intervals was dictated by the participant. Recovery between intervals was set at 30% PPO for cycling, 4 km/h for running, and 20% PPO for kayak sessions.

The validation session was designed to be similar to a typical training session, but different from the sessions in the initial arm of the study. sRPE was collected 10–15 min following exercise. Expired gas was measured during the last six of every 15 min period during the 30-min continuous cycling, and for the remainder of the session. Lactate was measured 30-s before and 30-s after the interval set. A Stryd power meter (Stryd, Boulder, CO, USA) was used to collect running power data [27], and stroke-by-stroke power was collected from the kayak ergometer. To increase generalizability, validation sessions could be performed at any time of day, but subjects refrained from eating in the 4-h pre-exercise window, with the exception of the same standardized snack consumed 45 min prior to exercise.

Data analysis

Carbohydrate utilization and energy expenditure during exercise were determined separately for aerobic and anaerobic energy systems. For each exercise session, breath-by-breath gas exchange data were interpolated into second-by-second values using the *whippr* R package [28]. To determine the contribution from aerobic energy production, $\dot{V}O_2$ values were converted to energy equivalents based on respiratory exchange ratio (RER) values using the conversion tables of Elia and Livesey [12]. This allows RER-specific energy conversions to be used. For example, the energy equivalent of 1 L O_2 is 4.687 kcal at an RER of 0.71, and 5.048 kcal at an RER of 1.0 [12]. This approach also allows the calculation of energy equivalents for RER values > 1.0 as CO_2 is not needed for calculations [12]. To account for excess (non-oxidative) CO_2 excretion, we considered 5.048 kcal/L as the maximum energy equivalent for the aerobic contribution if RER values were

> 1.0. Energy equivalents were calculated on a second-by-second basis for the entire session (step 1). The RER value was then used to estimate the percentage of carbohydrate and fat oxidation using the conversions of Elia and Livesey [12]. The percentage carbohydrate contribution was multiplied by the energy equivalent to calculate energy from carbohydrate sources (step 2).

To convert from energy (kcal) to mass (g) of carbohydrate, consideration of exercise intensity is required. This is because the energy yield from carbohydrate varies depending on the source, ranging from 3.719 kcal per gram of glucose to 4.187 kcal per gram of glycogen [29]. The equations of Jeukendrup and Wallis [4] vary based on exercise intensity, assuming 50% of the carbohydrate oxidation is derived from plasma glucose and 50% from muscle glycogen during low intensity exercise (40–50% $\dot{V}O_{2max}$), and 20% from glucose and 80% from muscle glycogen at moderate to high intensity exercise (50–75% $\dot{V}O_{2max}$). This results in carbohydrate oxidation yielding 3.95 kcal per gram of carbohydrate during low intensity exercise, and 4.07 kcal per gram of carbohydrate during moderate to high intensity exercise [4]. It has also been recommended that resting analyses should assume 100% glucose oxidation [4]. With this in mind, we used a scaled approach whereby the percent contribution from glycogen was assumed to be equal to the exercise intensity as a percentage of $\dot{V}O_{2max}$, allowing a second-by-second adjustment according to exercise intensity (step 3). The energy yield from glucose and glycogen-derived carbohydrate oxidation was then calculated and summed to get an intensity-adjusted energy yield from carbohydrate (step 4). The value for energy (kcal) from carbohydrate sources was divided by the adjusted energy yield to get a value of carbohydrate in grams for each second

(step 5), with these values summed to yield a session total for grams of carbohydrate utilized through the aerobic energy pathways.

A step-by-step example is shown in Box 1 for a $\dot{V}O_2$ of 2.9 L/min and an RER of 0.93, for someone with a $\dot{V}O_{2max}$ of 3.8 L/min.

Box 1. Example calculation of aerobic energy production

Step 1: Calculate energy expenditure per second.

RER of 0.93 yields 4.961 kcal per L O_2 per minute

2.9 L * 4.961 kcal per L = 14.39 kcal per minute/ 60 = 0.240 kcal per second

Step 2: Calculate energy expenditure from carbohydrate sources.

RER of 0.93 corresponds to a contribution from carbohydrate of 77.19%

0.240 kcal per second * .7719 = 0.185 kcal per second from carbohydrate

Step 3: Calculate percent contribution from glucose and glycogen sources, assuming percentage of glycogen is equivalent to percentage of $\dot{V}O_{2max}$.

2.9 L / 3.8 L = 76.3% from glycogen

1 – 76.3% = 23.7% from glucose sources

Step 4: Calculate energy yield from glucose- and glycogen-derived carbohydrate oxidation, summed for a total intensity-adjusted energy yield.

23.7% * 3.719 = 0.881 kcal/g from glucose

76.3% * 4.187 = 3.195 kcal/g from glycogen

0.881 + 3.195 = 4.076 kcal/g carbohydrate

Step 5: Calculate carbohydrate in grams per second.

0.185 kcal from carbohydrate per second (from step 2) / 4.076 kcal per g (from step 4) = 0.045 g carbohydrate per second

Step 6: Calculate the sum of the second-by-second values to get a session total. Expired gas was recorded for the last six minutes of each 15-min block during low intensity cycling, with the first minute of each collection period discarded. Therefore, values for the 5-min periods were multiplied by 3.

[End of Box 1]

Energy produced from anaerobic lactate metabolism was determined using the net lactate accumulation, body mass, and O_2 lactate equivalent method [3, 8], with example calculations shown in Box 2. Lactate was measured before and after the interval sets during the HIIT trials, with the change in lactate (post – pre) multiplied by 3 mL $O_2.kg^{-1}.mmol.L^{-1}$ to create an oxygen equivalent [3], which was then multiplied by 21.1 $kJ.L^{-1}$ [8], and divided by 4.184 to convert from kJ to kcal (step 1).

To convert from energy (kcal) to mass (g) of carbohydrate, consideration of the ATP yields from anaerobic glycolysis and aerobic oxidation of carbohydrate is required. The net yield of anaerobic glycolysis is 2.9 ATP when starting from glycogen (assuming 90% α-1,4 glycosidic bonds) and 2 ATP when starting from glucose [30]. The complete oxidation of glycogen yields 34.35 ATP, and complete oxidation of glucose yields 33.45 ATP [30]. During high-intensity exercise, we assume the substrate for anaerobic glycolysis is glycogen, implying it would require 11.845 times more carbohydrate (because 34.35 / 2.9 = 11.845) to produce the same amount of ATP via anaerobic, compared with aerobic, metabolism. Based on the aerobic yield of 4.187 kcal per gram of glycogen [12], we calculated grams of carbohydrate from anaerobic sources as kcal from step 1 divided by 4.187, multiplied by 11.845 to account for the inefficiency of ATP production from anaerobic glycolysis (step 2). This process was repeated for both interval sets. Total carbohydrate expenditure was calculated by summing the contributions from the aerobic and anaerobic systems. A step-by-step example of anaerobic energy calculation is shown in Box 2.

Box 2. Example calculation of anaerobic energy production

Step 1: Calculate kcal from anaerobic energy production.

Delta lactate = 9.3 mmol.L^{-1} (post) – 1.5 mmol.L^{-1} (pre) = 7.8 mmol.L^{-1}

Oxygen equivalent = 3 * 7.8 * 70 kg body mass = 1638 ml O_2 = 1.638 L O_2

1.638 L * 21.1 kJ per L = 34.56 kJ

34.56 kJ/ 4.184 = 8.26 kcal via anaerobic energy production

Step 2: Convert from kcal to grams of carbohydrate while accounting for the inefficiency of anaerobic energy production.

8.26 kcal / 4.187 * 11.845 = 23.4 g carbohydrate

[End of Box 2]

A visual overview of the pathways involved in energy production and rationale for this approach is provided in Figure 2.
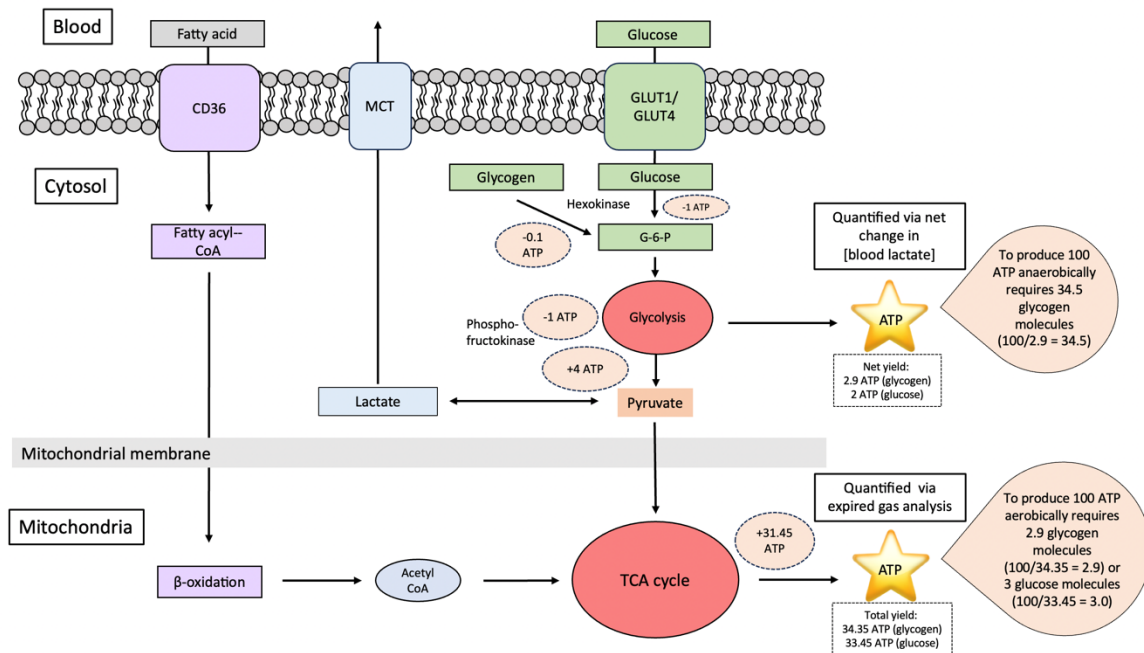

Figure 2. Overview of primary energy producing pathways in skeletal muscle. During glycolysis from glucose, 1 ATP is consumed at hexokinase and 1 ATP is consumed at phosphofructokinase to yield 2 trioses, each of which generates 1 ATP at phosphoglycerate kinase and 1 ATP at pyruvate kinase, for a net yield of 2 ATP/glucose. When starting from glycogen, less ATP is needed for the initial activation at hexokinase (~0.1 ATP), resulting in a greater net yield of 2.9 ATP [30]. An additional 31.45 ATP is produced from oxidative reactions, bringing the maximum total yield to 34.35 ATP from glycogen and 33.45 ATP from glucose. Glucose enters the cell via glucose transporters (GLUT)1 and 4. Lactate can be removed from the cell via monocarboxylate transporters (MCT). Fatty acids can enter the cell via fat transport proteins including cluster of

differentiation 36 (CD36). Differences in efficiency are highlighted by a comparison of ATP production; to produce 100 ATP requires 34.5 glycogen molecules via anaerobic energy production or 2.9 glycogen molecules via aerobic energy production

In addition to the strong theoretical and mechanistic rationale for this approach, we also tested our method using data from previously published studies which included metabolic tracers and/or muscle glycogen measurements. Calculations are provided in Supplementary File 1, showing good agreement (e.g., estimations of total carbohydrate utilization within ~1–4 g) when comparing our method with estimates using invasive techniques.

For each exercise session, six measures of training load were calculated as shown in Table 2.

Table 2. Measures of Training Load and their calculations.

Total Work Done (TWD-kJ)	A measure of total mechanical energy spent (kJ), collected from the ergometer (cycling and kayak) or Stryd power meter (running).
Session Rating of Perceived Exertion- Training Load (sRPE-TL)	Session RPE * Duration (min) [16]. This value was divided by 10 to account for the 100-pt scale, allowing easier comparisons to other research using the 10-pt scale.
Lucia training impulse (TRIMP)	(Duration (min) in zone 1 * 1) + (Duration in zone 2 * 2) + (Duration in zone 3 * 3). Zones were calculated using power (LuTRIMP$_W$) and HR (LuTRIMP$_{HR}$), with zone 1 below VT_1, zone 2 between VT_1 and VT_2, and zone 3 above VT_2 [17].
Training Stress Score (TSS)	[(seconds × NP × IF)/(FTP × 3600)] × 100 [18]. Power at VT_2 was used as an estimate of functional threshold power (FTP), as has been done by others [31]. Normalized power (NP) is calculated by creating rolling 30-s averages and raising each value to the fourth power, then taking the fourth root of the average of the fourth powers. Intensity factor (IF) is calculated as NP/ FTP.
Heart rate TSS (TSShr)	Calculated the same as TSS, substituting HR for power and using VT_2 HR as a measure of functional threshold HR.

Statistical Analysis

Primary study

To estimate differences in training load across the four sessions for each metric, a series of linear mixed models were fitted using the *lme4* R package with training load as the dependent variable, session as a fixed factor, and participant ID as a random intercept. Model-estimated means were calculated using the *emmeans* R package and contrasts between each session (within each training load metric) were estimated using the Holm correction for multiple comparisons. To examine the bivariate relationship between training load measures and the total carbohydrate and kcal cost of exercise, repeated-measures correlation was used, which allows analysis of repeated-measures data without violating independence assumptions [32]. To examine day-to-day variation all four trials began with the same 15-min period of cycling at 90% VT_1 power, allowing us to compute the Typical Error of Measurement (TEM) for VO_2, HR, carbohydrate oxidation, and RPE according to the approach of Hopkins [33].

To predict carbohydrate utilization and energy expenditure based on training load and other commonly measured variables known to influence substrate selection such as $\dot{V}O_{2max}$, sex, and dietary intake [34], multivariable models were created for each of the six training load measures predicting each of the two dependent variables (energy expenditure and carbohydrate utilization) using generalized estimating equations (GEE). GEE models provide population-averaged (e.g., marginal), rather than subject-specific models while accounting for repeated measurements within participants [35]. Quasi Information Criterion (QIC) was used for selecting an independence correlation structure as the working correlation matrix [36]. The following

variables were considered for the full models: training load, training load^2, session duration (min), session duration^2, prior-day sRPE-TL, type of session (continuous or interval training), prior day dietary carbohydrate and fat intake (g/kg), $\dot{V}O_2$ at VT_2 (L/min and % $\dot{V}O_{2max}$), $\dot{V}O_{2max}$ (mL/kg/min and L/min), blood lactate at the end of the $\dot{V}O_{2max}$ test, PFO (g/min), and sex. The following pre-specified interactions were also considered in the full model: prior-day sRPE-TL x prior day carbohydrate intake, session duration x training load, type of session x training load, type of session x $\dot{V}O_{2max}$, and type of session x $\dot{V}O_2$ at VT_2.

The top candidate models were identified using the *glmulti* R package [37], which performs a genetic search across possible models specified by a given set of predictors and selects the top models according to corrected Akaike Information Criterion (AICc). From the reduced pool of models, we performed participant-level leave-one-out cross-validation which fits a series of models on all but one of the participants, whose four sessions are used as a hold-out testing set [38], selecting the model with the lowest mean absolute error (MAE) as the final model for each measure of training load. The fit of each model was checked by visualizing the Q–Q and other residual plots to ensure approximate residual normality and homoscedasticity using the *performance* R package. Model performance is reported as the coefficient of determination (R^2), which represents the proportion of variance explained by the model, and the MAE, which quantifies the average absolute discrepancy between the observed and predicted values. These metrics were calculated using both in-sample data (i.e., the same data used to train the model and evaluate performance) and cross-validation, which offers a more realistic and unbiased (or least biased) estimate of model performance in the population in which the model is intended

[39]. Performance metrics for cross-validation are reported as mean [95% confidence intervals]. There were no missing data for models in the primary study.

Validation study

Data from the validation sessions were analyzed in the same manner as the primary study, with each session analyzed as only the 30-min low-intensity portion, and as the full session (30-min low intensity and the high-intensity intervals). Values of total carbohydrate utilization and energy expenditure for each session were predicted from the previously fit models for each measure of training load. Model performance was assessed using measures of overall fit R^2 (proportion of variance in explained in the external validation dataset, calculated using the traditional definition with sum of squares rather than the correlation between predicted and actual values) and MAE, and assessed for calibration which was quantified as calibration-in-the-large (the difference between mean observed and mean predicted outcome values, with 0 being ideal) and calibration slope (the agreement between predicted and observed values across the range of predicted values, with a slope of 1 being ideal) [39]. Finally, models were recalibrated using the intercept and slope of a linear model regressing the actual values on the predicted values [40], with measures of R^2 and MAE reported on the calibrated data. Due to technical issues, data for the low-intensity portion of one kayak trial and the total work done for one running trial were missing. Rather than using imputation, these data points were omitted from the predictions. One other kayak trial consisted of only three intervals due to equipment malfunction but data from the first 45 min of the session was included for analysis.

Based on Riley et al. [41], a minimum sample size of 15 was calculated for the primary study. This calculation used an estimated adjusted R^2 value of 0.89 and considered a model with up to 6 predictors. The choice of 6 predictors was derived from Riley et al. [42], to ensure a shrinkage factor of at least 0.9 and maintain a difference between adjusted and apparent R^2 values below 0.05. However, our sample size is below the minimum size of 240 (using the rule of 234 + number of predictors) needed for precise estimates of the residual standard deviation [41]. This means there will still be some uncertainty in the parameter estimates that can only be solved with very large sample sizes that extend beyond the capacity of this project. The approach of Archer et al. [43] was used to calculate the minimum sample sizes needed in the validation dataset to obtain precise estimates of R^2, calibration in the large, and calibration slope, assuming 90% confidence intervals with target widths of 0.2 for R^2 and 0.2 for calibration slope. Calculations were made separately for each model, resulting in a minimum requirement of 7–11 participants (kcal) and 13–19 participants (carbohydrate) in each validation arm depending on the model (Supplemental Table 1, supplemental R code). For the kayak arm of the validation study we were only able to recruit 17 athletes, which is sufficient for all energy expenditure models and four of the six carbohydrate models, but just below the target sample size of 18 for LuTRIMP$_W$ and 19 for sRPE-TL.

To determine the minimum sample size for detecting differences in training load across the four sessions in the primary study, we calculated the means and standard deviations for each session, estimated the pooled standard deviation, and determined the effect size (Cohen's f). Using these values, we performed a power analysis which indicated that a sample size of 6 participants per

group was required to achieve 95% power at a 5% significance level. All analyses were carried out with R version 4.3.1 (The R foundation for Statistical Computing, Vienna, Austria). Descriptive statistics are provided as mean ± SD, statistical significance was accepted at $p < 0.05$.

Results

Participant characteristics and self-selected dietary intake are shown in Table 1. The TEM of day-to-day variation during low intensity cycling was 2.9 b/min for HR, 0.08 L/min for $\dot{V}O_2$, 0.21 g/min for carbohydrate oxidation, and 4.0 arbitrary units (0–100 AU scale) for RPE. Training load, energy expenditure, and carbohydrate utilization for each training session are shown in Figure 3. All measures of training load were significantly different for each training session, except total carbohydrate utilization and TSS, which were not different between LIT-long and HIIT-short, and TWD-kJ, which was not different between HIIT-long and LIT-long. Repeated-measures correlation analysis indicated very large to near perfect correlations (0.71–0.98) between all training load metrics and both outcomes (Fig. 4).

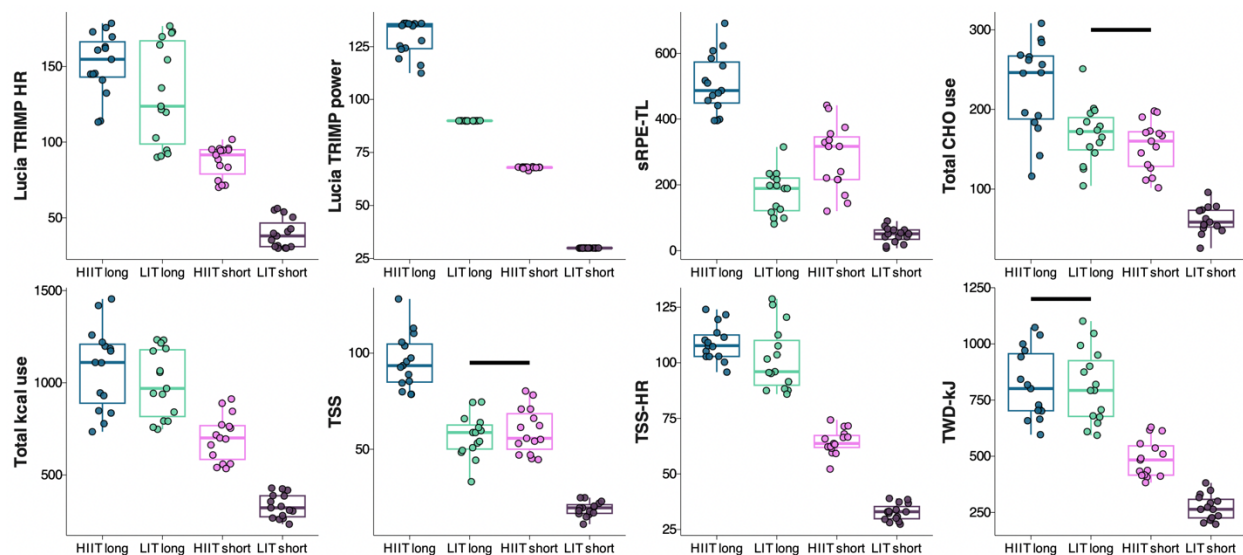**Figure 3.** Training load and kcal and carbohydrate (CHO) use for each training session in the primary study. Solid black lines indicate no significant difference between sessions (p > 0.05). HIIT: High-intensity training sessions, LIT: Low intensity training sessions, sRPE-TL: Session RPE-training load, TSS: Training Stress Score, TSS-HR: TSS calculated with heart rate, TWD-kJ: Total Work Done (kJ).

Figure 4. Repeated measures correlation values with 95% Confidence Intervals between training load measures and carbohydrate (CHO, top) and kcal use (bottom) from the primary study. sRPE-TL: Session RPE-training load, TSS: Training Stress Score, TSS-HR: TSS calculated with heart rate, TWD-kJ: Total Work Done (kJ)

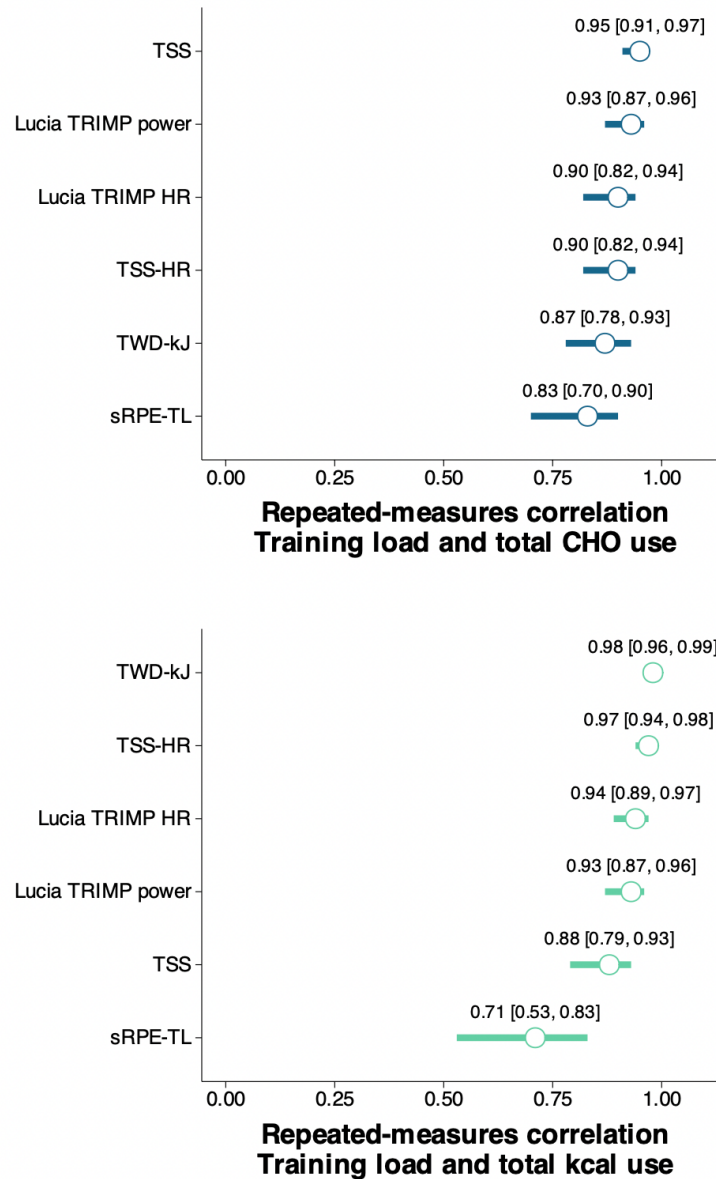Multivariable regression models were created for each of the six training load measures predicting the two dependent variables (energy expenditure and carbohydrate utilization –

resulting in 12 unique models), with additional variables that can be easily obtained from routine exercise physiology testing. Model coefficients for the included variables are shown in Table 3, along with model performance metrics R^2 and MAE using both in-sample data and leave-one-participant out cross-validation. All models explained a large amount of variance (R^2 values of 0.77–0.96) and displayed good accuracy (MAE of 16–21 g [10–14%] carbohydrate and 53–82 kcal [7–11%]

Table 3. Model performance measures and coefficients.

	R2	cv-R2	MAE	cv-MAE	Intercept	Type-Continuous	Sex-f	TL	TL2	Duration (min)	VT2 VO2 (L)	VT2 (%max)	VO2max (L)	VO2max (ml/kg/min)	Prior-day sRPE-TL
CHO															
TSS-HR	0.93	0.84 [0.74, 0.94]	16.0	17.6 [14.4, 20.8]	-22.77 (18.23)	-42.77 (5.85)	-28.83 (6.43)	1.585 (0.086)					23.62 (3.65)		-0.039 (0.007)
TSS	0.92	0.82 [0.71, 0.93]	16.1	18.2 [14.3, 22.1]	-28.01 (23.51)	-38.39 (7)	-21.69 (7.48)		0.007 (0.001)	1.43 (0.12)	32.51 (6.51)				-0.062 (0.012)
TWD-kJ	0.92	0.85 [0.77, 0.93]	16.5	17.9 [13.8, 22]	76.52 (9.3)	-49.93 (6.26)	-22.37 (4.92)	0.201 (0.009)							-0.06 (0.011)
Lucia TRIMP power	0.89	0.77 [0.62, 0.91]	18.2	20.1 [15.8, 24.4]	-69.16 (16.76)	-50.39 (7.23)			0.003 (0.001)	1.39 (0.14)	46.83 (4.56)				-0.058 (0.015)
Lucia TRIMP HR	0.90	0.79 [0.68, 0.89]	18.4	21.2 [17.6, 24.9]	-21.36 (31.21)	-61.47 (7.22)	-16.26 (9.38)		0.002 (0.001)	1.4 (0.2)	35.24 (8.85)				-0.046 (0.009)
sRPE-TL	0.88	0.77 [0.66, 0.88]	19.1	21.2 [16.8, 25.6]	-75.89 (16.46)	-54.53 (8.67)			1e-04 (4e-05)	1.8 (0.13)	46.45 (4.75)				-0.058 (0.014)
kcal															
TWD-kJ	0.96	0.93 [0.89, 0.97]	52.6	58.5 [38, 78.9]	135.04 (34.97)	-95.68 (15.56)	-41.79 (23.86)	1.18 (0.036)							
TSS	0.96	0.92 [0.89, 0.96]	55.1	59.6 [48, 71.1]	-709.07 (42.34)			5.3 (0.288)		7.56 (0.43)	225.35 (13.96)				
TSS-HR	0.94	0.90 [0.85, 0.95]	61.7	72.2 [54.3, 90.1]	-772.78 (198.04)	-116.84 (23.67)	-37.65 (24.65)	5.12 (0.807)		5.23 (1.2)		3.23 (1.615)	164.73 (21.65)		

	R2	cv-R2	MAE	cv-MAE	Intercept	Type-Continuous	Sex-f	TL	TL2	Duration (min)	VT2 VO2 (L)	VT2 (%max)	VO2max (L)	VO2max (ml/kg/min)	Prior-day sRPE-TL
Lucia TRIMP HR	0.93	0.88 [0.82, 0.93]	69.1	79.4 [62.3, 96.6]	-830.95 (261.22)	-177.24 (25.05)			0.006 (0.002)	9.7 (0.77)		4.376 (2.359)	172.96 (19.11)		
Lucia TRIMP power	0.93	0.88 [0.83, 0.93]	72.0	78.6 [63.6, 93.7]	-786.48 (71.93)			10.5 (1.619)	-0.03 (0.007)	4.05 (1.07)	222.81 (20.51)				
sRPE-TL	0.93	0.88 [0.83, 0.92]	73.1	81.7 [63.8, 99.6]	-615.45 (60.89)	-132.4 (40.94)		0.2576 (0.138)		10.69 (0.68)	208.62 (29.05)			1.453 (1.163)	

Model coefficients are shown with (std error). CHO = carbohydrate, MAE = Mean Absolute Error, cv-MAE = MAE from cross-validation, cv-R2 = R2 from cross-validation, TL = Training load, TL2 = Training load ^2, VT2 (%max) is on 0–100 scale

To test the application and generalizability of the models we performed a validation study using a different training session, and three new groups of athletes (cyclists, runners, and kayakers). Model-predicted values are shown compared to actual values for carbohydrate utilization (Fig. 5) and energy expenditure (Fig. 6). Summary values shown in Table 4, including MAE and R^2 (calculated using observed and model-predicted values), calibration-in-the-large (difference between the mean observed and the mean predicted outcome values), calibration intercept and slope (from regressing the actual values on the predicted values in a linear model), and MAE and R^2 calculated using the calibration-adjusted predictions and the actual values).

Table 4. Model accuracy from validation study

	MAE - raw	R^2 - raw	CITL	Calibration intercept	Calibration slope	MAE - calibrated	R^2 - calibrated
Carbohydrate utilization (g) - cycle							
sRPE-TL	21.9	0.85	-0.8	14.9	0.87	20.3	0.87
TSS	18.3	0.90	-2.3	8.7	0.91	17.8	0.91
TSS-HR	23.5	0.84	-11.2	5.8	0.87	19.3	0.89
TWD-kJ	16.7	0.90	-7.4	3.3	0.92	15.0	0.92
Lucia TRIMP HR	20.5	0.86	-3.2	12.7	0.87	19.1	0.88
Lucia TRIMP power	20.7	0.87	-1.5	11.7	0.89	19.6	0.88
Carbohydrate utilization (g) - kayak							
sRPE-TL	29.3	0.51	-20.9	15.8	0.68	15.0	0.87
TSS	31.6	0.46	-28.0	8.5	0.70	10.8	0.92
TSS-HR	28.7	0.58	-26.6	3.5	0.75	8.3	0.95
TWD-kJ	16.6	0.83	14.6	8.8	1.07	11.8	0.92
Lucia TRIMP HR	28.2	0.55	-23.6	10.2	0.71	12.6	0.90
Lucia TRIMP power	29.5	0.51	-23.5	12.6	0.69	14.2	0.89
Carbohydrate utilization (g) - run							
sRPE-TL	32.0	0.52	-28.4	6.3	0.75	18.6	0.86
TSS	36.9	0.38	-35.4	2.7	0.74	18.4	0.86
TSS-HR	29.5	0.62	-25.5	4.4	0.78	15.0	0.90
TWD-kJ	50.4	0.08	-50.0	-11.1	0.76	15.7	0.90
Lucia TRIMP HR	29.7	0.58	-25.6	7.5	0.76	16.4	0.88
Lucia TRIMP power	36.4	0.40	-33.7	6.3	0.73	18.5	0.86
Energy expenditure (kcal) - cycle							
sRPE-TL	79.9	0.90	8.7	80.2	0.89	70.6	0.92
TSS	68.1	0.93	2.2	45.0	0.93	65.3	0.94
TSS-HR	80.3	0.91	-20.7	51.9	0.89	72.4	0.92
TWD-kJ	53.0	0.95	-18.5	9.4	0.96	48.3	0.96
Lucia TRIMP HR	75.7	0.91	-3.0	73.9	0.88	69.8	0.92
Lucia TRIMP power	78.0	0.91	11.1	79.7	0.89	68.6	0.93
Energy expenditure (kcal) - kayak							
sRPE-TL	102.0	0.77	-45.5	94.3	0.75	58.3	0.91
TSS	87.1	0.83	-50.3	74.2	0.78	42.8	0.95
TSS-HR	77.8	0.86	-47.6	64.8	0.80	42.0	0.95

	MAE - raw	R^2 - raw	CITL	Calibration intercept	Calibration slope	MAE - calibrated	R^2 - calibrated
TWD-kJ	185.9	0.29	185.9	11.4	1.53	36.4	0.96
Lucia TRIMP HR	91.9	0.81	-45.5	81.8	0.77	56.9	0.92
Lucia TRIMP power	110.1	0.72	-90.7	42.4	0.78	56.3	0.92
Energy expenditure (kcal) - run							
sRPE-TL	91.6	0.83	-48.5	93.6	0.79	58.4	0.92
TSS	113.0	0.76	-103.2	11.1	0.84	54.1	0.94
TSS-HR	81.4	0.86	-38.1	90.2	0.81	55.1	0.94
TWD-kJ	178.3	0.47	-178.3	-28.1	0.81	39.5	0.96
Lucia TRIMP HR	85.2	0.85	-28.9	107.9	0.79	58.0	0.92
Lucia TRIMP power	125.9	0.71	-82.6	94.0	0.75	65.9	0.91

CITL = calibration-in-the-large (difference between the mean observed and the mean predicted outcome values), MAE = Mean Absolute Error, MAE-raw = MAE of predicted values, MAE-calibrated = MAE for each model following adjustment of each prediction as the calibration intercept + (predicted value * calibration slope). R2 values represent the coefficient of determination calculated using the traditional definition with sum of squares rather than the correlation between predicted and actual values.

The predictions for cycling displayed the highest accuracy, and most predictions were higher than actual values apart from TWD-kJ for kayak. This is evidenced by negative values for calibration-in-the-large and calibration slope values less than 1 (Table 4). The MAE ranged from 16.7–50.4 for carbohydrate utilization models and 53.0–185.9 for energy expenditure models.

After applying the calibration adjustment to each predicted value in the validation study (calibration intercept + predicted value * calibration slope), accuracy of all models was improved as shown in Figure 7 (carbohydrate utilization) and Figure 8 (energy expenditure). The calibrated MAE ranged from 8.3–20.3 for carbohydrate utilization models and 36.4–72.4 for energy expenditure models (Table 4).
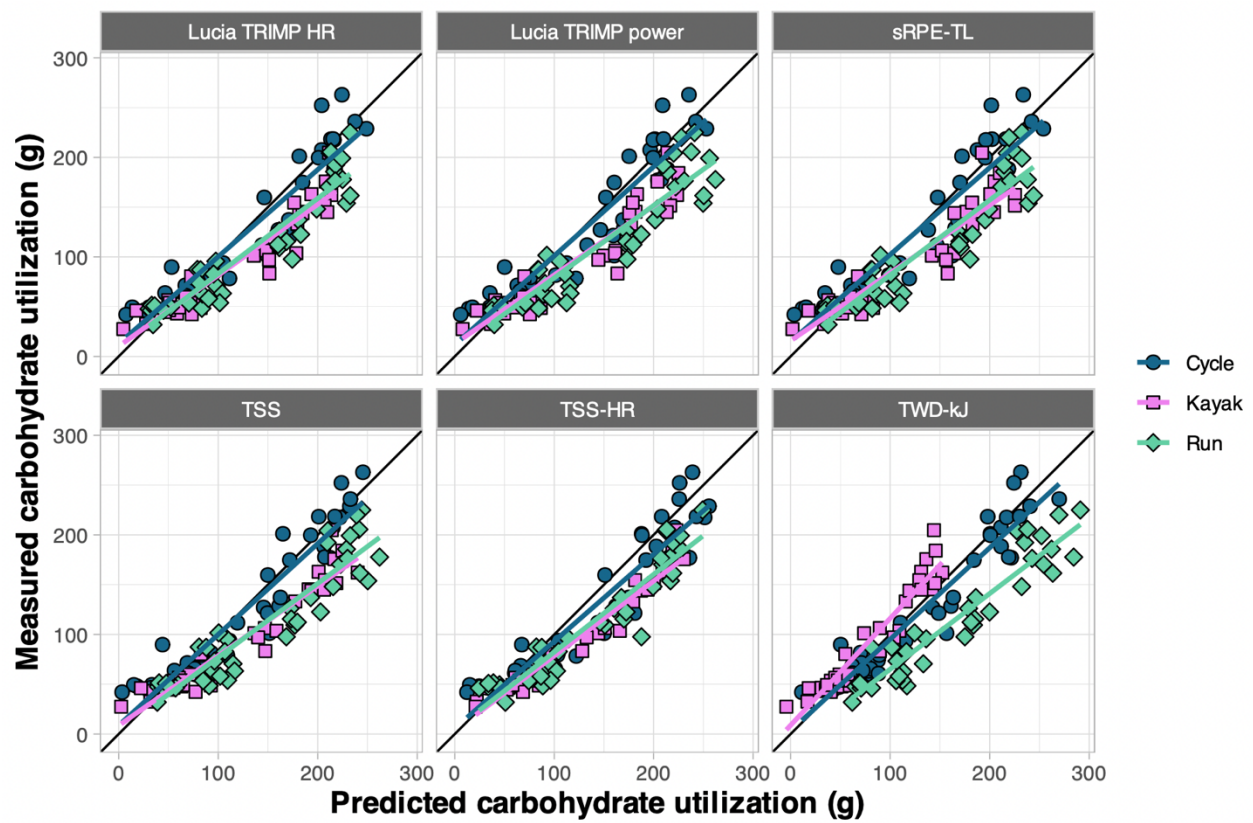
Figure 5. Predicted vs. measured values of carbohydrate utilization during validation sessions, separated by exercise mode
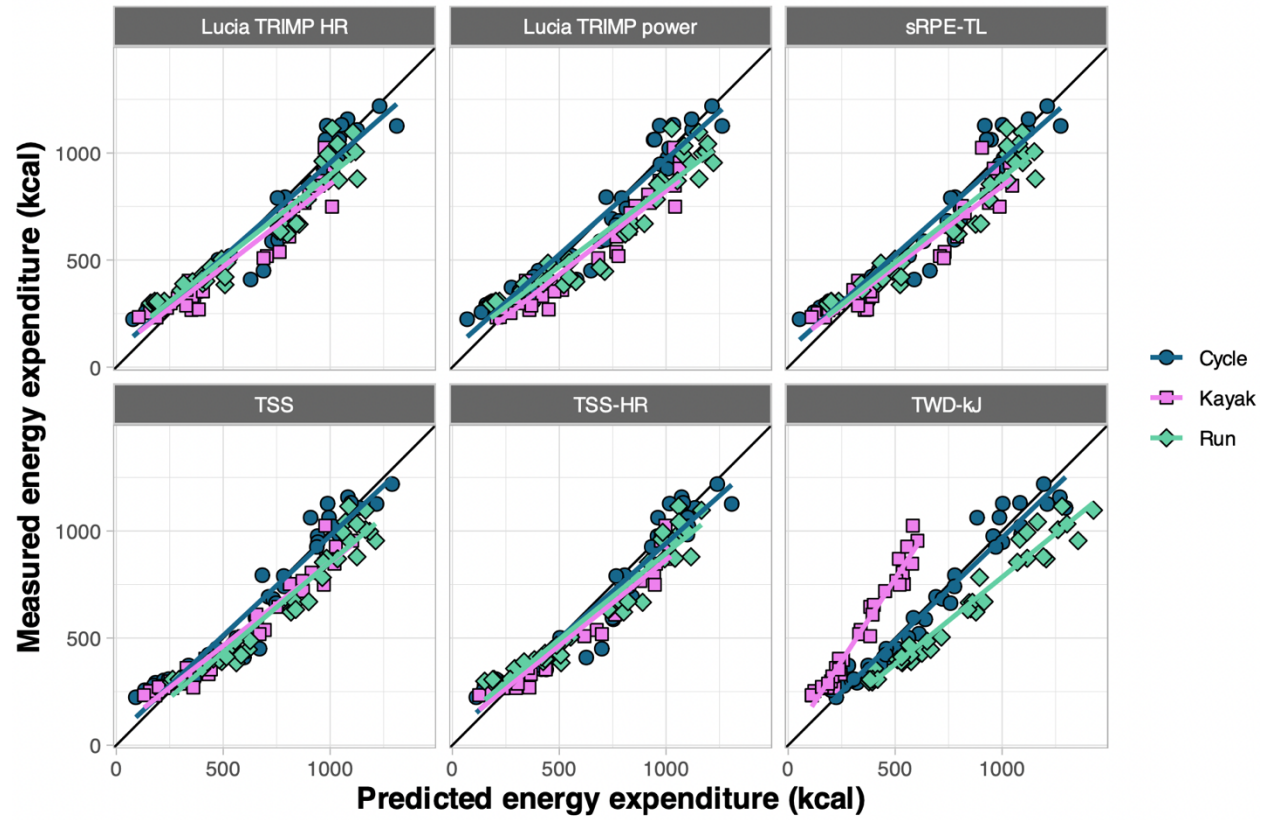
Figure 6. Predicted vs. measured values of energy expenditure during validation sessions, separated by exercise mode
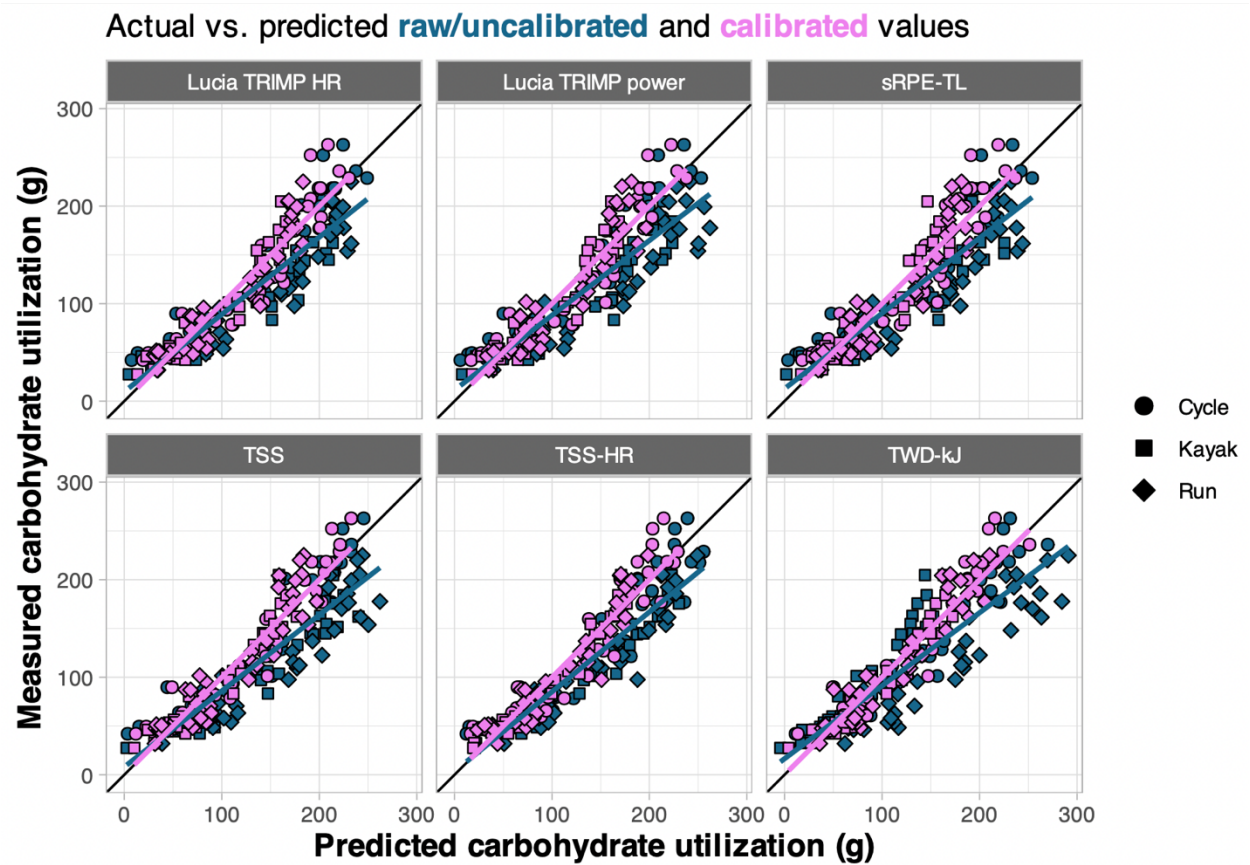
Figure 7. Predicted vs. measured values for calibrated carbohydrate models, with shapes denoting each sport and colors depicting the raw/uncalibrated values (blue) and calibrated values (violet)
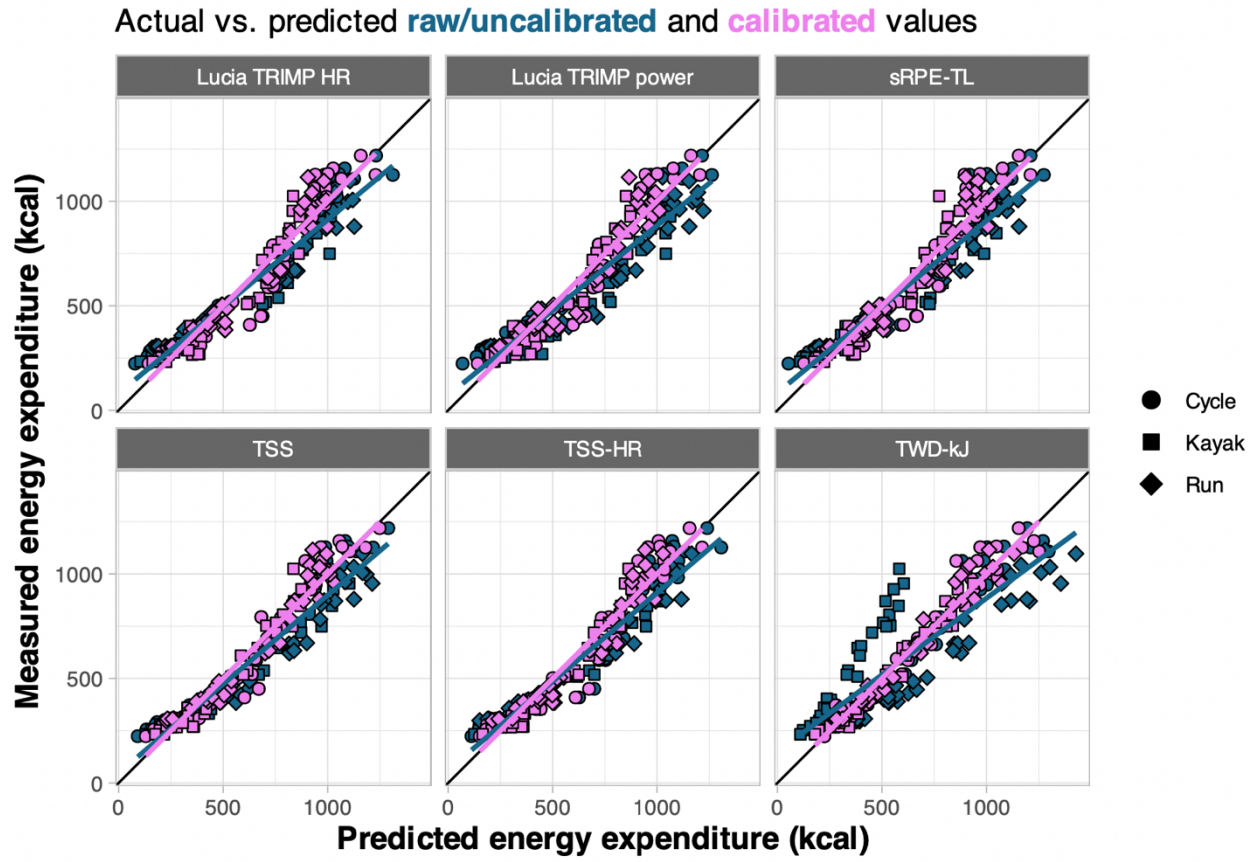
Figure 8. Predicted vs. measured values for calibrated energy expenditure models, with shapes denoting each sport and colors depicting the raw/uncalibrated values (blue) and calibrated values (violet)

Discussion

The main findings of our study were: 1) common measures of training load display very large to near-perfect associations (r = 0.71–0.98) with both carbohydrate utilization and energy expenditure during exercise, 2) TSS was the only measure of training load to accurately reflect similar total carbohydrate utilization between the longer-duration low intensity session and the shorter-duration high-intensity session, 3) carbohydrate utilization and energy expenditure during cycling could be predicted with a high degree of accuracy (MAE 17–24 g carbohydrate, 53–80 kcal) using measures of training load along with easily-obtainable laboratory measures, and 4) these models can be applied in running and kayaking when used with a calibration adjustment.

A key aim of the study was to quantify the bivariate relationships between both carbohydrate utilization and overall energy expenditure during exercise and commonly-used measures of training load across sessions of varying duration and intensity. This investigated the question "can training load be used as a proxy measure of carbohydrate/kcal expenditure to quantify the correlation between an athlete's intake and exercise expenditure?", as we have recently proposed [21]. Repeated-measures correlations showed very large to near perfect relationships between each of the training load measures and both carbohydrate utilization and energy expenditure, supporting the use of any of the training load measures by athletes looking for a way to match their carbohydrate or energy intake with their exercise. The weakest relationships were observed for sRPE-TL, but the correlation values 0.71–0.83 would still be considered very large and sRPE-TL is a more pragmatic option for many athletes and practitioners.

Despite the ubiquity of training load quantification and variety of methods for measuring load, there is no consensus on which methods best represent the true load of a training session [13, 31, 44]. This is context-dependent and relies heavily on the nature of the exercise stimulus (e.g., the sport or mode of training). Some suggest TSS is influenced more by intensity than other training load metrics such as TWD, LuTRIMP, and sRPE-TL, meaning TSS will be different for sessions where the same amount of energy is expended either at low or high intensity [19, 20]. This is related to the quadratic relationship between TSS and exercise intensity, whereas most other measures of training load have linear or exponential relationships [20]. The rationale for this quadratic relationship has been questioned based on the lack of quadratic relationship between exercise intensity and measures of internal load such as RPE, $\dot{V}O_2$, HR, blood lactate, as well as biochemical and hormonal responses [20]. However, the close relationship we observed between TSS and total carbohydrate use during exercise (Fig. 4), along with the observation that only TSS could accurately differentiate the carbohydrate needs of the various sessions (Fig. 3), suggests the quadratic relationship between TSS and exercise intensity may be a "feature", rather than a "bug". The inefficiency of the anaerobic energy pathways [30] offers a mechanism for why TSS might most accurately reflect total carbohydrate use during exercise of varying intensities. There has also been concern of a misestimation of the physiological impact of a long, easy endurance ride compared with a short high-intensity ride when expending the same amount of energy [20]. This concern is reflected in our study by the higher values of sRPE-TL for the HIIT-short compared with LIT-long sessions, despite TWD being higher for the LIT-long session, and TSS being the same between the two sessions (Fig. 3).

Although it has been suggested that studies comparing different training intensities should equalize sessions in terms of energy expenditure or work [45], TSS could be considered as a viable alternative for matching training load between groups. When combined with the known influence of post-exercise muscle glycogen concentrations on the molecular adaptations to exercise [46, 47], our findings can help explain the advantage of high-intensity, compared with moderate-intensity training, commonly observed when total work is matched between groups [48-51]. Retrospective analyses of training studies comparing moderate and high-intensity training using TSS could further explore this hypothesis.

This study used a novel method of estimating carbohydrate utilization and energy expenditure during exercise, extending methodological approaches from adjacent fields of biochemistry, physiology, and nutrition. Several challenges had to be overcome including energy conversions that are dependent on the relative contributions of fat and carbohydrate, the contributions of glucose and glycogen, differences in aerobic and anaerobic metabolism, the non-validity of gas exchange measures during high-intensity exercise, and the computational challenges associated with large data sets.

Carbohydrate oxidation is often calculated using measures of $\dot{V}O_2$ and $\dot{V}CO_2$, but this method assumes a stable bicarbonate pool and is thus unreliable at intensities above ~75% $\dot{V}O_{2max}$ due to shifting acid-base balance and excess (non-oxidative) CO_2 being excreted through hyperpnea [4]. Our approach, using the $\dot{V}O_2$-based calculations of Elia and Livesey [12], allows aerobic energy

expenditure to be calculated during higher intensities, and comparison of the two calculation methods during low intensity exercise revealed near-perfect correlation values of 0.97–0.99 (data not shown). Although lactate is produced during steady-state exercise, a major portion is eliminated through oxidation and so the rate of oxygen consumption effectively accounts for and reflects the energy generated via aerobic glycolysis [52].

Data were analyzed on a second-by-second basis, which allowed continuous adjustment of the energy yield of O_2 based on substrate utilization, and the intensity-dependent changes in the relative contribution of glucose and glycogen for carbohydrate oxidation. For the latter, we chose a scaled approach whereby the percent contribution from glycogen was assumed to be equal to the exercise intensity as a percentage of $\dot{V}O_{2max}$. This assumption was made based on the recommendations of Jeukendrup and Wallis [4] that resting analyses should assume 100% glucose oxidation, exercise around 40–50% $\dot{V}O_{2max}$ can assume 50% of carbohydrate oxidation is derived from plasma glucose and 50% from muscle glycogen, and exercise up to ~75% $\dot{V}O_{2max}$ can assume 20% from glucose and 80% from muscle glycogen. A limitation of this approach is that the estimations assume normal glycogen concentrations, yet the balance between glucose and glycogen-derived energy production is likely to change as the exercise duration extends and glycogen is depleted [53, 54]. However, the practical difference this would make to the carbohydrate calculation would be quite small as the substrate shifts more towards plasma glucose and away from muscle glycogen. We also adjusted the energy yield of O_2 based on substrate utilization. This is in contrast with many prior studies using a single value (e.g., 20.9 kJ/L O_2), an approach which has been criticized [55].

To account for differences in efficiency of anaerobic and aerobic energy production, a conversion factor was used based on work showing the complete oxidation of glycogen yields 34.35 ATP and the net yield of anaerobic glycolysis is 2.9 ATP [30]. Assuming the primary substrate during high-intensity exercise is muscle glycogen, this implies 11.845 times more carbohydrate would be required to produce the same amount of ATP via anaerobic, compared with aerobic, metabolism (Fig. 2). This inefficiency explains the observations of extremely high rates of muscle glycogen breakdown (~85–111 mmol/kg/dry mass) following a single 30-s bout of maximal cycling [56, 57]. For comparison, 120 min of cycling at 65–75% VO_{2max} would result in a similar depletion of muscle glycogen [58, 59]. Furthermore, our estimated values for energy contribution from anaerobic metabolism are in line with the previously reported range of 15–20 kcal nonoxidative energy capacity for a 70 kg human [52, data not shown]. Although this approach would be challenging to validate with any single method, it is derived from well-established methods for energy systems quantification and displays very close agreement with two different approaches to quantification carbohydrate utilization, described in detail in Supplemental File 1. Collectively, our approach overcomes the methodological challenges that have previously precluded this type of analysis from being performed. We also believe this approach can be extended to other exercise modalities and to different types of training sessions.

The second aim of the study was to model and predict carbohydrate utilization and energy expenditure during exercise using measures of training load. Models were built in the primary study of cyclists performing four laboratory-based training sessions. The best model for each

training load metric was selected from a pool of models containing commonly used and easily obtainable measures such as $\dot{V}O_{2max}$, ventilatory thresholds, dietary intake, sex, and prior-day training load. Using internal validation (i.e., testing the model on the same data that was used to train the model), R^2 values ranged from 0.88–0.96 and MAE values ranged from 16.0–19.1 for carbohydrate utilization and from 52.7–73.2 for energy expenditure. This is referred to as 'apparent performance', which, although commonly the only approach used in the field of sports science, typically provides overly optimistic values compared with when the model is evaluated in new data [39]. Therefore, it is recommended that prediction models should have additional internal-external, and external validation [60]. Internal-external validation refers to using a portion of the data for training a model and a separate portion of data for testing it. Because each participant in the primary study performed four training sessions, we used k-fold cross-validation splitting the data by participant. This means each participant was left out once for assessment of a model fit on all other participants, with the reported metrics based on the pooled assessment data [60]. Using this cross-validation approach in the primary study, R^2 values ranged from 0.77–0.93 and MAE values ranged from 17.6–21.2 g for carbohydrate utilization and 58.5–81.7 kcal for energy expenditure, representing a reasonable and expected decline in performance from the apparent performance (Table 3). From a practical perspective, the model prediction errors were lower than the errors from dietary quantification performed by sports nutritionists, which have been reported to be ~20–65 g carbohydrate and ~140–369 kcal per day [61].

External validation was then used to quantify how well the model predictions translated to a different set of athletes performing a different type of workout, across different exercise

modalities. To this end, we recruited 58 additional athletes to perform a single in-laboratory training session using one of three exercise modalities (cycling, running, and kayak). Overall, predictions for cycling were better than running and kayaking, and the between-sport differences were highly consistent (Fig. 5, 6). This implies differences inherent to performing the sport and/or the measurement devices used for each sport are responsible for much of the prediction error, rather than simply individual variation and/or noise in the data. This idea is further supported by the pronounced improvements with model calibration (Table 4, Figures 7, 8), which adjusts the predictions using a simple linear model. For example, the greatest overpredictions were the run models which used TWD-kJ (shown as the most negative calibration-in-the-large values in Table 4). Although it is acknowledged that estimating mechanical power for running is considerably more complex than it is for cycling, we chose to use the Stryd power meter because it is widely available and has demonstrated high repeatability and a consistent relationship with $\dot{V}O_2$ [27]. However, data from our study, and others [62], suggest the absolute power reported by the device may be overestimated compared with cycling power. In contrast, energy expenditure was considerably underestimated for kayak when using TWD-kJ, which can be reconciled by the differences in gross efficiency across exercise modes. Mean gross efficiency values in our study were 20.0% for cycling, 10.1% for kayak, and 26.2% for running (data not shown). As the models were trained using cyclists only, these differences in efficiency can explain both the large errors in predictions using TWD-kJ and the considerable improvements in prediction accuracy with model calibration (Table 4). From a practical perspective, people wishing to apply these models with the calibration adjustment can use the model coefficients from Table 3 to get a predicted value of carbohydrate or energy expenditure, then solve the equation calibration intercept +

(predicted value * calibration slope) from Table 4. Taken together, our data supports the use of prediction models (with sport-based calibration) to allow individuals to estimate their carbohydrate and energy expenditure based on commonly available, non-invasive measures.

In addition to measures of training load, other variables included in the prediction models were training load squared, type of training session (continuous or interval-based), sex, session duration, $\dot{V}O_2$ at VT_2, $\dot{V}O_{2max}$, and prior-day sRPE-TL. It is noteworthy that prior-day sRPE-TL appeared in all carbohydrate utilization models but none of the energy expenditure models (Table 3). Presumably, greater prior-day sRPE-TL values imply lower levels of starting muscle glycogen, which is known to shift substrate use away from carbohydrate and towards fat oxidation [34]. Practical application of the carbohydrate models may be challenged by people who do not use sRPE-TL, however this is among the easiest measures for someone to record, with the caveat that a validated scale is used [16]. A less optimal, yet potentially viable option for athletes who only use TSS would be to convert prior-day TSS values into estimated sRPE-TL using the regression equation -65.4 + (TSS * 5.18). Details of this equation are provided in Supplemental file 2, which is based on the strong relationship ($r = 0.86$) between TSS and sRPE-TL observed in our data.

The use of squared training load terms are included in half of the models, allowing non-linear relationships to be modelled. The second ventilatory threshold was also relevant in nearly all models, either as $\dot{V}O_2$ at VT_2 or as a percentage of $\dot{V}O_{2max}$ which ostensibly provides the same information when combined with $\dot{V}O_{2max}$. This could be expected, as VT_2 (also called the

respiratory compensation point) represents the highest $\dot{V}O_2$ (and therefore, energy expenditure) associated with steady-state lactate concentrations in the blood [63].

A strength of the study is the inclusion of male and female athletes across a range of fitness levels. Females typically have a reduced RER during submaximal endurance exercise compared males, indicating lower relative carbohydrate and higher relative fat oxidation [64]. This is reflected in the model coefficients shown in Table 3. Although resting glycogen concentrations may vary in females across the menstrual cycle [65], most studies have found no influence of menstrual phase on substrate oxidation during exercise, particularly in the fed state [66]. Therefore, despite the mechanistic rationale, we chose not to control for menstrual cycle phase in our female participants. For studies that have found an effect of menstrual phase on substrate oxidation, this has been observed during lower, but not moderate or higher intensity exercise [67]. For athletes taking oral contraceptive pills, it is possible that the active phase may be associated with increased fat oxidation during exercise than the inactive phase [68], but again most studies have found minimal effects [66] and so we did not control for this.

There are several limitations to this study which should be considered. First, we did not account for protein oxidation in our calculations. Measures of substrate oxidation during exercise are typically interpreted based on the assumption of negligible protein oxidation, but this assumption could be invalidated in the context of protein ingestion before or during exercise due to increased gluconeogenesis, which could decrease RER irrespective of any change in fat oxidation rate via transfer of the amino group to the urea cycle [4]. However, protein content was low in our

standardized pre-exercise meal, which is itself a practical limitation of the study. We standardized the pre-exercise meal, which means it is likely that carbohydrate use will decrease when training in the fasted state, at least during lower-intensity exercise [69], and increase following a higher-carbohydrate pre-exercise meal [34]. Future studies could test the robustness of these models across different types of pre-exercise meals. In addition, these models may not be applicable to very short (< 15 min) or very long (> 90 min) exercise durations. The carbohydrate models are more likely to overestimate use during longer sessions, as there is an expected decrease in carbohydrate reliance as exercise duration extends [34]. Some error may have also been introduced in the TSS calculation, which uses functional threshold power (FTP) that is often established using 95% of the average power attained during a 20-min time-trial [18]. To reduce participant burden, we used power at VT_2 as an estimate of FTP. This approach has been used elsewhere [31], although direct comparisons have reported VT_2 power as being both higher [70] and lower [71] than FTP. For cyclists applying our models using TSS, a corrected FTP based on fitness level could be recommended [72]. Although the use of FTP as a marker of intensity domains has been questioned [73], it could be considered that the best use of FTP may be in enabling an athlete to have a measure of TSS. Finally, it is possible that RPE-based training load values recorded in a laboratory setting may not be directly applicable to outdoor training, a finding observed in some [74] but not all [75] studies.

Future studies should consider longer exercise durations (>90 min), other exercise modalities such as rowing, and the influence of different pre-exercise nutritional intakes (e.g., fasted-state

training vs. pre-exercise carbohydrate or protein intake) and starting levels of muscle glycogen to further investigate the predictive capabilities of these models.

Conclusion

This study presents a novel method for measuring and estimating both carbohydrate utilization and energy expenditure during endurance exercise using easily available measures of training load and lab-based testing. We found all measures of training load displayed very large correlations with both carbohydrate and energy expenditure during exercise, but TSS was the only measure of training load to accurately reflect similar total carbohydrate use between longer sessions at low intensity and shorter sessions at high-intensity interval training. Our prediction models can be effectively applied in running and kayaking when used with a calibration adjustment.

Statements and Declarations

Author contributions JAR, SJM, and EM conceptualized and designed the study. JAR and SH performed data collection. JAR performed the analyses and wrote the first draft of the manuscript. All authors edited and revised the manuscript and approved the final version of the manuscript.

Funding No sources of funding were used to assist in the preparation of this article.

Conflict of interest The authors declare that they have no conflicts of interest relevant to this study.

Availability of data and material Data and code are available at https://osf.io/nqsu7/

References

1. Thomas DT, Erdman KA, Burke LM. American College of Sports Medicine Joint Position Statement. Nutrition and Athletic Performance. Med Sci Sports Exerc. 2016;48(3):543-68. doi:10.1249/MSS.0000000000000852.
2. Burke LM, Hawley JA, Jeukendrup A, Morton JP, Stellingwerff T, Maughan RJ. Toward a Common Understanding of Diet-Exercise Strategies to Manipulate Fuel Availability for Training and Competition Preparation in Endurance Sport. Int J Sport Nutr Exerc Metab. 2018;28(5):451-63. doi:10.1123/ijsnem.2018-0289.
3. di Prampero PE, Ferretti G. The energetics of anaerobic muscle metabolism: a reappraisal of older and recent concepts. Respir Physiol. 1999;118(2-3):103-15. doi:10.1016/s0034-5687(99)00083-3.
4. Jeukendrup AE, Wallis GA. Measurement of substrate oxidation during exercise by means of gas exchange measurements. Int J Sports Med. 2005;26 Suppl 1:S28-37. doi:10.1055/s-2004-830512.
5. Schytz CT, Ortenblad N, Gejl KD, Nielsen J. Differential utilisation of subcellular skeletal muscle glycogen pools: a comparative analysis between 1 and 15 min of maximal exercise. J Physiol. 2024;602(8):1681-702. doi:10.1113/JP285762.
6. Esbjornsson-Liljedahl M, Sundberg CJ, Norman B, Jansson E. Metabolic response in type I and type II muscle fibers during a 30-s cycle sprint in men and women. J Appl Physiol (1985). 1999;87(4):1326-32. doi:10.1152/jappl.1999.87.4.1326.
7. Horwath O, Envall H, Roja J, Emanuelsson EB, Sanz G, Ekblom B et al. Variability in vastus lateralis fiber type distribution, fiber size, and myonuclear content along and between the legs. J Appl Physiol (1985). 2021;131(1):158-73. doi:10.1152/japplphysiol.00053.2021.
8. Beneke R, Pollmann C, Bleif I, Leithauser RM, Hutler M. How anaerobic is the Wingate Anaerobic Test for humans? Eur J Appl Physiol. 2002;87(4-5):388-92. doi:10.1007/s00421-002-0622-4.
9. Davis P, Leithauser RM, Beneke R. The energetics of semicontact 3 x 2-min amateur boxing. Int J Sports Physiol Perform. 2014;9(2):233-9. doi:10.1123/IJSPP.2013-0006.
10. Milioni F, Zagatto AM, Barbieri RA, Andrade VL, Dos Santos JW, Gobatto CA et al. Energy Systems Contribution in the Running-based Anaerobic Sprint Test. Int J Sports Med. 2017;38(3):226-32. doi:10.1055/s-0042-117722.
11. de Campos Mello F, de Moraes Bertuzzi RC, Grangeiro PM, Franchini E. Energy systems contributions in 2,000 m race simulation: a comparison among rowing ergometers and water. Eur J Appl Physiol. 2009;107(5):615-9. doi:10.1007/s00421-009-1172-9.
12. Elia M, Livesey G. Energy expenditure and fuel selection in biological systems: the theory and practice of calculations based on indirect calorimetry and tracer methods. World Rev Nutr Diet. 1992;70:68-131. doi:10.1159/000421672.
13. Impellizzeri FM, Marcora SM, Coutts AJ. Internal and External Training Load: 15 Years On. Int J Sports Physiol Perform. 2019;14(2):270-3. doi:10.1123/ijspp.2018-0935.
14. Mujika I. Quantification of Training and Competition Loads in Endurance Sports: Methods and Applications. Int J Sports Physiol Perform. 2017;12(Suppl 2):S29-S217. doi:10.1123/ijspp.2016-0403.

15. van Erp T, Foster C, de Koning JJ. Relationship between various training-load measures in elite cyclists during training, road races, and time trials. Int J Sports Physiol Perform. 2019;14(4):493-500.
16. Foster C, Boullosa D, McGuigan M, Fusco A, Cortis C, Arney BE et al. 25 Years of Session Rating of Perceived Exertion: Historical Perspective and Development. Int J Sports Physiol Perform. 2021;16(5):612-21. doi:10.1123/ijspp.2020-0599.
17. Lucia A, Hoyos J, Santalla A, Earnest C, Chicharro JL. Tour de France versus Vuelta a Espana: which is harder? Med Sci Sports Exerc. 2003;35(5):872-8. doi:10.1249/01.MSS.0000064999.82036.B4.
18. Allen H, Coggan AR, McGregor S. Training and racing with a power meter. VeloPress; 2019.
19. Vermeire KM, Caen K, Bourgois JG, Boone J. Training Load and Acute Performance Decrements Following Different Training Sessions. Int J Sports Physiol Perform. 2023;18(3):284-92. doi:10.1123/ijspp.2022-0157.
20. van Erp T, Hoozemans M, Foster C, de Koning JJ. The influence of exercise intensity on the association between kilojoules spent and various training loads in professional cycling. Int J Sports Physiol Perform. 2019;14(10):1395-400.
21. Rothschild JA, Morton JP, Stewart T, Kilding AE, Plews DJ. The quantification of daily carbohydrate periodization among endurance athletes during 12 weeks of self-selected training: presentation of a novel Carbohydrate Periodization Index. medRxiv. 2022.
22. von Elm E, Altman DG, Egger M, Pocock SJ, Gotzsche PC, Vandenbroucke JP et al. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. BMJ. 2007;335(7624):806-8. doi:10.1136/bmj.39335.541782.AD.
23. Collins GS, Moons KGM, Dhiman P, Riley RD, Beam AL, Van Calster B et al. TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. BMJ. 2024;385:e078378. doi:10.1136/bmj-2023-078378.
24. Borg G. Borg's perceived exertion and pain scales. Human kinetics; 1998.
25. Ballmann CG. The Influence of Music Preference on Exercise Responses and Performance: A Review. J Funct Morphol Kinesiol. 2021;6(2). doi:10.3390/jfmk6020033.
26. Fanchini M, Ferraresi I, Modena R, Schena F, Coutts AJ, Impellizzeri FM. Use of CR100 Scale for Session Rating of Perceived Exertion in Soccer and Its Interchangeability With the CR10. Int J Sports Physiol Perform. 2016;11(3):388-92. doi:10.1123/ijspp.2015-0273.
27. Cerezuela-Espejo V, Hernandez-Belmonte A, Courel-Ibanez J, Conesa-Ros E, Mora-Rodriguez R, Pallares JG. Are we ready to measure running power? Repeatability and concurrent validity of five commercial technologies. Eur J Sport Sci. 2021;21(3):341-50. doi:10.1080/17461391.2020.1748117.
28. Maturana FM. whippr: Tools for Manipulating Gas Exchange Data. R package version 0.1.2 ed2022.
29. Livesey G, Elia M. Estimation of energy expenditure, net carbohydrate utilization, and net fat oxidation and synthesis by indirect calorimetry: evaluation of errors with special reference to the detailed composition of fuels. Am J Clin Nutr. 1988;47(4):608-28. doi:10.1093/ajcn/47.4.608.

30. Mookerjee SA, Gerencser AA, Nicholls DG, Brand MD. Quantifying intracellular rates of glycolytic and oxidative ATP production and consumption using extracellular flux measurements. J Biol Chem. 2017;292(17):7189-207. doi:10.1074/jbc.M116.774471.

31. Vermeire KM, Van de Casteele F, Gosseries M, Bourgois JG, Ghijs M, Boone J. The Influence of Different Training Load Quantification Methods on the Fitness-Fatigue Model. Int J Sports Physiol Perform. 2021;16(9):1261-9. doi:10.1123/ijspp.2020-0662.

32. Bakdash JZ, Marusich LR. Repeated Measures Correlation. Front Psychol. 2017;8:456. doi:10.3389/fpsyg.2017.00456.

33. Hopkins W. Spreadsheets for analysis of validity and reliability. Sportscience. 2015;19:36-42.

34. Rothschild J, Kilding AE, Stewart T, Plews DJ. Factors influencing substrate oxidation during submaximal cycling: a modelling analysis. Sports Med. 2022;52:2775–95.

35. McNeish D. A practical guide to selecting and blending approaches for clustered data: Clustered errors, multilevel models, and fixed-effect models. Psychol Methods. 2023. doi:10.1037/met0000620.

36. Pan W. Akaike's information criterion in generalized estimating equations. Biometrics. 2001;57(1):120-5. doi:10.1111/j.0006-341x.2001.00120.x.

37. Calcagno V, de Mazancourt C. glmulti: an R package for easy automated model selection with (generalized) linear models. J Stat Softw. 2010;34(12):1-29.

38. Molinaro AM, Simon R, Pfeiffer RM. Prediction error estimation: a comparison of resampling methods. Bioinformatics. 2005;21(15):3301-7. doi:10.1093/bioinformatics/bti499.

39. Collins GS, Dhiman P, Ma J, Schlussel MM, Archer L, Van Calster B et al. Evaluation of clinical prediction models (part 1): from development to external validation. BMJ. 2024;384:e074819. doi:10.1136/bmj-2023-074819.

40. Binuya MAE, Engelhardt EG, Schats W, Schmidt MK, Steyerberg EW. Methodological guidance for the evaluation and updating of clinical prediction models: a systematic review. BMC Med Res Methodol. 2022;22(1):316. doi:10.1186/s12874-022-01801-8.

41. Riley RD, Ensor J, Snell KIE, Harrell FE, Jr., Martin GP, Reitsma JB et al. Calculating the sample size required for developing a clinical prediction model. BMJ. 2020;368:m441. doi:10.1136/bmj.m441.

42. Riley RD, Snell KIE, Ensor J, Burke DL, Harrell FE, Jr., Moons KGM et al. Minimum sample size for developing a multivariable prediction model: Part I - Continuous outcomes. Stat Med. 2019;38(7):1262-75. doi:10.1002/sim.7993.

43. Archer L, Snell KIE, Ensor J, Hudda MT, Collins GS, Riley RD. Minimum sample size for external validation of a clinical prediction model with a continuous outcome. Stat Med. 2021;40(1):133-46. doi:10.1002/sim.8766.

44. Impellizzeri FM, Jeffries AC, Weisman A, Coutts AJ, McCall A, McLaren SJ et al. A clarification of the (mis) use of the term 'load'in sport and exercise science: why it is appropriate and scientific. SportRxiv. 2021. doi:https://doi.org/10.51224/SRXIV.16.

45. Andreato LV. High-Intensity Interval Training: Methodological Considerations for Interpreting Results and Conducting Research. Trends Endocrinol Metab. 2020;31(11):812-7. doi:10.1016/j.tem.2020.08.003.

46. Impey SG, Hearris MA, Hammond KM, Bartlett JD, Louis J, Close GL et al. Fuel for the work required: A theoretical framework for carbohydrate periodization and the glycogen threshold hypothesis. Sports Med. 2018;48(5):1031-48.

47. Rothschild JA, Islam H, Bishop DJ, Kilding AE, Stewart T, Plews DJ. Factors Influencing AMPK Activation During Cycling Exercise: A Pooled Analysis and Meta-Regression. Sports Med. 2022;52(6):1273-94. doi:10.1007/s40279-021-01610-x.

48. Helgerud J, Hoydal K, Wang E, Karlsen T, Berg P, Bjerkaas M et al. Aerobic high-intensity intervals improve VO2max more than moderate training. Med Sci Sports Exerc. 2007;39(4):665-71. doi:10.1249/mss.0b013e3180304570.

49. Gormley SE, Swain DP, High R, Spina RJ, Dowling EA, Kotipalli US et al. Effect of intensity of aerobic training on VO2max. Med Sci Sports Exerc. 2008;40(7):1336-43. doi:10.1249/MSS.0b013e31816c4839.

50. Reuter M, Rosenberger F, Barz A, Venhorst A, Blanz L, Hecksteden A et al. Does Higher Intensity Increase the Rate of Responders to Endurance Training When Total Energy Expenditure Remains Constant? A Randomized Controlled Trial. Sports Med - Open. 2023;9(1):35.

51. Maturana FM, Schellhorn P, Erz G, Burgstahler C, Widmann M, Munz B et al. Individual cardiovascular responsiveness to work-matched exercise within the moderate- and severe-intensity domains. Eur J Appl Physiol. 2021;121(7):2039-59. doi:10.1007/s00421-021-04676-7.

52. Brooks GA. Bioenergetics of exercising humans. Compr Physiol. 2012;2(1):537-62. doi:10.1002/cphy.c110007.

53. Watt MJ, Heigenhauser GJ, Dyck DJ, Spriet LL. Intramuscular triacylglycerol, glycogen and acetyl group metabolism during 4 h of moderate exercise in man. J Physiol. 2002;541(Pt 3):969-78. doi:10.1113/jphysiol.2002.018820.

54. Hargreaves M, McConell G, Proietto J. Influence of muscle glycogen on glycogenolysis and glucose uptake during exercise in humans. J Appl Physiol (1985). 1995;78(1):288-92. doi:10.1152/jappl.1995.78.1.288.

55. Gill PK, Kipp S, Beck ON, Kram R. It is time to abandon single-value oxygen uptake energy equivalents. J Appl Physiol (1985). 2023;134(4):887-90. doi:10.1152/japplphysiol.00353.2022.

56. Chen Z-P, McConell GK, Michell BJ, Snow RJ, Canny BJ, Kemp BE. AMPK signaling in contracting human skeletal muscle: acetyl-CoA carboxylase and NO synthase phosphorylation. Am J Physiol Endocrinol Metab. 2000;279(5):E1202-E6.

57. Birk JB, Wojtaszewski JF. Predominant alpha2/beta2/gamma3 AMPK activation during exercise in human skeletal muscle. J Physiol. 2006;577(Pt 3):1021-32. doi:10.1113/jphysiol.2006.120972.

58. Lee-Young RS, Palmer MJ, Linden KC, LePlastrier K, Canny BJ, Hargreaves M et al. Carbohydrate ingestion does not alter skeletal muscle AMPK signaling during exercise in humans. Am J Physiol Endocrinol Metab. 2006;291(3):E566-E73.

59. Jensen R, Ortenblad N, Stausholm MH, Skjaerbaek MC, Larsen DN, Hansen M et al. Heterogeneity in subcellular muscle glycogen utilisation during exercise impacts endurance capacity in men. J Physiol. 2020;598(19):4271-92. doi:10.1113/JP280247.

60. Steyerberg EW, Harrell FE, Jr. Prediction models need appropriate internal, internal-external, and external validation. J Clin Epidemiol. 2016;69:245-7. doi:10.1016/j.jclinepi.2015.04.005.

61. Stables RG, Kasper AM, Sparks SA, Morton JP, Close GL. An Assessment of the Validity of the Remote Food Photography Method (Termed Snap-N-Send) in Experienced and Inexperienced Sport Nutritionists. Int J Sport Nutr Exerc Metab. 2021;31(2):125-34. doi:10.1123/ijsnem.2020-0216.

62. Olaya-Cuartero J, Pueo B, Penichet-Tomas A, Villalon-Gasch L. Does Power Output at Critical Power Intensity Interchange between Cycling and Running? Applied Sciences. 2023;13(9):5511.

63. Keir DA, Iannetta D, Mattioni Maturana F, Kowalchuk JM, Murias JM. Identification of Non-Invasive Exercise Thresholds: Methods, Strategies, and an Online App. Sports Med. 2022;52(2):237-55. doi:10.1007/s40279-021-01581-z.

64. Tarnopolsky MA. Gender differences in substrate metabolism during endurance exercise. Can J Appl Physiol. 2000;25(4):312-27. doi:10.1139/h00-024.

65. McLay RT, Thomson CD, Williams SM, Rehrer NJ. Carbohydrate loading and female endurance athletes: effect of menstrual-cycle phase. Int J Sport Nutr Exerc Metab. 2007;17(2):189-205. doi:10.1123/ijsnem.17.2.189.

66. D'Souza AC, Wageh M, Williams JS, Colenso-Semple LM, McCarthy DG, McKay AKA et al. Menstrual cycle hormones and oral contraceptives: a multimethod systems physiology-based review of their impact on key aspects of female physiology. J Appl Physiol (1985). 2023;135(6):1284-99. doi:10.1152/japplphysiol.00346.2023.

67. Hackney AC, McCracken-Compton MA, Ainsworth B. Substrate responses to submaximal exercise in the midfollicular and midluteal phases of the menstrual cycle. Int J Sport Nutr. 1994;4(3):299-308. doi:10.1123/ijsn.4.3.299.

68. Wageh M, Williams JS, McCarthy DG, Tarnopolsky MA, MacDonald MJ, Gibala MJ et al. Hormonal Cycling And Substrate Oxidation At Rest And During Exercise: A Systematic Review And Meta-analysis: 473. Med Sci Sports Exerc. 2022;54(9S):115.

69. Rothschild JA, Kilding AE, Broome SC, Stewart T, Cronin JB, Plews DJ. Pre-Exercise Carbohydrate or Protein Ingestion Influences Substrate Oxidation but Not Performance or Hunger Compared with Cycling in the Fasted State. Nutrients. 2021;13(4). doi:10.3390/nu13041291.

70. Barranco-Gil D, Gil-Cabrera J, Valenzuela PL, Alejo LB, Montalvo-Perez A, Talavera E et al. Functional Threshold Power: Relationship With Respiratory Compensation Point and Effects of Various Warm-Up Protocols. Int J Sports Physiol Perform. 2020:1-5. doi:10.1123/ijspp.2019-0402.

71. Sitko S, Cirer-Sastre R, Corbi F, López-Laval I. Relationship between functional threshold power, ventilatory threshold and respiratory compensation point in road cycling. J Sports Med Phys Fitness. 2021;62(5):626-32.

72. Sitko S, Cirer-Sastre R, Lopez-Laval I. An Update Of The Allen & Coggan Equation To Predict 60-Min Power Output In Cyclists Of Different Performance Levels. Int J Sports Med. 2023;44(13):983-7. doi:10.1055/a-2079-1363.

73. Wong S, Burnley M, Mauger A, Fenghua S, Hopker J. Functional threshold power is not a valid marker of the maximal metabolic steady state. J Sport Sci. 2022;40(23):2578-84.

74. Jeffries O, Waldron M, Patterson SD, Galna B. An Analysis of Variability in Power Output During Indoor and Outdoor Cycling Time Trials. Int J Sports Physiol Perform. 2019:1273-9. doi:10.1123/ijspp.2018-0539.

75. Mieras ME, Heesch MW, Slivka DR. Physiological and psychological responses to outdoor vs. laboratory cycling. J Strength Cond Res. 2014;28(8):2324-9. doi:10.1519/JSC.0000000000000384.

Supplemental Table 1. Sample size calculations

	R2 apparent	Model parameters	R2 adjusted	Minimum sample size
Kcal				
TWD-kJ	0.96	3	0.95	7
TSS	0.96	3	0.95	7
TSS-HR	0.94	6	0.89	10
Lucia TRIMP power	0.93	4	0.90	11
Lucia TRIMP HR	0.93	5	0.89	11
sRPE-TL	0.93	5	0.89	11
CHO				
TSS-HR	0.93	5	0.89	11
TWD-kJ	0.92	4	0.89	13
TSS	0.92	6	0.86	13
Lucia TRIMP power	0.89	5	0.83	18
Lucia TRIMP HR	0.90	6	0.83	16
sRPE-TL	0.88	5	0.81	19

CHO = carbohydrate, R2 apparent = R2 from initial model, R2 adjusted = Adjusted R2 based on sample size and model parameters. Minimum sample size calculated as minimum data points required divided by two because each participant provides two data points.

Examples from published literature

This approach to energy and carbohydrate calculations can be 'stress-tested' using previously published research which has measured muscle glycogen and/or other contributors to energy expenditure.

Assumptions

- To convert from kcal to grams of carbohydrate, consideration of the carbohydrate source is necessary, because the energy yield is 3.719 kcal per gram from glucose and 4.187 kcal per gram from glycogen (Livesey and Elia 1988).
- Glycogen was converted from dry weight to wet weight using the conversion factor (0.23) of Areta and Hopkins (2018), and to grams of carbohydrate assuming 45% muscle mass and 25% active muscle mass (Medbo and Tabata 1993). A molecular weight of glycogen of 162 was used in the conversion (Schaubroeck 2022). Conversion steps are shown as follows:
  1) Glycogen (mmol/kg wet weight) = Glycogen (mmol/kg wet weight) * 0.23
  2) Activated muscle mass = body mass (kg) * 0.45 * 0.25
  3) Glycogen (mmol) = mmol/kg wet weight * Activated muscle mass (kg)
  4) Glycogen (g) = Glycogen (mmol) * (162 / 1000)

Romijn et al (1993) used stable isotope tracers and indirect calorimetry were to evaluate the regulation of endogenous fat and glucose metabolism in relation to exercise intensity and duration.

Extracting Fig. 8 allows us to use these values in calculating estimated energy expenditure.
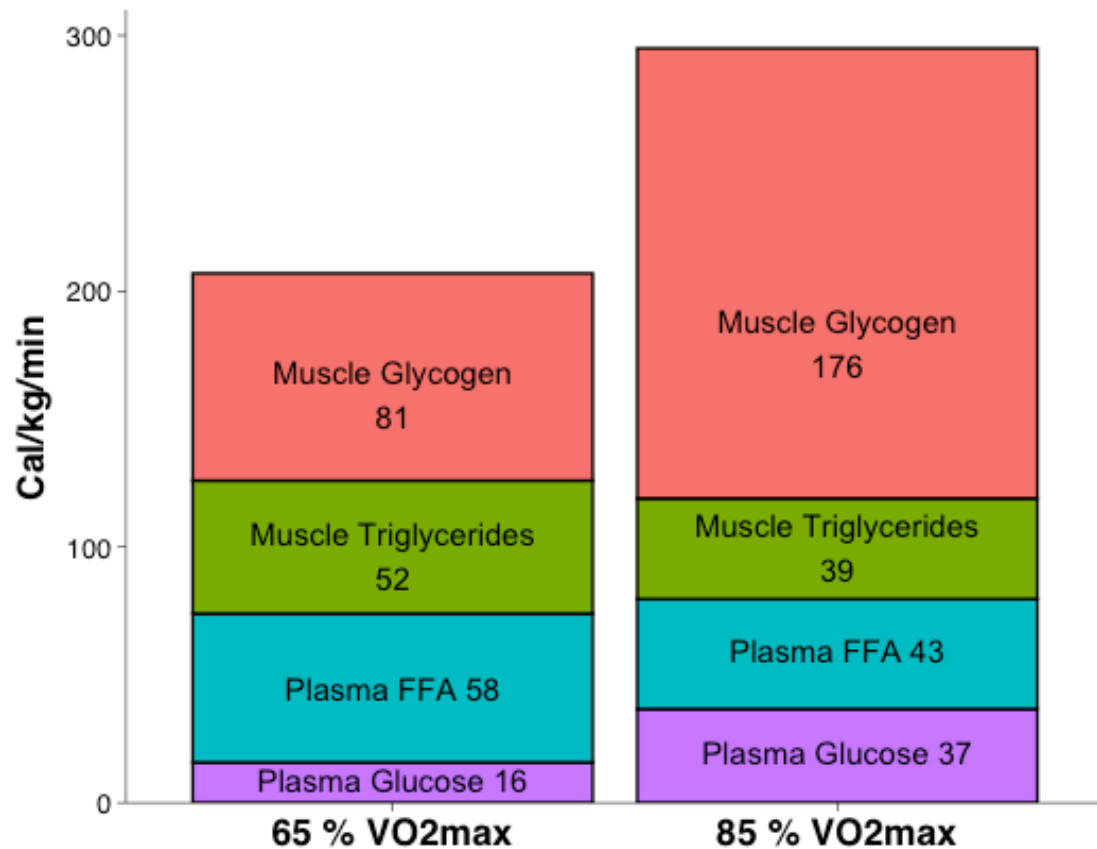
Extracted from Romijn et al (1993) Fig. 8

30-min at 85% VO$_{2max}$

Using data from Romijn et al (1993) we can estimate as follows based on the 30 min session at 85% VO$_{2max}$.

Source	Cal Kg Min	Body Weight	Kcal Min	Session Duration	Session Kcal
Plasma Glucose	36.7	75.2	2.76	30	82.8
Plasma FFA	43.0	75.2	3.23	30	96.9
Muscle Triglycerides	39.3	75.2	2.96	30	88.8
Muscle Glycogen	176.0	75.2	13.24	30	397.2

Summing everything together gives an estimate of **666 total kcal**, and summing together the contributions from glucose and glycogen after dividing by the appropriate yields of kcal/g from glucose (3.719) and glycogen (4.187), gives **117 g carbohydrate**.

We can also estimate the session expenditure using the summed aerobic (VO_2) + anaerobic (lactate) systems approach, assuming 3.77 mmol lactate accumulation during the 30 min of cycling (reported at 85% VO_{2max} in Romijn 1992), and an average RER value of 0.91.

```
romijn_delta_lactate <-  4.55- 0.78 # based on Romijn 1992 at 85% VO2max

romij_vo2_calc <- tibble(
    vo2_max = vo2max_romijn, # 5.04 L, reported in paper
    pct_max = .85,  # reported in paper
    vo2 = pct_max * vo2_max, # get absolute VO2
    rer = .91  # reported in paper
 ) %>%

 # get RER-adjusted VO2-kcal conversions from Elia and Livesey 1992
    left_join(kj_lookup_tbl, by = "rer") %>%

 # calculate kcal and carb for 1 minute
 mutate(
    kcal_from_vo2 = vo2 * kcal_L ,
    kcal_from_carb = pct_carb/100 * kcal_from_vo2,
    pct_from_glucose = 1- pct_max,  # intensity-adjusted
    yield_from_glucose = pct_from_glucose * glucose_yield, #glucose yields
3.719 kcal/g
    yield_from_glycogen = pct_max * glycogen_yield,  #glycogen yields 4.187
kcal/g
    glucose_glycogen_yield = yield_from_glucose + yield_from_glycogen,
    carb_ox_g_min = kcal_from_carb/glucose_glycogen_yield
 )
```

These calculations are reflected below:

	Value
vo2_max	5.04
pct_max	0.85
vo2	4.28
rer	0.91
kJ_L	20.65
kcal_L	4.94
pct_carb	70.53
kcal_from_vo2	21.14
kcal_from_carb	14.91
pct_from_glucose	0.15
yield_from_glucose	0.56
yield_from_glycogen	3.56
glucose_glycogen_yield	4.12
carb_ox_g_min	3.62

Multiplying the per-minute values by 30 and adding the anaerobic component:

```r
romij_vo2_lactate_calc <- romij_vo2_calc %>%
  #multiply by 30 for 30 min session
  summarise(
    aerobic_carb_g = carb_ox_g_min * 30,
    aerobic_kcal = kcal_from_vo2 * 30
  ) %>%
  # calculate anaerobic energy expenditure using net lactate accumulation
method
  mutate(
    ml_o2_lactate = romijn_delta_lactate * 3 * bw_romijn,
    lactate_kj = ml_o2_lactate/1000 * 21.1,
     # kJ to kcal to grams * 11.845 to account for anaerobic inefficiency
    carb_from_lacate_g = lactate_kj/4.184/4.187  * 11.845,  #
    total_carb_g = round(carb_from_lacate_g + aerobic_carb_g,0),
    total_kcal = round(aerobic_kcal + (carb_from_lacate_g *
glycogen_yield),0),
    estimate = "Romijn VO2-lactate calculation",
  )
```

The above calculations result in the following values:

```
## # A tibble: 1 × 4
##   `Aerobic CHO (g)` `CHO (g) from lactate` `Aerobic kcal` `Total kcal`
##              <dbl>                  <dbl>          <dbl>        <dbl>
## 1              109                     12            634          685
```

When compared together they display good agreement.

Estimate	Total kcal	Total CHO (g)
Romijn additive calculation	666	117
Romijn VO2-lactate calculation	685	121

Wingate tests

We also calculated kcal and carbohydrate use from a study reporting muscle glycogen reductions during a single 30-s Wingate sprint (Esbjornsson-Liljedahl et al 1999), which also reported data for males and females separately.

We extracted VO_2 data from a separate study (Beneke et al 2002), adjusting for VO_{2max} to match each group

To estimate contributions from plasma glucose and muscle triglycerides we adapted the numbers from Romijn 1993, and used each group's reported values of muscle glycogen breakdown. Considering both the intensity and short duration, FFA contribution was deemed negligible.
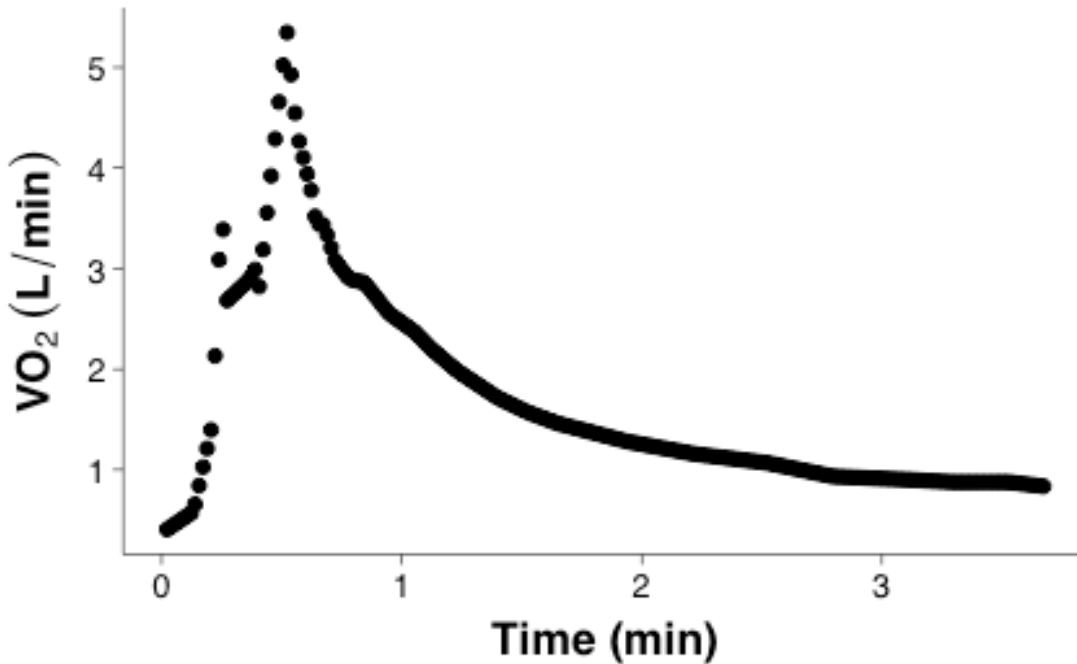
Extracted from Beneke at al (2002) Fig. 1

This study also reported fiber-type specific glycogen reduction, which was then adjusted based on relative area of fiber types reported in Table 1 of the paper. We assumed for average RER value of 1.0 during the 3.5 min window.

```
esbj_session_duration <- 3.5
wingate_glucose_cal_kg_min <- 36.7
wingate_tg_cal_kg_min <- 39.3
wingate_avg_rer <- 1
beneke_vo2max <- 4.22

Esbj_additive_calcs <- tibble(
  sex = c("m", "f"),
  mass = c(75, 65),
  session_duration = esbj_session_duration,
  fat_cal_min = wingate_tg_cal_kg_min,
  fat_kcal = fat_cal_min * mass * session_duration/1000,
  glucose_cal_min =  wingate_glucose_cal_kg_min,
  glucose_kcal = glucose_cal_min * mass * session_duration/1000,
  glucose_g = glucose_kcal/ glucose_yield,
  quad_mass = mass * .45 * .25,
  type1pct = c(56, 66)/100,
  type2pct = 1-type1pct,
  glycogen_delta_type1 = c((452-326), 428-355),
  glycogen_delta_type2 = c((526-395), (542-393)),
  weighted_glycogen_delta_dm = (glycogen_delta_type1 * type1pct +
glycogen_delta_type2 * type2pct),
```

```
  weighted_glycogen_delta_wm = weighted_glycogen_delta_dm * .23,
  glycogen_g = convert_glycogen_to_grams(weighted_glycogen_delta_wm,
quad_mass),
  glycogen_kcal = glycogen_g * glycogen_yield,
  total_kcal = fat_kcal + glucose_kcal + glycogen_kcal,
  total_carb_g = glycogen_g + glucose_g,
  estimate = c("Esbjornsson additive calculation - males", "Esbjornsson
additive calculation - females"),
)
```

These calculations result in the following values:

Calculation	M	F
mass	75.00	65.00
session_duration	3.50	3.50
fat_cal_min	39.30	39.30
fat_kcal	10.32	8.94
glucose_cal_min	36.70	36.70
glucose_kcal	9.63	8.35
glucose_g	2.59	2.25
quad_mass	8.44	7.31
type1pct	0.56	0.66
type2pct	0.44	0.34
glycogen_delta_type1	126.00	73.00
glycogen_delta_type2	131.00	149.00
weighted_glycogen_delta_dm	128.20	98.84
weighted_glycogen_delta_wm	29.49	22.73
glycogen_g	40.30	26.93
glycogen_kcal	168.75	112.76
total_kcal	188.70	130.05
total_carb_g	42.89	29.18

```
Esbj_lactate_calcs <- tibble(
  sex = c("m", "f"),
  mass = c(75, 65),
  lactate_delta = c(11.8- 1.46, 9.35-1.35),
  ml_o2_lactate = 3 * lactate_delta * mass,
  lactate_kj = ml_o2_lactate/1000 * 21.1,
  carb_from_lacate_g = round(lactate_kj/4.184/4.187  * 11.845,1)
)

Esbj_lactate_calcs

## # A tibble: 2 × 6
##   sex    mass lactate_delta ml_o2_lactate lactate_kj carb_from_lacate_g
```

```
##    <chr> <dbl>       <dbl>       <dbl>    <dbl>          <dbl>
## 1 m         75        10.3        2326.    49.1           33.2
## 2 f         65         8          1560     32.9           22.3
```

```r
# Adjust VO2 values for each group's VO2max
Esbj_vo2max_tbl <- tibble(
  sex = c("m", "f"),
  mass = c(75, 65),
  est_vo2max = c(55, 45),
  vo2max_l = est_vo2max * mass/ 1000,
  adjustment_factor_from_beneke = vo2max_l/ beneke_vo2max   # adjust for
group-specific VO2max values
)


vo2_calcs_tbl <- wingate_df %>%
  interpolate() %>%
  mutate(
    m = vo2 * Esbj_vo2max_tbl$adjustment_factor_from_beneke[1],
    f = vo2 * Esbj_vo2max_tbl$adjustment_factor_from_beneke[2],
    rer = wingate_avg_rer
  ) %>% select(-vo2) %>%

  pivot_longer(m:f, names_to = "sex", values_to = "vo2") %>%
  nest(.by = sex) %>%
  left_join(Esbj_vo2max_tbl %>% select(sex, vo2max_l), by = "sex") %>%
  unnest(data) %>%
  mutate(pct_max = vo2/vo2max_l) %>%
  nest(.by = sex) %>%
  mutate(
    # apply second-by-second aerobic energy calculations function
    totals = map(data, calculate_carb_kcal_fn)
  ) %>% unnest(totals) %>% select(-data)   %>%
  left_join(Esbj_lactate_calcs, by = "sex") %>%
  mutate(
    estimate = c("Esbjornsson VO2-lactate calculation - males", "Esbjornsson
VO2-lactate calculation - females"),
    total_kcal = aerobic_kcal + carb_from_lacate_g * glycogen_yield,
    total_carb_g = aerobic_carb_g + carb_from_lacate_g, .keep = "unused"
  ) %>% select(estimate, total_kcal, total_carb_g)
```

Good agreement is seen for both males and females

| Estimate	Total kcal	Total CHO (g)
Esbjornsson additive calculation - males	189	43
Esbjornsson VO2-lactate calculation - males	169	41
Estimate	Total kcal	Total CHO (g)
Esbjornsson additive calculation - females	130	29
Esbjornsson VO2-lactate calculation - females	115	28

Converting TSS to sRPE-TL

For people wishing to use the carbohydrate models who don't have access to prior-day sRPE-TL data, prior-day TSS could be substituted by using a conversion to estimated sRPE-TL (albeit with some acknowledged trade-offs in terms of accuracy).

Using data combined from the primary and validation studies, there is a strong correlation ($r = 0.86$) between TSS and sRPE-TL. Therefore, a regression line can be estimated and used to convert prior day TSS to a value of sRPE-TL which can then be used in the prediction models.

```
tss_convert1 <- lmer(sRPE_TL ~ TSS + (1 | subject_id), mod_refit_tbl)

tss_convert1

## Linear mixed model fit by REML ['lmerModLmerTest']
## Formula: sRPE_TL ~ TSS + (1 | subject_id)
##    Data: mod_refit_tbl
## REML criterion at convergence: 1973.78
## Random effects:
##  Groups     Name        Std.Dev.
##  subject_id (Intercept) 41.56
##  Residual               75.91
## Number of obs: 169, groups:  subject_id, 70
## Fixed Effects:
## (Intercept)          TSS
##     -65.389        5.177
```
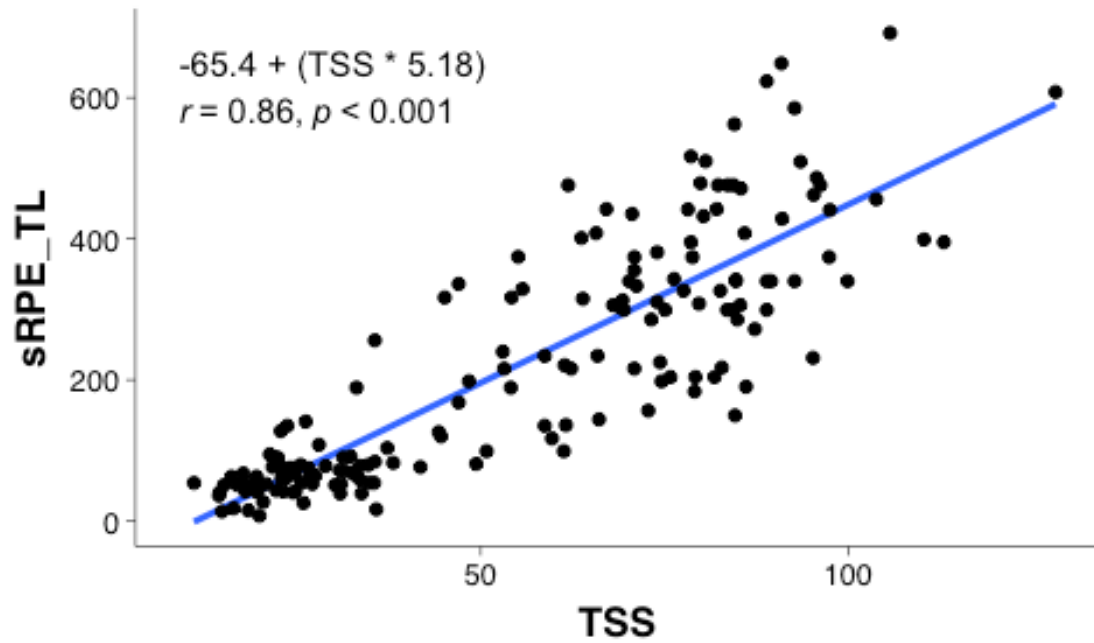
So for example if you have a TSS of 100, use the equatoion -65.4 + (TSS * 5.18) to estimate a value of 453.