ORIGINALIA

ACCEPTED: January 2022

Mentz L, Winkert K, Schlegel S, Rohlmann F, Steinacker JM, Treff G. Variability of mechanical power output in elite rowers during ergometer testing. Dtsch Z Sportmed. 2022; 73: 61-69. doi:10.5960/dzsm.2021.517

Mentz L[1], Winkert K[1], Schlegel S[2], Rohlmann F[2], Steinacker JM[1], Treff G[3]

Variability of Mechanical Power Output in Elite Rowers during Ergometer Testing

Variabilität der Zyklusleistung von Elite-Ruderern bei Ergometertests

1. UNIVERSITY HOSPITAL ULM, *Division of Sports and Rehabilitation Medicine, Ulm, Germany*
2. ULM UNIVERSITY, *Institute of Epidemiology and Medical Biometry, Ulm, Germany*
3. PARACELSUS MEDICAL UNIVERSITY,*University Institute of Sports Medicine, Prevention and Rehabilitation, Salzburg, Austria*

Summary

- **Purpose:** Incremental step tests are defined by constant duration, increment, and constant mechanical power output (P) within each step. P is not externally adjustable on air-braked rowing ergometers like the Concept-2 (C2), hence variability of P depends on the rower. The actual variability of P is currently unclear. To this end, we evaluated it during rowing ergometer testing in elite rowers.
- **Methods:** Twenty-three elite rowers of the German National squad (10 female; age: 24.5±3.5 years) performed an incremental step test on a C2 rowing-ergometer modified with an external reference system (REF) allowing evaluation of stroke-to-stroke variability during rowing ergometer testing.
- **Results:** The coefficient of variation (CV) for P_{REF} was 4.7% (95% CI [4.3, 5.3]). CV was significantly lower in male vs. female rowers (4.0% [3.4, 4.6] vs. 5.5% [5.0, 6.1]; $p < 0.001$). When the first three strokes were excluded, total CV decreased to 3.0% [2.7, 3.3] and variability was similar between male vs. female (3.0% [2.7, 3.3] vs. 3.0% [2.6, 3.3]; $p = 0.822$). The highest differences between individual CVs were found in step 1, ranging from 2.9% to 11.8%.
- **Conclusion:** Stroke-to-stroke variability during rowing ergometer testing amounts to 4.7%, which is probably higher than during testing on externally controllable bicycle ergometers, where resistance is controlled as a function of cadence. We observed considerable inter-individual differences even within this homogenous groups of elite rowers.

KEY WORDS:
High-Performance, Concept 2, Indoor Rowing, Step Test

Zusammenfassung

- **Einleitung:** Stufentests sind durch konstante Dauer, Stufenlänge und eine konstante mechanische Leistung (P) innerhalb jeder Stufe definiert. P ist bei luftgebremsten Ruderergometern wie dem Concept-2 (C2) nicht extern justierbar, so dass die Variabilität P vom Ruderer abhängt. Die tatsächliche Variabilität von P ist derzeit unklar. Zu diesem Zweck haben wir sie bei Elite-Ruderern untersucht.
- **Methoden:** Dreiundzwanzig Elite-Ruderer der deutschen Nationalmannschaft (10 weiblich; Alter: 24,5±3,5 Jahre) führten einen Stufentest auf einem C2 durch, das mit einem externen Referenzsystem (REF) modifiziert wurde, um die Schlag-zu-Schlag-Variabilität während des Tests zu evaluieren.
- **Ergebnisse:** Der Variationskoeffizient (CV) für P_{REF} betrug 4,7% (95 % CI [4.3, 5.3]). Der CV war bei männlichen Ruderern signifikant niedriger als bei weiblichen (4.0% [3.4, 4.6] vs. 5.5% [5.0, 6.1]; $p < 0.001$). Wenn die ersten drei Schläge von der Analyse ausgeschlossen wurden, sank der Gesamt-CV auf 3.0% [2.7, 3.3] und die Variabilität war bei Männern und Frauen ähnlich (3.0 % [2.7, 3.3] vs. 3.0% [2.6, 3.3]; $p = 0.822$). Die individuell größte Spannweite der CVs war in Stufe 1 (2,9%–11,8 %) zu beobachten.
- **Schlussfolgerung:** Die Schlag-zu-Schlag-Variabilität bei Ruderergometertests beträgt 4.7% und liegt damit wahrscheinlich höher als bei Fahrradergometern, bei denen der Widerstand in Abhängigkeit von der Trittfrequenz automatisch gesteuert wird. Selbst innerhalb dieser homogenen Gruppe von Elite-Ruderern wurden erhebliche interindividuelle Unterschiede beobachtet.

SCHLÜSSELWÖRTER:
Leistungssport, Concept 2, Indoor-Rudern, Stufentest

CORRESPONDING ADDRESS:
Lennart Mentz
Division of Sports and Rehabilitation Medicine
University Hospital Ulm
Leimgrubenweg 14
89075 Ulm
✉: lennart.mentz@uni-ulm.de

Introduction

In elite sports like Olympic rowing even minor changes in performance can be meaningful, because racing times vary by only 1% and smallest worthwhile changes are quantified as low as 0.3% (11). To assess such small changes in physical performance, rowing ergometers are widely used, because they allow for measurements within a controllable laboratory environment that markedly reduces measurement error and noise (19–21). Notably, rowing ergometer tests demonstrate a valuable transferability to on-water rowing (4, 26).

The probably most common used rowing ergometer is the air-braked Concept 2 (C2) (Concept 2, Morrisville, USA). Like with every air-braked ergometer, resistance increases with the amount of circulated air that results out of the rower's acceleration of

›

the flywheel. Consequently, mechanical power output (P) is not externally adjustable and the accuracy of actual vs. targeted P mainly depends on the rower's skills to match the target P. This is a relevant difference to e.g., electromagnetically braked cycling ergometers, where continuous control of the torque as a function of the actual cadence allows to keep P virtually constant. Furthermore, the C2 underestimates P by approximately 14–25 W (~ 6.8–7.4%) (7, 14). This error depends, amongst other things, on the variability of stroke-to-stroke power in individual rowers. Thus, a steady pace with low stroke to stroke variations in intensity and pace seems to be a necessity when testing elite rowers on the C2.

Aside from specific tests like the 2000 m (12), ramp (23), 1-h, and 20-s all out tests, stepwise incremental tests have been established in several international rowing programs to assess basic endurance performance, lactate-based "thresholds", and maximal performance in a time-efficient manner, thereby enabling training prescription and evaluation of changes in performance (1). Briefly, incremental step tests consist of a sequence of steps of constant duration that increase by a given constant, e.g., 50 W each 4 min. Typically, blood samples are collected after completion of each step, assuming a constant P within each step to allow for steady state conditions during low intensities and allowing to mirror the exponential increase in blood lactate above the first lactate "threshold" (27).

It is noteworthy that data on variability of P in rowing ergometer tests are scarce. Shimoda et al. (14) calculated a stroke-to-stroke consistency between 1.7±0.5% and 3.0±0.9% in 2-min steps and reported a decreasing variability with increasing P. However, this study was limited by a nowadays outdated C2 ergometer model (model C) and the lack of a reference system. Smith and Spinks (19) applied an external reference system, but used an unusual wheeled rowing ergometer and they did not report variability within steps, which is also a limitation of the study conducted by Boyas et al. (7). In addition, none of the aforementioned studies evaluated whether variability differs between male and female elite rowers. This is at least conceivable, due to well documented differences in muscle mass, especially in the upper body (9, 15) and differences in length ratios of trunk and extremities (5, 6). These differences potentially influence the force generation during the rowing stroke, which is the result of simultaneous or sequenced work of legs and trunk, leading to a rather rectangular or triangular stroke shape (13).

Finally, previous research has not evaluated whether the power estimation of the C2 accurately reflects actual variability as determined by a reference system. This is questionable, because the C2-calculaton is based on the acceleration and deceleration of the ergometer's flywheel and not on direct measurements of work per time (25).

To this end, we aimed to evaluate stroke-to-stroke variability of P in male and female elite rowers during stepwise incremental testing for the first time and to evaluate differences between sex, while assessing P with an external reference system in order to allow insights into underlying mechanisms of any observed particularities.

Material and Methods

Participants

Twenty-three elite rowers (10 female) of the German National Team participated in this observational, cross-sectional study with two sub-groups (i.e., male and female). Inclusion criterion was the active membership of the German National Team. Participants were therefore well familiarized for years with the testing procedure. All data were collected within one day in a performance laboratory of the German Rowing federation during the Olympic cycle 2017-21. To reduce pre-analytical bias, athletes followed the pre-test standardization of the German Rowing Federation, i.e. a well-balanced diet was maintained, no high-intensity training was conducted 48 h prior testing, and the last low intensity training ended 20 h prior to the test or earlier to avoid fatigue, glycogen deficiency, and hypohydration (17). All participants provided informed written consent for the scientific evaluation of their data. The study was conducted according to the declaration of Helsinki and approved by the ethical review board of the University of Ulm (472/18).

Procedures and Testing

After a standardized warm-up of 20 min with a P of 160 W (male) or 100 W (female), athletes performed a stepwise incremental test on a rowing ergometer. In line with the standard procedure of the German Rowing Federation (17), the test started with 200 W or 140 W and the increment amounted to 50 W/step or 40 W/step, in male or female rowers, respectively. Step duration was 4 min separated by a 30-s break for blood sampling. Depending on individual performance level, athletes completed 5 or 6 steps. The exact number was determined based on previous tests, targeting a blood lactate concentration of 4 mmol/L after the last step.

The drag factor, a native Concept 2 factor that is linked to the air feed of the flywheel and thereby to its behavior, was set to 145 or 130 in male or female rowers. Rowing stroke frequency was 18 strokes/min at the first step and increased by 2±1 following-strokes/min with each step. Stroke frequency was unconstrained at steps 5 and 6.

A rowing ergometer (Concept 2, Type D, Morrisville, USA), was used for all testing. The ergometer was equipped with the C2-standard PM5-monitor that calculates and displays measures of rowing performance C2-data (based on angular velocity of the flywheel) were logged stroke-by-stroke by the FLOAT app (Float, Ergstick Lmt., Cambridge, UK). Rowers were not able to see the PM5-monitor.

To allow for correct power calculation as physical measurement (i.e. work per time), the C2 was additionally equipped with an external reference system (REF) as described elsewhere (23). P of REF (P_{REF}) and C2 (P_{C2}) were logged simultaneously. Briefly, the REF-System FES Ruderergo (Institut für Forschung und Entwicklung von Sportgeraeten (FES), Berlin, Germany) uses a load cell placed between chain and handle bar to measure stroke force and an incremental encoder placed on the axis of the flywheel to measure displacement of the handlebar. According to the manufacturer, the error of measurement of the FES-set-up is equal to or smaller than 1.5%. REF-data were logged and visualized in real time by the software FES-Ruderergo Dyno v.2017 (FES, Berlin, Germany). The device was calibrated before the test according to the manufacturer's recommendations. REF data were visualized in real-time on a large computer monitor to provide feedback on stroke frequency and P_{REF} in front of the rower.

Data Analysis

The sample size was not calculated a priori, because we already acquired the maximum sample of homogeneous elite rowers in Germany for this observational study. Thus, we were forced to choose a convenience sample. For each rower, arithmetic mean ± standard deviation of the power output of all strokes at each step and for each system were calculated. Based on these data, the coefficient of variation (CV, equation 1) was calculated, representing the stroke-to-stroke variability of a single athlete at one step and for one system, i.e., REF or C2.

$$CV\% = \frac{SD}{\bar{x}} * 100 \qquad (1)$$

where SD represents standard deviation and x̄ arithmetic mean. This CV was calculated for all strokes (STROKE_{ALL}) of each step and system and without the first three strokes ($\text{STROKE}_{w/o1\text{-}3}$). The first three strokes were excluded because other studies have already shown that P_{C2} calculation of the starting strokes differs from the following strokes. This is related to the initial acceleration of the flywheel from standstill. Median CVs per step and system were presented as boxplots (Figure 1), subdivided by male and female.

To calculate statistical differences of the CVs between (i) measurement systems, (ii) steps, and (iii) sex, a mixed model structure based on a linear regression analysis with repeated measurements was applied, accounting for dependent data from repeated measurements within rowers [each rower provided two related measures (REF and C2 data) in a repeated way (steps)]. The same model was applied for the calculation of least square (LS) means of individual rowers' mean power output and individual standard deviation (accounting for the repeated structure). These LS means were tabulated per step and system for male and female. Differences in target P and actual P (i.e., P_{REF}) were also calculated by of LS means of the mixed model.

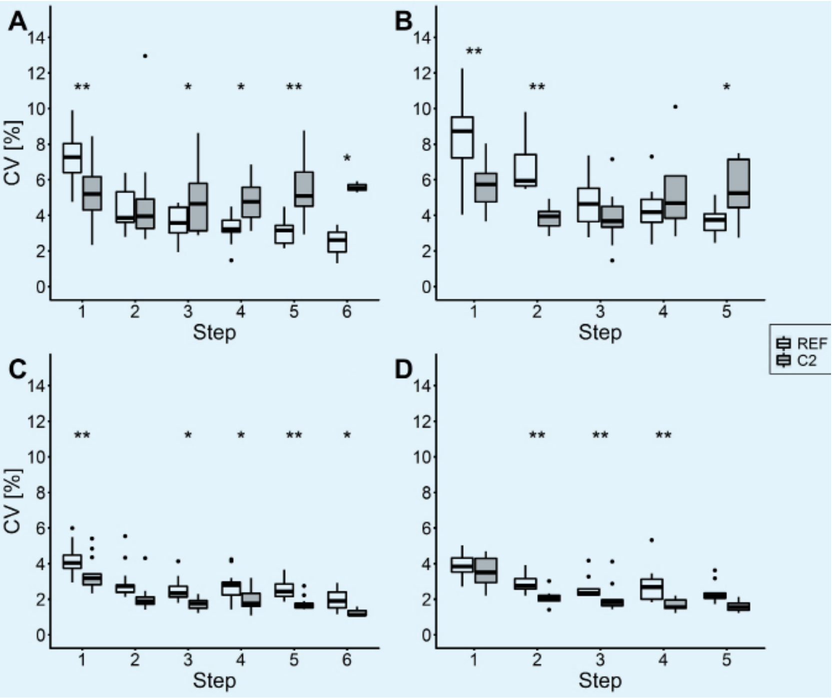

Figure 1

Coefficients of variation (CV) of stroke-to-stroke mechanical power output measured with a reference system (REF) and the Concept 2 rowing ergometer display PM5 (C2) during 1-5 steps (three rowers completed 6 steps) of incremental step testing in male (A and C) and female (B and D) elite rowers. A and B include all strokes of each step, C and D exclude strokes 1-3. * and ** indicate p<0.05 or p< 0.001, respectively, between REF and C2 at a given step.

To allow for intra-sex comparisons, we used a model applying either male or female rowers, where *step* [1-5(6)], *measurement system* (REF or C2), and the interaction effect *step*measurement system,* were used as fixed effects. For inter-sex comparisons an additional model was implemented, where the fixed effects *sex* (male or female) and the interaction effect *measurement system*sex* were added to the aforementioned effects. All models were calculated for STROKE_{ALL} and $\text{STROKE}_{w/o1\text{-}3}$.

Except for the median data shown in Figure 1, all results presented here are based on the mixed model. P-values for main effects were calculated within the models and in case of significant main effects pairwise post-hoc testing was performed via Bonferroni tests. Level of significance was set to p ≤ 0.05 for each model. Effect sizes (partial eta squared η2) were considered as small (≥ 0.01< 0.06), medium (≥ 0.06 < 0.14) and large (≥ 0.14) (8). Correlation between LS mean CV and differences between REF and C2 was assessed using simple linear regression and Pearson's correlation coefficient. For all procedures, the statistics software package SPSS 25 (IBM, Armonk, NY, SA) was applied.

Results

Descriptive Data

Table 1 shows the descriptive data of the participants.

Variability of Stroke-to-Stroke Mechanical Power Output

The coefficient of variation (CV) for P_{REF} was 4.7% (95% CI [4.3, 5.3]). Table 2 shows the variability in P (REF and C2) averaging all steps in male and female rowers including STROKE_{ALL} and $\text{STROKE}_{w/o1\text{-}3}$ in male and female rowers.

For REF, CVs were significantly different (p<0.05) between male and female rowers when taking STROKE_{ALL} into account (Table 2). CVs were also significantly different (p<0.05) between measurement systems (REF or C2) (Table 2) and between steps (p<0.001) (Table S1) for STROKE_{ALL} and $\text{STROKE}_{w/o1\text{-}3}$. Moreover, changes in CV between steps were significantly different between C2 and REF if STROKE_{ALL} were included, as indicated by the significant (p<0.001) interaction of measurement *system*step*. The interaction was not significant for $\text{STROKE}_{w/o1\text{-}3}$. The fixed effect of steps came along with large effect sizes for male and female rowers for STROKE_{ALL} and $\text{STROKE}_{w/o1\text{-}3}$, while effect sizes between measurement systems were of medium to large size. Effect sizes for sex were small to medium. Detailed f-values, p-values, and effect sizes are presented in Table 1 and Table 2.

Stroke-to-Stroke Variability

Median CVs of each step are shown in Figure 1A-B. In male rowers, the CV (95% CI [lower limit, upper limit]) based on the REF data gradually decreased from 7.3% [6.4, 8.2] (step 1) to 2.4% [0.3, 5.2] (step 6). Based on the C2-data, the CV amounted to 5.2% [4.1, 6.4] in step 1, decreased slightly to 4.7% [3.1, 6.4] in step 2, and increased continuously afterwards to 5.6% [4.8, 6.4] in step 6, thereby forming a U-shape. In female rowers the trend was the same with REF-data constantly decreasing from 8.5% [6.8, 10.1] (step 1) to 3.7% [3.1, 4.3] (step 5), while the CVs based on C2-data decreased from 5.7 % [4.8, 6.7] in step 1 to 3.8 % [3.4, 4.4] in step 2 and increased continuously afterwards to 5.5 % [4.2, 6.7] (step 5), also indicating a U-shaped progression. The highest differences between individual CVs (P_{REF}) were found in step 1, ranging 2.9% to 11.8%.

When the first three strokes were excluded from the analysis (Figure 1C-D), the CV of the first step was significantly higher (p<0.001) compared to each subsequent step in male rowers, irrespective of the measurement system. This was similar in

›

Table 1

Descriptive data of participants. P4: Power at blood lactate concentration of 4 mmol/l; 2k-TT: 2000-m-time-trial on rowing ergometer. Data is presented as mean +/- standard deviation.

SEX	N	AGE [YEARS]	HEIGHT [CM]	BODY MASS [KG]	P4 [W]	2K-TT [MIN:S]
Male	13	24.2 ± 3.4	191.5 ± 5.8	93.5 ± 5.7	393 ± 6.3	5:56.5 ± 6.8
Female	10	24.9 ± 3.6	179.7 ± 4.7	74.4 ± 5.0	290 ± 11.9	6:51.7 ± 5.9

Table 2

Stroke-to-stroke variability during incremental rowing ergometer step testing in 13 male and 10 female elite rowers. Data are least square mean coefficients of variation (CV) and 95% confidence [95%CI] intervals. For each comparison (within and between male and female), a mixed model structure was applied. See text for details. Analysis of $\text{STROKE}_{\text{ALL}}$ includes all strokes, while $\text{STROKE}_{\text{w/o1-3}}$ omits strokes 1-3 of each step. Mechanical power output was assessed either with an external reference system (REF) or via the Concept 2 ergometer's PM5 display (C2).

ANALYSIS	MALE			FEMALE			MALE VS. FEMALE	
	REF	C2	p	REF	C2	p	REF (p)	C2 (p)
$\text{STROKE}_{\text{ALL}}$	4 [3.4-4.6]	5 [4.5-5.6]	<0.001	5.5 [5.0-6.1]	4.9 [4.3-5.4]	0.022	0.002	0.655
$\text{STROKE}_{\text{W/O1-3}}$	3 [2.7-3.3]	2.2 [1.9-2.5]	<0.001	3 [2.6-3.3]	2.2 [1.8-2.5]	<0.001	0.822	0.998

Table 3

Actual mechanical power output during incremental step testing in 13 male and 10 female elite rowers measured with a reference system (REF) and the Concept 2 rowing ergometer's display PM5 (C2) versus targeted power output. Data are least square means of the individual rowers' mean power output and least square means of individual rowers' standard deviations (SD). Analysis of $\text{STROKE}_{\text{ALL}}$ includes all strokes, while $\text{STROKE}_{\text{w/o1-3}}$ omits strokes 1-3 of each step. Three rowers completed step 6 (450 W). SF = Stroke frequency.

ROWERS		TARGET	SF [1/MIN]	REF				C2			
Sex and Step Number		W	[1/MIN]	$\text{STROKE}_{\text{All}}$		$\text{STROKE}_{\text{w/o1-3}}$		$\text{STROKE}_{\text{All}}$		$\text{STROKE}_{\text{w/o1-3}}$	
				Mean [W]	SD [W]	Mean [W]	SD [W]	Mean [W]	SD [W]	Mean [W]	SD [W]
male	1	200	18.8 ± 0.7	204.9	17.3	202.6	9.9	188.8	10.6	187.6	6.4
	2	250	20.4 ± 0.7	251.4	11.4	250.3	7.4	234.1	9.7	234.4	4.8
	3	300	22.2 ± 0.7	301.2	11.3	300.3	7.5	282.8	11.4	283.8	4.9
	4	350	24.2 ± 0.6	352.3	11.5	352.1	9.8	333.5	13.4	334.9	6.5
	5	400	26.7 ± 1.0	403.3	12.7	402.6	10.2	383.9	16.7	388.7	6.7
	6	450	29.4 ± 1.4	455.1	11.1	454.4	9.1	436	18.6	438.1	5.6
female	1	140	18.2 ± 0.5	143.4	14	141.1	5.5	130.3	7.9	129.4	4.7
	2	180	19.8 ± 0.3	181.2	13.9	179.1	5.1	167.2	6.6	166.6	3.5
	3	220	22.1 ± 0.4	221.4	11	220.2	6.2	206.8	7.3	207	3.9
	4	260	24.6 ± 0.9	261.3	12.3	260	7.5	246.1	10.3	246.9	4.3
	5	300	27.4 ± 1.0	301.1	11.6	299.9	7	285.2	12.1	286.4	4.6

female rowers, where the CV of step 1 was significantly higher than at each subsequent step (all p-values <0.05).

Differences in mechanical Power Output between REF and C2

Table 3 shows the LS mean of the actual (≠ target) stroke-to-stroke P and the LS mean standard deviation, respectively, of each step in male and female rowers. A representative example of stroke-to-stroke data of a single step in a male elite rower can be found in Figure 2. In addition, Figure 1 shows stroke-to-stroke variability during a complete test in one exemplary rower.

In male and female rowers, the average P of all steps determined with the C2 was significantly lower compared to REF (18.2±1.3 W; p<0.001 and 14.7±1.1 W; p<0.001, respectively).

Regression and Correlation Analysis of Variability vs. Differences in mechanical Power Output

The linear regression of CV and the mean difference between REF and C2 at each step and rower is shown in Figure 3. Correlation analysis indicated a moderate, significant correlation (r=0.64, p<0.001).

Discussion

The main results of this study that aimed to assess stroke-to-stroke variability in elite rowers during stepwise incremental testing on air-braked rowing ergometers are (i) a stroke-to-stroke variability in elite rowers amounting to 4.0-5.5%, (ii) an average variability that was significantly lower in male than in female rowers, if all strokes were considered, and

(iii) differences both in magnitude of the CV as well as in its course during the test between REF and C2.

Stroke-to-Stroke Variability in Elite Rowers

As indicated by the REF data in Table 2, average stroke-to-stroke variability amounted to 4.7% (4.0% in male and 5.5% in female rowers). It is difficult to interpret the magnitude of these values, because to the best of our knowledge there are no data – not to mention clear cut-offs –available to judge biological variability of P within constant load steps, neither for rowing nor for any other type of constant ergometer work. However, the inconsistency in P appeared surprisingly high to us, especially in comparison to torque-controlled, electromagnetically braked cycling ergometers. In those ergometers, torque or resistance is a function of power, thereby allowing to keep P virtually constant. The observed variability appears also high in light of the fact that all athletes within this study were of world elite level (2) including world champions and Olympic medalists, who were very familiar with ergometer training and testing. Furthermore elite rowers are supposed to have the most constant rowing skills (19). Our findings deviate from Shimoda et al. (18), reporting an average CV of 1.7 – 3.0% (P_{C2}) for college rowers, which is considerably lower than the corresponding CVs calculated with P_{C2} of 4.9% and 5.0% in our study. Nevertheless, stroke-to-stroke variability of our elite rowers is much lower than in national level or club rowers, reported to approximately 15.4%, and even slightly lower than 6.7% reported for elite rowers by Smith and Spinks (19). Of note, there was a considerable individual variability, with CVs ranging from 2.9% to 11.8% within one step, indicating that even within such a homogeneous group, the ability to meet the target power output differs markedly.

Sex Differences in Stroke-to-Stroke Variability

Our REF-based data further suggest, at least on the first view, that average variability (%CV) is lower in male than in female rowers (p=0.002). Interestingly, the statistical difference disappears if the first three strokes are excluded from the analysis (Table 2), leading to an identical CV of 3.0%. Hence, the sex differences in variability are solely due to the start strokes. “On the course” stroke-to-stroke precision is virtually the same (please see below for a discussion of the particularities of strokes 1–3). Therefore, the “sex associated difference” in Table 2 should, in our opinion, be interpreted very cautiously. First of all, absolute standard deviations are very similar between male and female rowers with the difference ranging -2.5 to 1.1 W (Table 3), except for step 1 where the difference is 3.3 W. Here, if anywhere, physical reasons for the differences might actually be suspected, as it is the absolutely more powerful men who clearly overpace in the lowest stage, regardless whether starting strokes are considered or not. Apart from this, the higher CVs of female rowers are apparently due to the lower absolute P in the denominator (i.e., $\bar{x}$ in equation 1), due to the lower target value. So, even if male and female rowers tend to surpass the targeted P similarly, relative variability is higher in the female athletes.

Differences between C2 and Reference Measurement

The results discussed so far rely exclusively on REF data. In addition, our study indicates significant differences between REF and C2 regarding variability and absolute P. For a more comprehensive understanding of the results, a detailed discussion of the results seems warranted, even though this might be beyond the initial focus of this paper.

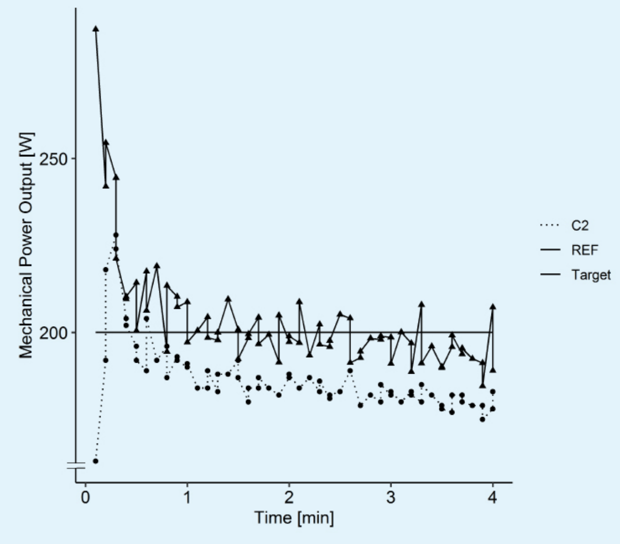

Figure 2

Stroke-to-Stroke data of mechanical power output measured with a reference system (REF) and Concept 2 rowing ergometer (C2) in one male rower during a 200 Watt step of incremental testing.

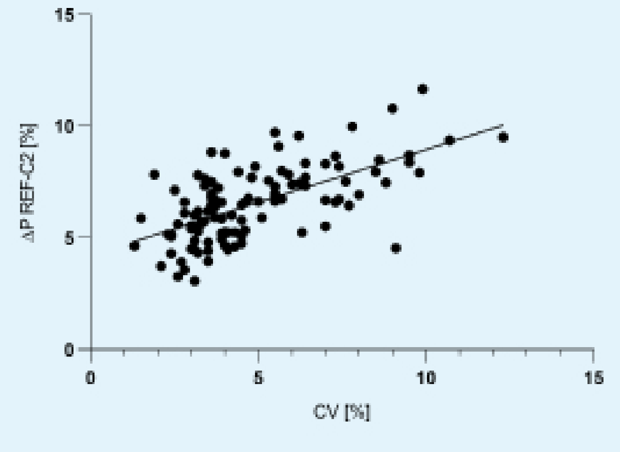

Figure 3

Linear regression between coefficient or variation in stroke-by-stroke mechanical power output versus the difference in power output between the Concept 2 rowing ergometer’s PM5 display (C2) and a reference system (REF) in 13 male and 10 female rowers. Each data point represents one step of an incremental step test ranging 1-5 steps (three rowers completed 6 steps). r2=0.39; y=0.4726x + 4.225.

Difference in Variability between REF and C2 when all Strokes are Considered

Table 2 shows the average CVs of all steps and indicates that variability is measured significantly different by REF and C2 for all rowers, regardless if $\text{STROKE}_{\text{ALL}}$ or $\text{STROKE}_{\text{w/o1 3}}$ are considered. Furthermore, CVs calculated from the C2-dataset are significantly smaller than those obtained by REF, with the analysis of $\text{STROKE}_{\text{ALL}}$ in male rowers being the only exception. Finally, Table 2 shows that variability measured by REF differs significantly between male and female rowers, but is almost equal when C2 data are analyzed. Even though these “over-all” findings already point to the interpretation that C2 power output calculation might mitigate the actual variability, probably attributable to the stable, inertia related nature of an air-braked ergometer once the flywheel has been accelerated (16), a detailed analysis allows to understand the more complex › origination of the all-strokes mean values shown in Table 2: In Figure 1A-B a continuous and marked decrease of the mean CVs for REF is obvious, while the means of the C2 data form a U-shape, both in male and female rowers. Hence, the CVs calculated from REF are not always higher than C2, indeed they are higher only at the lowest two steps, while the relationship is reversed in subsequent steps with higher workloads. Of note, the difference shown is clearly underlined by the significant interaction-effect of *measurement system*step*, indicating that the CVs changed differently between REF and C2 throughout the test (Table 1).

When including absolute variability (Table 3), expressed as standard deviation in absolute power (Watt) into the STROKE$_{ALL}$-analysis, the actual differences between REF and C2 become even clearer: While the standard deviations tend to decrease in REF and range within a small span from the third step at the latest (11.1-12.7 W (min-max) in male and 11.0-12.3 W in female), the C2-standard deviations increase continuously following the nadir in step 2. Hence, the decreasing CVs of REF in Figure 1 are mainly due to increasing average P at each step (which is the denominator in equation 1), while absolute variability remains relatively stable after an initial decrease. The situation is different with the C2-data, where CVs increase due to continuously increasing absolute variability (i.e. standard deviations) from step 2 onwards, being relatively high compared to the increase in average power.

Differences in Variability between REF and C2 without Strokes 1-3

It has already been reported here and elsewhere (7), that the first three to five strokes play a "special role" in the C2-calculation of P. They almost inevitably affect variability. Unsurprisingly Table 2 shows that all average CVs are significantly lower when the first three strokes are excluded from the analysis – irrespective of measurement system and sex. Also, average variability for a given measurement system (REF or C2) is no longer sex related. Figure 1 indicates, that the CVs based on REF-data are always and mostly significantly higher than for C2. This is in line with current research, also reporting that stroke-to-stroke variability is underestimated in rowers of lower performance level by the C2 (10). In addition, the course of the average CVs is similar for REF and C2 in male and female, i.e., after an initial decrease in steps 1 and 2, the CVs form a plateau, remaining roughly within a 1%-span. Concomitantly, the mixed model analysis indicates that CVs changed in a similar way between systems, due to an insignificant *measurement system*step*-interaction effect (p= 0.32) for STROKE$_{w/o1\text{-}3}$ in male and female rowers (Table 1).

The averaged standard deviations in Table 3 clearly show that the CVs for STROKE$_{w/o1–3}$ (Figure 1) are lower than those for STROKE$_{ALL}$ (Figure) due to a considerably lower absolute variation (i.e., standard deviation) in P, while the mean values are only mildly influenced by the exclusion of the start strokes. Up to this point it remains to be summarized, that the C2 calculation underestimates stroke-to-stroke variability in elite rowers during 4-min step tests if the first strokes are excluded from the analysis and that those first three strokes have a considerable impact on the variability – irrespective if measured via REF or C2. This underestimation might partly explain the high reliability of C2 tests (2000 m time trial) reported in other studies (16), because due to the longer test duration and number of strokes, the impact of the starting strokes on total reliability is reduced. This raises the question for the reason of the impact of strokes 1-3 on variability.

First, the REF data in Figure 1 suggest that at each step (always preceded by a 30-s break for blood sampling), especially the first ~3 strokes are characterized by a strong difference relative to the target value. This difference is not due to the measurement system, but due to the rowers, who obviously need some strokes to adjust force and stroke frequency to the targeted power output, a behavior that has been reported before (23). Figure 1 illustrates, that in STROKE$_{ALL}$ analysis, the CV is increased due to a factual higher variation in P and, of note, this leads to a high difference in CV between STROKE$_{ALL}$ vs. STROKE$_{w/o1–3}$. Consequently, rowers overpace during the first strokes and this is measured correctly by REF.

Second, Figure 2 shows a huge underestimation of the first three strokes by the C2-system, which is followed by a relatively small difference between both measurement systems subsequent to stroke #5. This underestimation has already been reported before (7), but we add the information, that the underestimation is not systematic. This notion is also supported by Figure 2 (where the first strokes of REF and C2 do neither proportionally nor inversely mirror each other) and further by the different course of the CVs in Figure 1 – where the behavior of the C2 in strokes 1–3 leads to a U-shape of the average CVs of all steps, while a flattening of the average CVs is obvious when REF is considered or strokes 1-3 are excluded. This means that stroke power of the first strokes is underestimated and not measured correctly by C2.

The main reason for the contrasting behavior of the C2 during the first strokes is directly related to the C2 measurement and the PM5-power calculation. Briefly, the C2 power calculation is based on the measurement of angular velocity of the flywheel (which is the only measured variable) and its known mass (25). The REF-system, in contrast, measures stroke work (i.e., integration of force-displacement curve) per duration of the whole rowing cycle. This leads to different results, especially if stroke-by-stroke variability is high, e.g., during starting strokes. In other words, the REF system measures the mechanical power applied by the rower, whereas the C2 measures the impact of the rower's mechanical power on the behavior of the flywheel, but it does not account for the energy that is stored in the flywheel during the first strokes. These details explain the unsystematic differences in variability of the first three strokes between REF and C2 observed in this study.

Potential Impact of our Findings

The practical implications of our results remain speculative at this moment and need to be evaluated in future studies. First, due the impact of the first strokes on the overall variability of each stage, long stage durations of at least 4 min seem to be advantageous especially at low intensities, because the more strokes, the less the influence of the first strokes. This is especially true when P is measured with the C2 only, where the first strokes are almost contrasting actual P. Moreover, a very high individual variability within steps of up to 19 W or 13.0% (highest individual within-step variation found in this study) probably influences the outcome of physiological measurements, like e.g., blood lactate concentrations, especially at higher workloads due to higher muscular energy turnover (3). Hence, considering the small within-subject variability of lactate-based fitness markers in elite rowers (24), small changes between a particular rower's tests might possibly be related to differences in stroke consistency instead of aerobic performance. Finally, it has been suggested that a consistent P facilitates higher efficiency for a given distance of rowing. Hence, rowers should focus on constant pace and avoid unnecessary fluctuations in P including the start strokes (18). This notion is in line with recent research (22) Finally, we would like to highlight that stroke-to-stroke variability is a potentially relevant variable for rowing ergometer testing, but not for e.g., cycle ergometer testing, where mechanical power demand is externally controlled as a function of cadence and torque, leading to a variability approximating zero.

Limitations

It is worth mentioning that the sample size of 13 male and 10 female rowers might be too small to draw generalizable conclusions. However, we aimed to describe the variability in elite rowers and finally acquired two complete National Teams. An inclusion of e.g. U23 rowers would have meant a reduction of the sample's quality. Another limitation is that the metabolic implications of our findings remain unclear. We would like to emphasize in this context that we aimed to investigate the variability of P to determine whether there is a potentially relevant problem. The evaluation of the physiological implications requires an experimental design that was not implementable in this study.

Conclusion

We report a non-negligible variability of stroke-to-stroke P in elite rowers during incremental step testing. A relatively high percentage of the variability is attributable to the first three strokes that influence the coefficient of variation within each 4-min step in male and female rowers due to an actual overpacing and due to a considerable underestimation of mechanical power output by the C2. If strokes 1-3 are excluded from the analysis, the underestimation of stroke-to-stroke variability by the C2 becomes obvious, but there are no differences in accuracy between male and female rowers. Our results point to the notion that rowers should aim to minimize fluctuations in mechanical power output during step testing, especially when metabolic variables are evaluated. Aside from that it is worth mentioning that probably every certified cycle ergometer allows for much lower variability during constant load ergometer testing than an air-braked rowing ergometer like the Concept 2. However, further investigations are warranted to elucidate why variability of the C2 power calculation is slightly different from a reference system even after the first start strokes. Studies employing test rigs for rowing ergometers are necessary to evaluate these questions.

Acknowledgements

The support of the German Rowing Federation's coaches Marcus Schwarzrock and Marcin Witkowski is gratefully acknowledged. We would like to thank Mr. Mark Amort for his technical support during data acquisition and all athletes for their enthusiasm and the time they dedicated to this study. We acknowledge the clarification of technical details by Peter Dreissigacker (Concept 2). ■

Conflict of Interest

The authors have no conflict of interest.

References

(1) **ALTENBURG D, MATTES K, STEINACKER JM.** Manual of rowing training: Technique, high performance and planning. 2. rev. ed. Limpert: Wiebelsheim; 2012.

(2) **ARAÚJO CGS, SCHARHAG J.** Athlete: a working definition for medical and health sciences research. Scand J Med Sci Sports. 2016; 26: 4-7. doi:10.1111/sms.12632

(3) **BENEKE R, VON DUVILLARD SP.** Determination of maximal lactate steady state response in selected sports events. Med Sci Sports Exerc. 1996; 28: 241-246. doi:10.1097/00005768-199602000-00013

(4) **BOURDIN M, MESSONNIER L, LACOUR JR.** Laboratory blood lactate profile is suited to on water training monitoring in highly trained rowers. J Sports Med Phys Fitness. 2004; 44: 337-341.

(5) **BOURGOIS J; CLAESSENS AL.**; Janssens M; Van Renterghem B; Loos R; Thomis M; Philippaerts R; Lefevre J; Vrijens J. Anthropometric characteristics of elite female junior rowers. J Sports Sci. 2001; 19: 195-202. doi:10.1080/026404101750095358

(6) **BOURGOIS J, CLAESSENS AL, VRIJENS J, PHILIPPAERTS R, VAN RENTERGHEM B, THOMIS M, JANSSENS M, LOOS R, LEFEVRE J.** Anthropometric characteristics of elite male junior rowers. Br J Sports Med. 2000; 34: 213-216, discussion 216-217. doi:10.1136/bjsm.34.3.213

(7) **BOYAS S, NORDEZ A, CORNU C, GUÉVEL A.** Power responses of a rowing ergometer: mechanical sensors vs. Concept2 measurement system. Int J Sports Med. 2006; 27: 830-833. doi:10.1055/s-2006-923774

(8) **COHEN J.** Statistical power analysis for the behavioral sciences. 2nd ed. L. Erlbaum Associates: Hillsdale, N.J; 1988.

(9) **HEGGE AM, MYHRE K, WELDE B, HOLMBERG H-C, SANDBAKK Ø.** Are gender differences in upper-body power generated by elite cross-country skiers augmented by increasing the intensity of exercise? PLoS One. 2015; 10: e0127509. doi:10.1371/journal.pone.0127509

(10) **HOLT AC, HOPKINS WG, AUGHEY RJ, SIEGEL R, ROUILLARD V, BALL K.** Concurrent Validity of Power From Three On-Water Rowing Instrumentation Systems and a Concept2 Ergometer. Front Physiol. 2021; 12. doi:10.3389/fphys.2021.758015

(11) **HOPKINS WG, SCHABORT EJ, HAWLEY JA.** Reliability of power in physical performance tests. Sports Med. 2001; 31: 211-234. doi:10.2165/00007256-200131030-00005

(12) **INGHAM SA, WHYTE GP, JONES K, NEVILL AM.** Determinants of 2,000 m rowing ergometer performance in elite rowers. Eur J Appl Physiol. 2002; 88: 243-246. doi:10.1007/s00421-002-0699-9

(13) **KLESHNEV V.** Power in Rowing. In: Proceedings of XVIII Symposium of ISBS, Hong-Kong, 2000; 96-99.

(14) **LORMES W, BUCKWITZ R, REHBEIN H, STEINACKER JM.** Performance and blood lactate on Gjessing and Concept II rowing ergometers. Int J Sports Med. 1993; 14: S29-S31. doi:10.1055/s-2007-1021220

(15) **MAYHEW JL, SALM PC.** Gender differences in anaerobic power tests. Eur J Appl Physiol Occup Physiol. 1990; 60: 133-138. doi:10.1007/BF00846033

(16) **SCHABORT EJ, HAWLEY JA, HOPKINS WG, BLUM H.** High reliability of performance of well-trained rowers on a rowing ergometer. J Sports Sci. 1999; 17: 627-632. doi:10.1080/026404199365650

(17) **SCHWARZROCK M, TREFF G, VIEDT C AND WORKING GROUP.** Trainingsmethodische Grundkonzeption 2017-2020: Hannover, 2017.

(18) **SHIMODA M, FUKUNAGA T, HIGUCHI M, KAWAKAMI Y.** Stroke power consistency and 2000 m rowing performance in varsity rowers. Scand J Med Sci Sports. 2009; 19: 83-86. doi:10.1111/j.1600-0838.2007.00754.x

(19) **SMITH RM, SPINKS WL.** Discriminant analysis of biomechanical differences between novice, good and elite rowers. J Sports Sci. 1995;13:377-385. doi:10.1080/02640419508732253.

(20) **SMITH TB, HOPKINS WG.** Variability and predictability of finals times of elite rowers. Med Sci Sports Exerc. 2011; 43: 2155-2160. doi:10.1249/MSS.0b013e31821d3f8e

(21) **SOPER C, HUME PA.** Reliability of power output during rowing changes with ergometer type and race distance. Sports Biomech. 2004; 3. doi:10.1080/14763140408522843

(22) **TREFF G, MENTZ L, MAYER B, WINKERT K, ENGLEDER T, STEINACKER JM.** Initial Evaluation of the Concept-2 Rowing Ergometer's Accuracy Using a Motorized Test Rig. Frontiers in Sports and Active Living. 2021.

(23) **TREFF G, WINKERT K, MACHUS K, STEINACKER JM.** Computer-Aided Stroke-by-Stroke Visualization of Actual and Target Power Allows for Continuously Increasing Ramp Tests on Wind-Braked Rowing Ergometers. Int J Sports Physiol Perform. 2018; 13: 729-734. doi:10.1123/ijspp.2016-0716

(24) **TREFF G, WINKERT K, SAREBAN M, STEINACKER JM, BECKER M, SPERLICH B.** Eleven-Week Preparation Involving Polarized Intensity Distribution Is Not Superior to Pyramidal Distribution in National Elite Rowers. Front Physiol. 2017; 8: 515. doi:10.3389/fphys.2017.00515

(25) **VAN HOLST.** Behind the ergometer display. 2014. http://home.hccnet.nl/m.holst/ErgoDisp.pdf [15th January 2014].

(26) **VOGLER AJ, RICE AJ, GORE CJ.** Physiological responses to ergometer and on-water incremental rowing tests. Int J Sports Physiol Perform. 2010; 5: 342-358. doi:10.1123/ijspp.5.3.342

(27) **WASSERMAN K, BEAVER WL, DAVIS JA, PU JZ, HEBER D, WHIPP BJ.** Lactate, pyruvate, and lactate-to-pyruvate ratio during exercise and recovery. J Appl Physiol. 1985; 59: 935-940. doi:10.1152/jappl.1985.59.3.935