



ARTICLE

SCI & Exercise

# The test-retest reliability of individualized $\text{VO}_2\text{peak}$ test modalities in people with spinal cord injury undergoing rehabilitation

Nicolaj J. Holm<sup>1</sup> · Fin Biering-Sørensen<sup>1</sup> · Lone H. Schou<sup>2</sup> · Tom Møller<sup>3</sup>

Received: 19 March 2020 / Revised: 19 August 2020 / Accepted: 20 August 2020 / Published online: 4 September 2020  
© The Author(s), under exclusive licence to International Spinal Cord Society 2020

## Abstract

**Study design** Test-retest study.

**Objective** To describe the test-retest reliability of an individualized approach for peak oxygen uptake ( $\text{VO}_2\text{peak}$ ) testing in a clinical context, accommodating inpatients with heterogeneous levels of physical functioning.

**Setting** Inpatient rehabilitation.

**Method** Test-retest on two types of devices (recumbent stepper and arm-crank ergometer) with two predefined protocols of varying load increments on each device. No participants performed all four test modalities.  $\text{VO}_2\text{peak}$ , respiratory exchange ratio (RER), peak workload, and the reason for termination were reported. Of 23 participants included, 21 completed the test-retest procedure. Participants were inpatients,  $\geq 18$  years, who sustained a spinal cord injury (SCI) within the last 12 months at the time of admission. The median age was 52 years and 74% were male. Median time since injury at the time of the first test (T1) was 5.5 months. Seventeen were classified as American Spinal Injury Association Impairment Scale D. Intraclass correlation coefficients (ICC) and Bland–Altman plots for  $\text{VO}_2\text{peak}$  was calculated to determine intra- and intertester reliability of  $\text{VO}_2\text{peak}$ .

**Results** Mean  $\text{VO}_2\text{peak}$  obtained at T1 was 1.91 L/min (95%CI: 1.31–2.51); 1.26 L/min and 2.69 L/min for the recumbent stepper protocols; 0.61 L/min and 1.39 L/min for the arm-crank ergometer protocols. ICC for intra- and intertester reliability was 0.997 (95%CI: 0.986–0.999) and 0.994 (95%CI: 0.978–0.998) respectively. Bland–Altman plots showed a mean difference of  $-0.005$  (SD 0,12) for overall test-retest.

**Conclusion** Test-retest reliability was high. Mean  $\text{VO}_2\text{peak}$  at discharge from rehabilitation was higher than previously reported and individualized test modalities seemed to yield more optimal results in a heterogeneous and representative sample during rehabilitation.

## Introduction

Physical deconditioning of people with spinal cord injury (SCI) is a contributing and major risk factor of cardiovascular disease, and the clinical practice guideline for identification and management of cardiometabolic risk after SCI

recommends exercise testing of all people with SCI when possible [1]. At the Department for Spinal Cord Injuries, Rigshospitalet, Denmark  $\text{VO}_2\text{peak}$  testing is performed during rehabilitation as part of standard care, by different testers in patients with heterogeneous levels of physical functioning.

The most common modalities used for  $\text{VO}_2\text{peak}$  testing in people with SCI have been arm-crank ergometers (ACE) or hand-bikes and manual wheelchair protocols on rolls or treadmill [2]. Arm ergometry protocols were suitable for people without motor function in the lower extremities, but underestimated aerobic capacity in people with an incomplete SCI and preserved motor function in the lower extremities [3]. This may be a problem when comparing individual test results to reference fitness values. A seated recumbent stepper with simultaneous involvement of the upper and lower extremities was found feasible in people

✉ Nicolaj J. Holm  
nicolaj.jersild.holm@regionh.dk

<sup>1</sup> Department for Spinal Cord Injuries, Neuroscience Center, Hornbæk, Denmark

<sup>2</sup> Department of Nursing and Nutrition, University College Copenhagen, Copenhagen, Denmark

<sup>3</sup> University Hospitals Centre for Health Care Research, University hospital Copenhagen, Rigshospitalet, Denmark

with an incomplete SCI, and may be an alternative testing modality although reliability was not investigated [4]. High test-retest reliability of VO<sub>2</sub>peak was found in previous studies using ACE and robot-assisted treadmill protocols in people with SCI [3, 5]. However, these studies were not conducted in an inpatient rehabilitation setting with different testers involved.

Studies assessing VO<sub>2</sub>peak in a clinical inpatient setting during rehabilitation are few, and included arm ergometry protocols in wheelchair-dependent people only [2, 6–9]. However, inpatients during rehabilitation are heterogeneous due to differences in the level and severity of SCI, and a high number of people have a motor incomplete SCI at admission with a potential for improving motor functioning during rehabilitation [10, 11]. Therefore, a set-up with individualized test modalities for testing cardiovascular fitness may be more appropriate for achieving VO<sub>2</sub> peak results presumed to be the most optimal in accordance with pre-defined criteria for VO<sub>2</sub> peak.

The aim of this study was to investigate the test-retest reliability of individualized testing of VO<sub>2</sub>peak in a rehabilitation setting involving several testers and a heterogeneous sample of participants reflecting clinical practice. We hypothesized that the predefined test modalities used in this context would be reliable for use in standard care based on measures of relative and absolute reliability. We also hypothesized that selecting individualized test modalities, based on the participants' motor function and the testers clinical judgement, would result in more optimal VO<sub>2</sub>peak values by achieving better thresholds for standard VO<sub>2</sub>peak criteria.

## Methods

### Participants and eligibility criteria

This study was part of a prospective cohort study with a consecutive enrollment of all patients with a new SCI during a 10 months period [12] at the Department for Spinal Cord Injuries, Rigshospitalet, Denmark. Participants were > 18 years, having sustained a SCI within the last 12 months before admission to rehabilitation (Table 1). Due to concerns with completing a valid exercise test, individuals with motor-complete injuries (American Spinal Injury Association (ASIA) Impairment Scale (AIS) A and B)) [13], at or above C4 requiring artificial ventilation, were excluded from the study. Other and previously used exclusion criteria were the presence of decubiti, severe spasticity or musculoskeletal problems at risk of exacerbation or aggravation during testing or preventing completion of the test [6]. If diagnosed cardiovascular disease was present a medical doctor decided whether the participant could engage in the exercise test.

### Protocols and equipment

Two different test devices and two predefined protocols for each device were used for VO<sub>2</sub>peak testing. Primarily, one of the test modalities was used for each participant depending on functional level and overall condition. However, in some instances it was decided to switch to one of the other test modalities to reach a VO<sub>2</sub>peak presumed to be more optimal (Fig. 1).

A recumbent stepper (NuStep T5XR®, Ann Arbor, MI, USA) was primarily used for participants with a motor incomplete SCI and higher lower extremity motor score (LEMS) (Fig. 1). The two stage protocols incorporated were a Total Body Recumbent Stepper Exercise Test protocol (TBRS-XT) starting at 50 watts (W) with 115 steps per minute (SPM) and a 25 W incremental increase every 2 min during the first three stages and 30 W increments thereafter. The other protocol was a modified TBRS-XT (mTBRS-XT) starting at 25 W with 80 SPM and with 15 W increments every 2 min. Both protocols were described by Billinger et al. [14, 15].

Two different ACE protocols (SCI FIT Pro1®, Tulsa, OK, USA) were used for participants with a motor complete SCI, very de-conditioned participants or participants with an incomplete SCI but a low LEMS [2] (Fig. 1). The protocols were designed as incremental stage protocols starting at 5 W and increased 5 W each minute in people with tetraplegia (Tetra-ACE) and 10 W in people with paraplegia (Para-ACE) with 60 revolutions per minute. The protocols were entered into the stress test program of the SCI FIT Pro1® software before testing. The ACE was easy adjustable for optimal positioning of the upper extremities and was set up for synchronous arm-crank movements.

### Testing procedure

Three physiotherapists performed the tests reflecting clinical practice. Participants were allocated by randomization to two VO<sub>2</sub>peak test sessions of either intra- or interrater reliability within the last two weeks before discharge from rehabilitation. However, independent of randomization, the physiotherapists who were skilled in VO<sub>2</sub>peak testing choose one of the four test modalities depending on the functional level of the participant. The choice of test modality was based on information provided by the participants' primary physiotherapist and the testers clinical judgement. The flowchart in Fig. 1 guided the process of choosing between the recumbent stepper protocols or arm crank protocols. However, the protocol choice between the two recumbent stepper protocols was made from clinical judgement and feedback from the participant after trying the initial resistance and step cadence (Fig. 1). Participants were instructed to continue the test for as long as possible, and

**Table 1** Individual participant characteristics and results from the VO<sub>2</sub>peak test for all 23 participants.

Severity group	Participant number	Gender	Age	Cause	TSI	NLI	AIS	UEMS/ LEMS	Test modality	VO <sub>2</sub> peak (L/min)	RER	W	Time
<i>CI-C8 A,B,C</i>	1	M	52	T	12	C3	A	11/0	Tetra ACE	0.61	1.0	15	02.30
<i>T1-S5 A,B,C</i>	2	F	83	T	5	Para	A	–	Para ACE	0.69**	1.02	55	04.30
	3	M	26	NT	6	T1	C	50/ 24	Para ACE	1.21	1.13	85	09.00
	4	M	21	T	7	T5	C	50/0	Para ACE	1.46	1.03	75	08.00
	5	M	70	NT	7.5	T11	C	50/23	mTBRSt- XT	1.60	1.18	85	09.15
	6	M	44	T	4	T10	A	50/2	Para ACE	1.87	1.26	85	11.15
<i>All AIS D</i>	7	F	76	T	5	C5	D	37/45	mTBRSt- XT	0.70	0.86	40	02.45
	8	M	65	T	5.5	C1	D	50/49	mTBRSt- XT	0.87	0.88	40	03.00
	9	M	71	NT		C1	D	37/40	mTBRSt-XT	0.88**!	0.90	40	02.30
	10	F	69	NT	5	T7	D	50/49	mTBRSt- XT	1.03	1.03	70	06.45
	11	M	51	NT	4	C5	D	50/40	mTBRSt- XT	1.14	0.98	65	05.15
	12	M	70	NT	6	T2	D	50/26	mTBRSt- XT	1.24	1.17	70	06.30
	13	F	41	NT	11	C8	D	50/46	mTBRSt- XT	1.68	1.10	100	10.15
	14	M	61	NT	9	T9	D	50/5	Para ACE	1.75	1.19	115	11.15
	15	M	42	T	4	C4	D	50/44	mTBRSt- XT	1.89	1.12	115	12.20
	16	M	75	NT	3	T12	D	50/40	mTBRSt-XT (T1a)/ TBRSt-XT (T1b)	1.99	1.08	200	16.00
	17	F	27	NT	6.5	L4	D	50/50	mTBRSt-XT (T1a)/ TBRSt-XT (T1b)	2.05	1.06	130	16.00
	18	M	40	NT	4	C1	D	50/50	TBRSt- XT	2.51	1.05	200	12.00
	19	M	56	NT	3	C6	D	44/45	mTBRSt-XT (T1a)/ TBRSt-XT (T1b)	2.56*	1.03	115	10.00
	20	M	45	NT	2	L4	D	50/49	TBRSt- XT	2.91	1.09	200	11.00
	21	M	43	T	7	C3	D	45/50	TBRSt-XT	3.07	1.11	200	10.45
	22	M	41	NT	3.5	C2	D	50/49	TBRSt-XT	3.18	1.04	200	17.00
	23	M	27	NT	3	Para	D	–	mTBRSt-XT (T1a)/ TBRSt-XT (T1b)	3.38*	1.17	200	14.15

Severity group based on neurological level and severity of spinal cord injury (American Spinal Injury Association (ASIA) Impairment Scale (AIS) classification – A, B, C, and D).

Gender: *M* male or *F* female, Cause: *T* traumatic or *NT* non-traumatic, *TSI* Time since injury (months) at the time of VO<sub>2</sub>peak testing, *UEMS* Upper extremity motor score, and *LEMS* lower extremity motor score, *Tetra ACE* Test modality: tetraplegia arm crank ergometer modality, *Para ACE* paraplegia arm crank ergometer modality, *mTBRSt- XT* modified total body recumbent stepper exercise test, *TBRSt- XT* total body recumbent stepper exercise test, *VO<sub>2</sub>peak* (L/min) peak oxygen uptake, respiratory exchange ratio at VO<sub>2</sub>peak (RER), *W* Watts resistance at test termination, *test 1a (T1a)* T1a test duration (minutes) (Time) for participants at test 1 test 1a, *T1b* test duration (minutes) (Time) for participants at test 1 test 1b.

All values obtained at T1b(\*).

Two participants did not perform a retest (\*\*).

One of them terminated testing due to knee pain (!).

the test was terminated due to volitional exhaustion, when it was not possible to maintain cadence despite verbal encouragement or if an adverse event occurred.

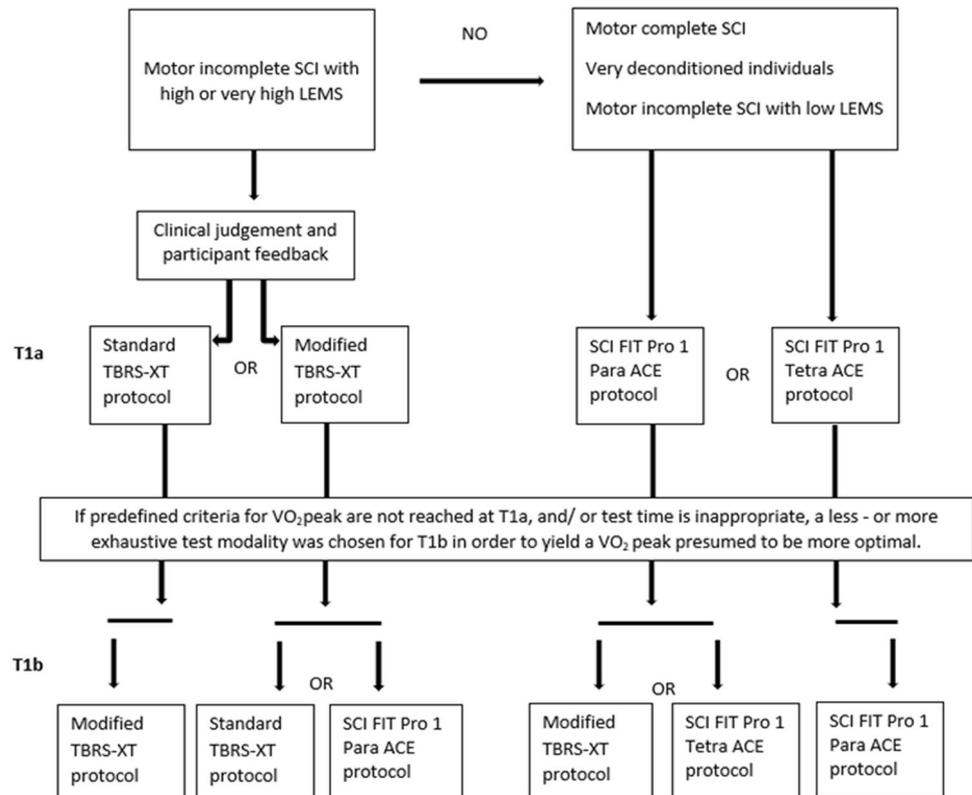
Criteria for VO<sub>2</sub>peak were defined as the highest mean VO<sub>2</sub> (L/min) recorded over a 30 s sample, and a corresponding respiratory exchange ratio (RER) > 1.0 [2, 6, 16, 17]. If predefined criteria for VO<sub>2</sub>peak were not reached during the first test (T1a), another protocol was chosen in order to yield a VO<sub>2</sub> peak presumed to be more optimal. The new protocol was tested (T1b) and retested (T2b), resulting in three tests performed in total. VO<sub>2</sub>peak

and RER from T1a and T1b were compared to determine if any improvement occurred. However, this approach was not possible if the Tetra-ACE protocol was too hard and the criteria for VO<sub>2</sub>peak were not reached. A protocol switch to a more demanding protocol was decided if the test duration exceeded 15 min in order to reach a VO<sub>2</sub>peak presumed to be more optimal. Likewise, if the test duration was shorter than 5 min a less demanding protocol was chosen if possible (Fig. 1).

Test-retest was performed at the same time of the day with 48 h to 5 days in between. Participants were instructed

**Fig. 1** Flow chart showing the process of identifying an appropriate test protocol by clinical judgement based on functional level and overall condition of the participant.

Abbreviations: Spinal cord injury SCI; lower motor extremity motor score LEMS; total body recumbent stepper exercise Test protocol TBRS-XT; arm crank ergometer ACE (adjusted for para (Para)- and tetraplegic (Tetra) individuals respectively); test 1a T1a; test 1b T1b.



to refrain from caffeine, alcohol and intensive physical exercise prior to the test on the day of testing, as well as tobacco smoking two hours before testing. Bladder emptying was carried out immediately before testing. A warmup for maximum 5 min was performed if desired by the participant.

VO<sub>2</sub>peak was measured as absolute VO<sub>2</sub> L/min by a portable metabolic system for cardiopulmonary exercise testing (Metamax 3B from Cortex Biophysik GmbH, Leipzig, Germany) and a Hans Rudolph mask with breath-by-breath measurements averaged every 15 s. The equipment was calibrated prior to every test according to the manufacturer recommendations as described previously [17].

## Statistical analysis

Descriptive statistics were calculated for demographics and expressed as median and interquartile range (IQR). Analysis of VO<sub>2</sub>peak at discharge and VO<sub>2</sub>peak in relation to protocol and severity of SCI was based on data from T1 (T1a or T1b) and expressed as mean and 95% confidence interval (95%CI). Participants were divided into three subgroups based on SCI severity and described as C1-C8 AIS A,B,C; T1-S5 AIS A,B, C and ALL AIS D regardless of injury level [18].

Based on VO<sub>2</sub>peak relative reliability was analyzed for test-retest and intra-rater reliability by intra class correlation coefficient (ICC) with 95% CI, using a two-way mixed

effects model, absolute agreement, single measurement. ICC for inter-rater reliability was analyzed using a two-way random effects model, consistency, single measurement [19]. Absolute reliability was analyzed by standard error of measurement (SEM) calculated as  $SD \cdot \sqrt{1 - ICC}$ . Also, limits of agreement was determined by Bland-Altman plots and linear regression was performed. An alpha level of 0.05 was used to indicate statistical significance. All statistical analyses were performed using IBM SPSS statistics version 22.

## Results

### Descriptives

Of 46 who consented, 23 were able to participate in the VO<sub>2</sub>peak test with 21 participants completing the test-retest procedures (Table 1). One participant did not complete T1a due to worsening of knee pain and one was discharged before T2a. T1a was performed within the last two weeks before discharge (median 5.5 days, IQR 4.0).

Individuals excluded from participation were classified as AIS A-D (52% AIS D) and characterized by a median age of 63 years (IQR 23.7), 12 men (52%) and 11 women (48%) with 16 (70%) having a non-traumatic and 7 (30%) a traumatic SCI. Limited motor function ( $N = 8$ ), pain or

**Table 2** Reasons for exclusion from peak oxygen uptake test ( $\text{VO}_2\text{peak}$ ) at discharge from primary rehabilitation for 23 participants.

Reasons for exclusion from $\text{VO}_2\text{peak}$ test	Number
Not enough voluntary motor function	8
Pain/ aggravation of symptoms	3
Refused test participation specifically	2
Cardiovascular restrictions	2
Pulmonary restrictions	2
Stress due to discharge situation	2
Discharge before scheduled	2
Pressure ulcer	1
Mask discomfort	1

aggravation of symptoms were the primary reasons for exclusion ( $N = 3$ ) (Table 2).

### $\text{VO}_2\text{peak}$

All participants exercised until volitional exhaustion except one who terminated testing due to knee pain, and thus  $\text{VO}_2\text{peak}$  from this participant was excluded for analysis. Mean  $\text{VO}_2\text{peak}$  obtained at T1 (T1a or T1b) was 1.91 L/min (95%CI: 1.31–2.51) with a minimum of 0.61 L/min achieved on the Tetra-ACE protocol and a maximum of 3.38 L/min on the TBRS-XT protocol (Fig. 2c). The highest  $\text{VO}_2\text{peak}$  values were obtained among participants classified as AIS D (1.98 L/min, 95%CI: 1.55–2.41) and were significantly higher than  $\text{VO}_2\text{peak}$  in participants with a T1-S5 level of injury classified as T1-S5 AIS A, B, or C (1.36 L/min, 95%CI: 0.81–1.92),  $p < 0.001$  (Fig. 2a). The AIS D group also reached the highest workloads at  $\text{VO}_2\text{peak}$  (Fig. 2b).  $\text{VO}_2\text{peak}$  obtained on the TBRS-XT protocol (2.69 L/min, 95%CI: 2.24–3.14) was significantly higher than  $\text{VO}_2\text{peak}$  obtained on the mTBRS-XT protocol (1.26 L/min, 95%CI: 0.91–1.61);  $p < 0.000$  (Fig. 2c). Overall, participants with a low LEMS were tested on one of the ACE protocols and participants with higher LEMS scores, ranging from 40 to 50 points, were tested on one of the protocols in the recumbent stepper as expected (Table 1).

A protocol shift from the mTBRS-XT protocol at T1a occurred in seven participants classified as AIS D. Of those, five participants did not reach predefined criteria for  $\text{VO}_2\text{peak}$ . Test duration was short in three participants, but a switch to a less demanding ACE protocol was not possible due to upper extremity or neck problems. Two participants reached a higher  $\text{VO}_2\text{peak}$  after switching to the TBRS-XT protocol on T1b, and improved  $\text{VO}_2\text{peak}$  from 2.24 L/min (RER 0.92) at T1a to 3.38 L/min (RER 1.17) and from 1.94 L/min (RER 0.96) at T1a to 2.56 with L/min (RER 1.03), respectively (Table 1).

Two participants reached  $\text{VO}_2\text{peak}$  at T1a after a prolonged test session exceeding 15 minutes. Switching to the TBRS-XT protocol on T1b resulted in faster volitional exhaustion after 10.00 and 10.30 min, respectively, but the more demanding protocol did not improve  $\text{VO}_2\text{peak}$ . Values obtained at T1a and T1b were 2.05 L/min vs. 2.00 L/min, and 1.99 L/min vs. 1.93 L/min for these two participants.

Compared to reference values for able-bodied men and women  $\text{VO}_2\text{peak}$  was lower in the study participants. In young men (20–29 years) mean  $\text{VO}_2\text{peak}$  was 1.63 L/min (SD 0.53) compared to 4.30 L/min (SD 0.73), and 0.72 L/min (SD 0.18) vs. 1.79 L/min (SD 0.07) in women  $\geq 70$  years [20].  $\text{VO}_2\text{peak}$  obtained by 40–49-year-old men and SCI severity AIS D ( $n = 5$ ) was high compared to the other participants (2.71 L/min, SD 0.52), but still much lower than the reference value (4.01 L/min, SD 0.62) [20]. Relative  $\text{VO}_2\text{peak}$  values ( $\text{mL kg}^{-1} \text{min}^{-1}$ ) were all in the lowest category regardless of gender and age compared to reference values for able bodied. Compared to reference fitness values for people with SCI [21], mean absolute  $\text{VO}_2\text{peak}$  for our male participants with paraplegia (AIS A-D) was 1.85 L/min (SD 0.21) and for tetraplegia (AIS A-D) 1.73 L/min (SD 0.29). These are categorized as “excellent”. Our male participants with paraplegia (AIS A, B, C) had a mean  $\text{VO}_2\text{peak}$  of 1.51 L/min (SD 0.19) and were in the category “good” [21].

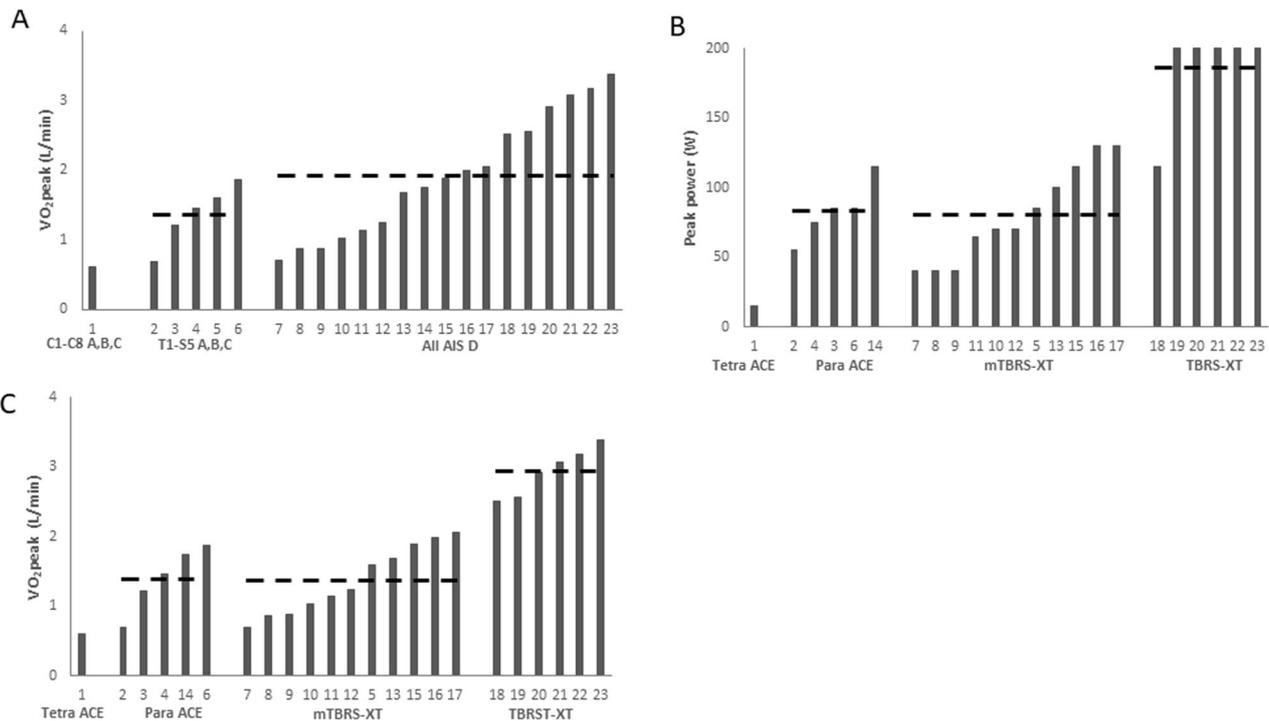
### Test-retest reliability

The test-retest sessions T1a and T2a were separated by a median of 2 days (IQR 1.5). When a protocol shift was made the test-retest sessions T1b and T2b were separated by a median of 2.5 days (IQR 3.25).

The overall ICC for the test-retest sessions for all four protocols together was 0.994 (95%CI: 0.986–0.998), SEM 0.041 L/min. Twelve participants (all AIS D) were randomized to assessment of the inter-tester reliability using the mTBRS-XT or TBRS-XT protocols, and nine participants were randomized to intra-tester reliability. Overall ICC for intra- and inter-tester reliability was 0.997 (95% CI: 0.986–0.999), SEM 0.02 L/min and 0.994 (95%CI: 0.978–0.998), SEM 0.04 L/min, respectively.

The Bland–Altman plots showed a mean difference of  $-0.005$  (SD 0.12) for overall test-retest (Fig. 3a), 0.02 L/min (SD 0.08) for intra-tester reliability (Fig. 3b) and  $-0.02$  L/min (SD 0.15) for inter-tester reliability, respectively (Fig. 3c). Linear regression analysis showed non-significant  $t$ -scores ranging from 0.441 for inter-tester test-retest results to 0.581 for intra-tester test-retest results.

For the respective protocols, ICC for test-retest reliability was 0.992 for the Para-ACE, 0.961 for the mTBRS-XT and 0.990 for the TBRS-XT protocol. One participant only was



**Fig. 2 Peak oxygen uptake and peak power.** **a** Peak oxygen uptake (VO<sub>2</sub>peak) (L/min) according to neurological level of injury and severity of spinal cord injury (SCI) according to American Spinal Injury Association (ASIA) Impairment Scale (AIS) A, B, C, and D. Participant numbers are specified for each column in the same order as shown in Table 1. Dotted line indicates mean VO<sub>2</sub>peak in each severity group; **b** Peak power (W) achieved at VO<sub>2</sub>peak for each of the test modalities. Participant numbers are specified for each column. Dotted line indicates mean peak power for each test modality;

**c** VO<sub>2</sub>peak (L/min) according to the different test modalities used. Participant numbers are specified for each column. Dotted line indicates mean VO<sub>2</sub>peak for each test modality. Abbreviations: Tetra ACE test protocol for people with complete tetraplegia on arm-crank ergometer Para ACE test protocol for people with complete paraplegia on arm-crank ergometer; mTBRS-XT modified Total Body Recumbent Stepper Exercise Test (modified protocol); TBRS-XT Total Body Recumbent Stepper Exercise Test (standard protocol).

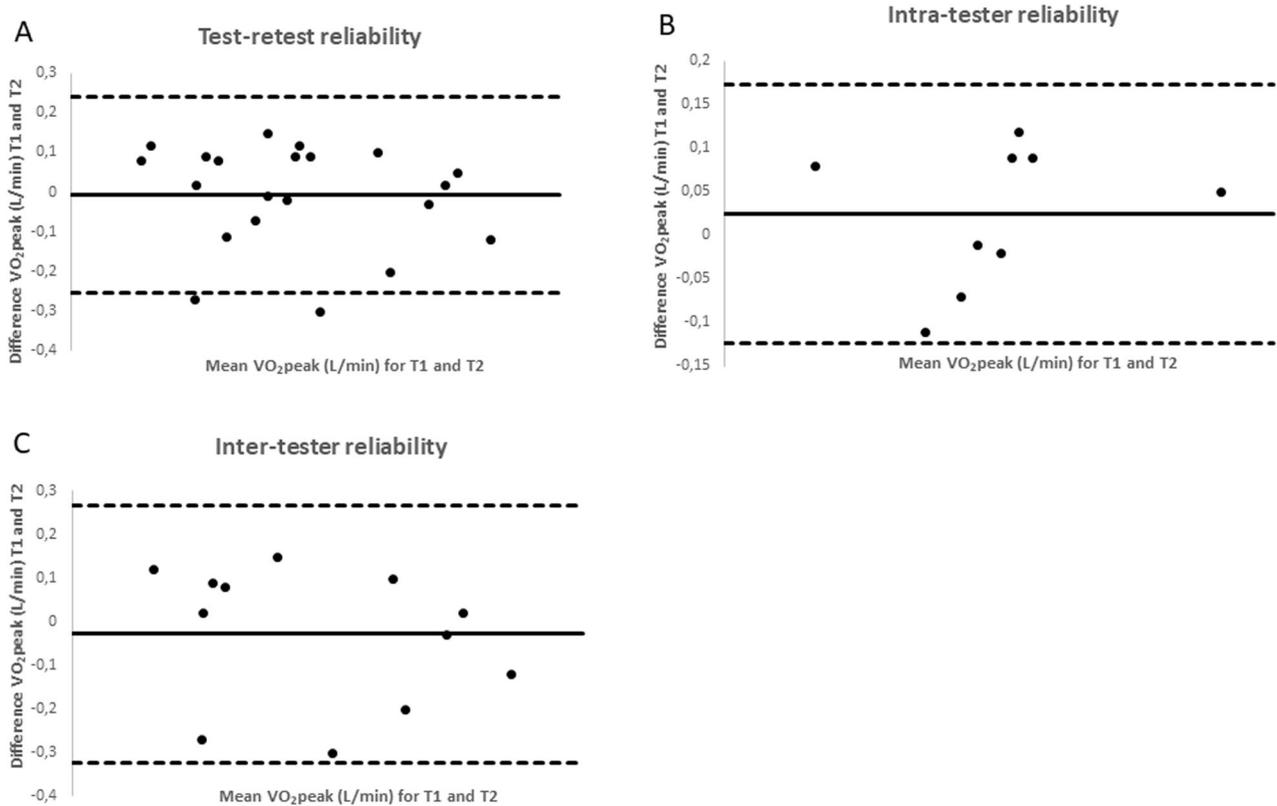
assigned to the Tetra-ACE protocol, thus ICC was not calculated.

## Discussion

This study described the content and test-retest reliability of an individualized approach for VO<sub>2</sub>peak testing developed for a clinical context. The outcomes reported reflected the recommendations for reporting by Eerden et al. including VO<sub>2</sub>peak, RER, workload at the end of the protocol and reason for termination [2]. Criteria for VO<sub>2</sub>peak were defined as the highest mean VO<sub>2</sub> (L/min) recorded over a 30 s sample, and a corresponding respiratory exchange ratio (RER) > 1.0. Although low, RER > 1.0 has been used previously and recommended as criterion in people ≥ 65 years [2, 16]. Reaching maximal VO<sub>2</sub> defined as reaching a plateau in VO<sub>2</sub> despite an increased work rate may be challenging for some patients and sedentary people, and therefore this definition was not used [16]. Participants in our study exercised until volitional exhaustion.

The protocols used were based on previous studies and a systematic review in which small increments in workload resulting in termination of the test between 8 and 12 min is recommended [2, 14, 15]. Protocol length for the maximal exercise tests described was in general 6–15 min with one study reporting termination after 4.51 min in people with tetraplegia using arm cycle ergometry [2]. Therefore we considered a protocol shift if test duration was shorter than 5 min or exceeded 15 min and volitional fatigue was not reached, even if RER exceeded 1.0.

VO<sub>2</sub>peak is highly dependent on the level and completeness of SCI as well as testing modality. This study included participants with various neurological levels and SCI severity (AIS A–D) regardless of mobility status at discharge from rehabilitation. Mean VO<sub>2</sub>peak (1.91 L/min (95%CI: 1.31–2.51), obtained at T1 (T1a or T1b) was higher than previously reported in studies on VO<sub>2</sub>peak at discharge from inpatient rehabilitation [6, 9, 22]. These studies only included wheelchair-dependent participants (AIS A–D) in the age of 18–65 years. Mean VO<sub>2</sub>peak reported in two of the studies was 1.32 L/min (±0.37) and 1.21 L/min (±0.4) in people with paraplegia, who



**Fig. 3** Bland-Altman plots for test-retest reliability. **a** Overall test-retest reliability of peak oxygen uptake (VO<sub>2</sub>peak) in individuals with spinal cord injury ( $N = 21$ ); **b** Intra-tester reliability ( $N = 9$ ); and **c** Inter-tester reliability ( $N = 12$ ). Abbreviations: First test (T1) and retest (T2).

performed a peak wheelchair test on a motor-driven treadmill [6, 9]. However, both studies had more participants included. The third study used a handcycle protocol and reported a VO<sub>2</sub>peak of 1.38 L/min ( $\pm 0.54$ ) [22]. The last study included people with an incomplete SCI but all participants were tested on a wheelchair protocol at discharge from rehabilitation resulting in a mean VO<sub>2</sub>peak of 1.05 L/min ( $\pm 0.5$ ) [23].

The higher mean VO<sub>2</sub>peak in our study was most likely due to the high proportion of participants classified as AIS D, and a high LEMS, making it possible for them to use more muscle mass during testing. This was probably the reason why the participants were categorized as “excellent” according to the SCI reference values, because these were primarily based on a sample of untrained people with motor complete SCI [21]. However, the high proportion of participants classified as AIS D in our sample was more representative for individuals with SCI capable of performing a VO<sub>2</sub>peak test during rehabilitation in the Nordic countries, where 65–80% of individuals with SCI are discharged with AIS D in Denmark and Norway [24]. In individuals not capable of performing a VO<sub>2</sub>peak test, and thus excluded from participation, limited motor function

was the most common reason as expected. VO<sub>2</sub>peak increased with more incomplete injuries supporting the rationale of using different test modalities (Fig. 2a, c) (Table 1). In participants using the Para-ACE protocol, mean VO<sub>2</sub>peak was 1.51 L/min (95%CI: 1.10–2.04), which was also slightly higher than previously reported at this timepoint [25]. Despite the higher mean VO<sub>2</sub>peak the participants were deconditioned compared to reference values for able bodied [20].

Eight participants did not reach RER > 1.0 at T1a but in four participants a protocol shift to a less demanding ACE protocol was not possible. One was already tested on the least demanding ACE protocol, and three participants had upper extremity or neck problems. Smaller increments in workload per stage could have enhanced the participants’ ability to reach VO<sub>2</sub>peak, securing a stronger relationship between workload and oxygen uptake, but the pre-programmed protocols in the recumbent stepper were not possible to adjust manually [2].

Four participants made a protocol shift to the more demanding TBRS-XT protocol. Two participants did not improve VO<sub>2</sub>peak, while the other two improved VO<sub>2</sub>peak and reached a RER > 1.0. This illustrated that it was

challenging for the testers to decide the appropriate protocol in some motor incomplete participants. However, while the effect of a protocol shift was inconclusive the shifts illustrated that even though the two recumbent stepper protocols were quite similar, the choice of protocol made a difference in those improving VO<sub>2</sub>peak illustrating the importance of individualized protocols.

The relative reliability for VO<sub>2</sub>peak in this study was high, with ICC ranging from 0.96 for the mTBRs-XT protocol to 0.99 for the Para-ACE protocol, when compared to other reliability studies involving ACE and wheelchair protocols (ICC 0.81–0.83) [26, 27]. However, the interpretation of the ICC values in this study must be taken with caution due to the heterogeneous sample of participants, which may increase the between-participants variability and improve the ICC [28].

Compared to wheelchair protocols, where wheelchair propulsion technique may have an impact on performance, the technical and biomechanical demands on the equipment used in our study seemed modest [29]. In addition, the participants in our study were accustomed to the equipment as it was used during rehabilitation, which may have contributed to limiting some of the expected variability.

Physical deconditioning which may be a major risk factor for cardiometabolic disease was present in all participants, why continual exercise testing of people with SCI throughout lifespan has been recommended [1]. In this perspective, selecting an individualized test modality may ensure that the VO<sub>2</sub>peak presumed to be most optimal is reached. The VO<sub>2</sub>peak test provided an opportunity to address the evidence-based exercise guidelines for adults with SCI, recommending moderate and hard exercise intensities [30]. A clinical observation was that many participants stated they had not exercised at this intensity during rehabilitation. Some stated that the VO<sub>2</sub>peak test, including the test result, motivated them to exercise at these intensities and that the VO<sub>2</sub>peak test provided an experience of what moderate and hard exercise intensities feel like. This observation contributes to the clinical relevance of testing VO<sub>2</sub>peak during rehabilitation.

## Limitations and strengths

This study described the test-retest reliability of VO<sub>2</sub>peak tests in a heterogeneous sample, where most participants were male classified as AIS D and tested using one of the recumbent stepper protocols. This is also a strength as it largely is in accordance with the patient population today in our SCI clinic. In this study, a low number of participants were tested in each test modality and reliability needs to be further

established in future studies with more participants included. However, testing reliability in a clinical inpatient setting with different clinicians performing the tests was a strength. A limitation of the study was the small sample size due to consecutive enrollment of participants with a new SCI who were able to perform the test at discharge from rehabilitation. However, this aspect also describes the feasibility of performing a VO<sub>2</sub>peak test among a heterogeneous group of patients in rehabilitation which is a strength.

## Conclusion

Test-retest reliability of the protocols tested was high based on ICC values and narrow limits of agreement, but further investigation of reliability is needed due to the heterogeneous sample used and the low number of participants. An individualized approach to exercise testing seemed to yield more optimal results. Mean VO<sub>2</sub>peak was higher than previously described at discharge from inpatient rehabilitation. This was probably due to a large proportion of motor incomplete participants. Thus, the sample was representative of people with SCI undergoing rehabilitation in the Nordic countries and capable of performing a VO<sub>2</sub>peak test. Despite the higher VO<sub>2</sub>peak physical deconditioning was present in all participants. While exercise testing has been recommended, future studies performed during rehabilitation should investigate the effect of individualized test modalities on VO<sub>2</sub>peak depending on motor functioning. Also, establishing the reliability of VO<sub>2</sub>peak based on the recumbent stepper test modalities, with a larger sample size, is needed.

## Data archiving

Data will be available by request from the corresponding author.

**Acknowledgements** We thank all the participants and physiotherapists Charlotte Lindholm, Tobias Benjamin Petersen and Lennart Beckwith conducting the test sessions, as well as all physiotherapists at the department contributing to the implementation of the test.

**Funding** This work was supported by a research program, “Centre for Integrated Rehabilitation of Cancer Patients (CIRE)—Neuro/Psychology”, conducted collaboratively by the University Hospitals Centre for Health Care Research, University Hospital Copenhagen, Rigshospitalet, University College Copenhagen, Department of Nursing and Nutrition, and the Neuroscience Centre, Rigshospitalet.

**Author contributions** NJH was responsible for the study, including the design and implementation of the VO<sub>2</sub>peak test. FB-S contributed with the idea of implementing a VO<sub>2</sub>peak test during rehabilitation and the

design of the approach. TM contributed with insights about the variables collected. All authors contributed with feedback on the text writing and approved the final version of the paper.

## Compliance with ethical standards

**Conflict of interest** The authors declare that they have no conflict of interest.

**Statement of ethics** We certify that all applicable institutional and governmental regulations concerning the ethical use of human volunteers were followed during the course of this research. The project is approved by the Committees on Health Research Ethics in the Capital Region of Denmark on 10.07.2018 (Journal-nr.: H-18018325). All participants received oral and written information about the study, before written consent was obtained. The project is registered at Clinicaltrials.gov (NCT03369080).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## References

- Nash MS, Groah SL, Gater DR Jr, Dyson-Hudson TA, Lieberman JA, Myers J, et al. Identification and management of cardiometabolic risk after spinal cord injury: clinical practice guideline for health care providers. *Top Spinal Cord Inj Rehabil.* 2018;24:379–423.
- Eerden S, Dekker R, Hetinga FJ. Maximal and submaximal aerobic tests for wheelchair-dependent persons with spinal cord injury: a systematic review to summarize and identify useful applications for clinical rehabilitation. *Disabil Rehabil.* 2018;40:497–521.
- Jack LP, Purcell M, Allan DB, Hunt KJ. Comparison of peak cardiopulmonary performance parameters during robotics-assisted treadmill exercise and arm crank ergometry in incomplete spinal cord injury. *Technol Health Care.* 2010;18:285–96.
- DiPiro ND, Embry AE, Fritz SL, Middleton A, Krause JS, Gregory CM. Effects of aerobic exercise training on fitness and walking-related outcomes in ambulatory individuals with chronic incomplete spinal cord injury. *Spinal Cord.* 2016;54:675–81.
- Gorman PH, Geigle PR, Chen K, York H, Scott W. Reliability and relatedness of peak VO<sub>2</sub> assessments during body weight supported treadmill training and arm cycle ergometry in individuals with chronic motor incomplete spinal cord injury. *Spinal Cord.* 2014;52:287–91.
- Haisma JA, Bussmann JB, Stam HJ, Sluis TA, Bergen MP, Dallmeijer AJ, et al. Changes in physical capacity during and after inpatient rehabilitation in subjects with a spinal cord injury. *Arch Phys Med Rehabilitation.* 2006;87:741–8.
- Nooijen CF, de Groot S, Postma K, Bergen MP, Stam HJ, Bussmann JB, et al. A more active lifestyle in persons with a recent spinal cord injury benefits physical fitness and health. *Spinal Cord.* 2012;50:320–3.
- Hjeltmes N, Wallberg-Henriksson H. Improved work capacity but unchanged peak oxygen uptake during primary rehabilitation in tetraplegic patients. *Spinal Cord.* 1998;36:691–8.
- van Koppenhagen CF, Post M, de Groot S, van Leeuwen C, van Asbeck F, Stolwijk-Swuste J, et al. Longitudinal relationship between wheelchair exercise capacity and life satisfaction in patients with spinal cord injury: A cohort study in the Netherlands. *J Spinal Cord Med.* 2014;37:328–37.
- Spieß MR, Müller RM, Rupp R, Schuld C, Group E-SS, van Hedel HJA. Conversion in ASIA impairment scale during the first year after traumatic spinal cord injury. *J Neurotrauma.* 2009;26:2027–36.
- Halvorsen A, Pettersen AL, Nilsen SM, Halle KK, Schaanning EE, Rekan T. Epidemiology of traumatic spinal cord injury in Norway in 2012–2016: a registry-based cross-sectional study. *Spinal Cord.* 2019;57:331–8.
- Holm NJ, Møller T, Adamsen L, Dalsgaard LT, Biering-Sorensen F, Schou LH. Health promotion and cardiovascular risk reduction in people with spinal cord injury: physical activity, healthy diet and maintenance after discharge- protocol for a prospective national cohort study and a preintervention- postintervention study. *BMJ Open.* 2019;9:e030310.
- Kirshblum SC, Burns SP, Biering-Sorensen F, Donovan W, Graves DE, Jha A, et al. International standards for neurological classification of spinal cord injury (revised 2011). *J Spinal Cord Med.* 2011;34:535–46.
- Billinger SA, Loudon JK, Gajewski BJ. Validity of a total body recumbent stepper exercise test to assess cardiorespiratory fitness. *J Strength Cond Res.* 2008;22:1556–62.
- Billinger SA, Tseng BY, Kluding PM. Modified total-body recumbent stepper exercise test for assessing peak oxygen consumption in people with chronic stroke. *Phys Ther.* 2008;88:1188–95.
- Edvardsen E, Hem E, Anderssen SA. End criteria for reaching maximal oxygen uptake must be strict and adjusted to sex and age: a cross-sectional study. *PLoS ONE.* 2014;9:e85276.
- Macfarlane DJ, Wong P. Validity, reliability and stability of the portable Cortex Metamax 3B gas analysis system. *Eur J Appl Physiol.* 2012;112:2539–47.
- Biering-Sørensen F, DeVivo MJ, Charlifue S, Chen Y, New PW, Noonan V, et al. International spinal cord injury core data set (version 2.0)-including standardization of reporting. *Spinal Cord.* 2017;55:759–64.
- Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med.* 2016;15:155–63.
- Aspenes ST, Nilsen TIL, Skaug E-A, Bertheussen GF, Ellingsen Ø, Vatten L, et al. Peak oxygen uptake and cardiovascular risk factors in 4631 healthy women and men. *Med Sci sports Exerc.* 2011;43:1465–73.
- Simmons OL, Kressler J, Nash MS. Reference fitness values in the untrained spinal cord injury population. *Arch Phys Med Rehabil.* 2014;95:2272–8.
- Nooijen CF, Stam HJ, Bergen MP, Bongers-Janssen HM, Valent L, van Langeveld S, et al. A behavioural intervention increases physical activity in people with subacute spinal cord injury: a randomised trial. *J Physiother.* 2016;62:35–41.
- Dallmeijer AJ, van der Woude LH, Hollander AP, van As HH. Physical performance during rehabilitation in persons with spinal cord injuries. *Med Sci Sports Exerc.* 1999;31:1330–5.
- Halvorsen A, Pettersen AL. NordicSCIR Årsrapport 2018. Available from: <https://stolav.no/Documents/NordicSCIR%20%C3%85rsrapport%202018.doc.pdf>.
- Haisma JA, van der Woude LHV, Stam HJ, Bergen MP, Sluis TAR, Bussmann JBJ. Physical capacity in wheelchair-dependent persons with a spinal cord injury: a critical review of the literature. *Spinal Cord.* 2006;44:642–52.
- Hol AT, Eng JJ, Miller WC, Sproule S, Krassioukov AV. Reliability and validity of the six-minute arm test for the evaluation of cardiovascular fitness in people with spinal cord injury. *Arch Phys Med Rehabilitation.* 2007;88:489–95.
- Gauthier C, Arel J, Brosseau R, Hicks AL, Gagnon DH. Reliability and minimal detectable change of a new treadmill-based progressive workload incremental test to measure cardiorespiratory

- fitness in manual wheelchair users. *J Spinal Cord Med.* 2017;40:759–67.
28. Weir JP. Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. *J Strength Cond Res.* 2005;19:231–40.
29. de Groot S, de Bruin M, Noomen SP, van der Woude LHV. Mechanical efficiency and propulsion technique after 7 weeks of low-intensity wheelchair training. *Clin Biomech (Bristol, Avon).* 2008;23:434–41.
30. Martin Ginis KA, van der Scheer JW, Latimer-Cheung AE, Barrow A, Bourne C, Carruthers P, et al. Evidence-based scientific exercise guidelines for adults with spinal cord injury: an update and a new guideline. *Spinal Cord.* 2017;56:308–321