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.

Sophia Eerden, Msc, a Rienk Dekker, PhD, b,c Florentina J. Hettinga, PhD d

a. University of Groningen, University Medical Center Groningen, Center for Human Movement Sciences, Groningen, The Netherlands
b. Department of Rehabilitation Medicine, Center for Rehabilitation, University Medical Center Groningen, Groningen, The Netherlands
c. Center for Sports Medicine, University Medical Center Groningen, Groningen, The Netherlands
d. University of Essex, School of Biological Sciences, Centre of Sport and Exercise Science, Colchester, United Kingdom

Corresponding author:
Florentina J. Hettinga
Address: Wivenhoe Park CO3 4SQ Colchester
Telephone: +441206872046
E-mail: fjhett@essex.ac.uk
**Purpose:** To summarize the available maximal and submaximal aerobic exercise tests for wheelchair-dependent persons with a spinal cord injury and to identify useful applications for clinical rehabilitation.

**Method:** The databases of PubMed, CINAHL®, EMBASE and PsycINFO® were searched for English-language studies published prior to March 2015. Two independent raters identified and examined studies that reported on laboratory-based aerobic exercise tests in persons with a spinal cord injury, according to the PRISMA statement.

**Results:** The test protocols of maximal (n = 105) and submaximal (n = 28) exercise tests, covered by 95 included studies, were assessed. A large variety in patient characteristics, test objectives, test protocols, exercise modes and outcome parameters was reported. Few studies reported on adherence to recommendations, adverse events and peak outcome validation.

**Conclusion:** An incremental test protocol with small, individualized, increments per stage seems preferable for testing maximal aerobic capacity, but additional validation of the available test modes is required to draw conclusions. Submaximal testing is relevant for assessing the performance at daily life intensities and for estimating VO$_2$peak. Consensus regarding reporting test procedures and outcomes needs to be achieved to enhance comparability of rehabilitation results.

**Keywords:** cardiopulmonary exercise test; rehabilitation outcome; wheelchair; upper extremity; spinal cord injuries.
INTRODUCTION

Individuals with a spinal cord injury (SCI) have difficulties to engage in physical activities since they experience poor accessibility and fewer opportunities to be physically active. As a result, these persons often show lower physical activity levels when compared with ambulatory individuals and, consequently, are at risk for the development of medical complications (1-3). Increasing the aerobic capacity of persons with a SCI during rehabilitation is essential for the prevention of low physical fitness levels (4). In order to monitor and optimize effects of rehabilitation training it is recommended to quantify changes in the aerobic capacity of patient with SCI during rehabilitation (5). To do so, it is important that the characteristics of the available aerobic exercise tests for individuals with a SCI are explored and judged on their applicability in the rehabilitation practice. The current review will therefore summarize the available maximal and submaximal exercise tests for wheelchair-dependent persons with a SCI.

Over the past few decades, a variety of different upper-body exercise testing modes and protocols has been conducted in the SCI population. As indicated in the study of Valent et al. (2007), differences in exercise test designs to measure physical capacity might influence the test results. The validity of the reported improvements in peak oxygen uptake (VO$_{2\text{peak}}$) and peak power output (PO$_{\text{peak}}$) after training is therefore questionable (6). The VO$_{2\text{peak}}$ and PO$_{\text{peak}}$ parameters are, according to the American College of Sports Medicine (ACSM), considered to be the gold standard for indicating a persons’ peak physical capacity (7, 8). The disparity in testing protocols and outcomes hampers the process of interpreting findings, makes it difficult to compare trends across studies, and impedes generalization of the results to the larger SCI population (9). At the same time, the implementation of evidence-based practice in SCI health care has become increasingly important over the past ten years. Furthermore, as pointed out by De Groot et al. (2010), there is a strong basis for
implementing standardized tests in SCI rehabilitation centers, which emphasizes the practical possibilities of the development of a standardized aerobic exercise test (5). These findings emphasize the necessity of a thorough evaluation of the available aerobic exercise test for people with a SCI, as a first step towards the development of standardized testing.

Regular testing a patient with SCI throughout the rehabilitation process with a standardized aerobic exercise test can provide very valuable information. It enables rehabilitation professionals to monitor and evaluate the patients’ progress and to make specific adjustments in the training program. This adequate training will support patients in the performance of daily life activities, which is an important goal for rehabilitation, and it would contribute to improve rehabilitation outcomes (4, 5). In order to develop evidence-based exercise and fitness monitoring in rehabilitation practice, a first step is to explore the available aerobic testing protocols that have been applied in the SCI population. Therefore, the aims of this systematic review are to summarize the available maximal and submaximal aerobic exercise tests for wheelchair-dependent persons with a SCI (i) and to identify useful applications for clinical rehabilitation (ii).

**METHODS**

**Search strategy**

This systematic review was conducted in accordance with the recommendations of PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-analysis) (10). The electronic databases of PubMed, CINAHL®, EMBASE and PsycINFO® were systematically searched on studies published prior to May 2013. An updated search was performed in March 2015 and May 2016, by using the same search strategy. A comprehensive search strategy was built,
consisting of a combination of database-specific MeSH terms, free text, ‘wild cards’ (words truncated by using “*”) and Boolean operators (“AND”, “OR”, “NOT”). The search was structured into three parts, with the first part concerning population keywords (spinal cord injury, paraplegia, tetraplegia, wheelchair). The second part of the search strategy refers to studies about wheelchair propulsion-related aerobic exercise tests. The used keywords were exercise test, maximal, submaximal, physiologic fitness and training. For the third part of the search, that covered the possible outcome measures of exercise tests, keywords were i.e. oxygen consumption, power output and heart rate. All three parts of the search were combined using the Boolean operator “AND”. Retrieved papers (n = 1211) were combined in a single database and duplicates (n = 191) were removed.

**Study selection**

In order to be included in the current review, studies had to meet the following criteria: (1) >80% of the experimental study group has a SCI, (2) a laboratory-based aerobic exercise test is included, (3) and a description of the initial settings and stages of the testing protocol is provided. Exclusion of studies occurred if they only reflected on anaerobic testing, body weight support training, respiratory training, functional electrical stimulation, quality of life assessment, body temperature examination, activities of daily living, electromyography, electrocardiography, homeostatic processes or metabolic responses, since these outcomes were not directly related to physical capacity. Additionally, any other type of article than an experimental or observational research article was excluded, including a review of the literature or a comment to the editor.

**Screening**
The flow diagram of literature searches and results is shown in figure 1. After removing duplicates, 1020 articles were identified. In the first and second screening stage, two authors (RD and SE) independently screened the titles and abstracts respectively, according to inclusion and exclusion criteria. In case of persisting disagreement during any of these two assessment phases, a third observer (FJH) gave a binding verdict. Agreement between the authors during the title- and abstract assessment phase, expressed with Cohens Kappa, was $\kappa = 0.572$ and $\kappa = 0.487$ ($p < 0.001$) respectively. Full agreement (100%) was achieved during a consensus meeting that was held for each phase. Ninety-six articles were retained for full text assessment, but nine of these 96 articles were unavailable despite several attempts of the authors to retrieve them. In the second screening stage, 87 articles were read by RD and SE and were included when both reviewers felt they met all the inclusion criteria. Subsequently, 24 were excluded based on these inclusion criteria. Respectively three and four more articles were included after the updated searched in March 2015 and May 2016. Additionally, 25 eligible articles were found after checking the reference lists. A total of 95 articles were rated as eligible to be included for review.

Data extraction

The three authors together established a data extraction form. Author SE completed these data extraction forms for the included 95 studies accordingly. Relevant study characteristics were extracted and described: (i) population characteristics, (ii) the test protocol used to conduct the aerobic exercise test and termination guidelines referred to, (iii) the criteria used to determine maximal performance, (iv) adverse events during testing and (v) key measurement
outcomes reported, namely oxygen uptake, power output, respiratory exchange ratio and heart rate.

RESULTS

A total of 89 incremental maximal exercise tests, 14 intermittent maximal exercise tests, 21 constant load maximal exercise tests and 28 submaximal exercise tests were conducted among the 95 included studies. The extracted study and population characteristics are shown in Table 1. Table 2 and 3 present the protocol details and outcomes for the maximal aerobic tests and submaximal aerobic tests, respectively.

Insert table 1.

Patient characteristics

Based on 95 articles, a total of 2,725 participants were included in the analysis. The number of participants included in a study ranged from 1 (33) to 185 (4). Mean age ranged from 24 (33, 66) to 50.0 (50) years. Most studies included more men than women, but 46 studies included only men. Mean time since injury (TSI) ranged from 78 days (52) to 28.7 years (25) and lesion level ranged from C1 (82) to S2 (43, 45). Forty-four studies included only persons with a paraplegia, whereas 10 studies only included persons with a tetraplegia. Forty studies described both persons with a tetraplegia and paraplegia. One study did not report on the lesion level of the included participants. Completeness of the injury was assessed in 67 of the 95 studies. A total of 46 of these 95 studies included both subjects with a complete and incomplete lesion, whereas 21 studies included solely persons with a complete lesion. The
reported fitness of the participants ranged from persons with a low physical fitness status (rehabilitants, sedentary, untrained and inactive people) to persons with a high physical fitness status (athletes, active, trained people).

Study designs

In the majority of the included studies, a single measure design was applied (n = 44). Twenty-two studies were registered as a pre-post training design, whereas 17 studies conducted repeated measures. Nine studies applied a prospective cohort design, of which eight studies were the result of the cohort study titled ‘Physical strain, work capacity, and mechanisms of restoration of mobility in the rehabilitation of persons with spinal cord injuries’. Other study designs were registered as well, including a randomized controlled trial (n = 2) and a case study (n = 1). Sixteen of the included studies included a control group in the study design, which consisted of either persons with a SCI (n = 3) or able-bodied persons (n = 13). The remaining 79 studies did not include a control group.

Test objectives

The main test objectives identified for the aerobic exercise tests were to determine physiological responses (max: n = 48, submax: n = 8), to assess the effect of training or rehabilitation on physical capacity (max: n = 26, submax: 5) or to describe the relationship between two parameters (max: n = 13, submax: 4). Other identified objectives were to screen for contraindications for training (max: n = 1), to determine VO2peak for additional training or testing protocols (max: n = 2), to examine the reliability of the six-minute push test (max: n = 1) and a graded submaximal test (submax: n = 1), to determine measurement properties of fitness measures (n = 1), to determine increments per stage for a subsequent maximal test (submax: n = 1) or to determine the a steady state submaximal performance submax: n = 1).
The test objective of seven submaximal tests was not reported.

**Exercise modes**

In 52 of the 105 performed maximal exercise tests, an arm crank ergometer was used to conduct the exercise test. The wheelchair ergometer was used in 44 tests and the hand cycle in 6 tests. Other identified exercise modes were supine arm crank ergometry (n = 1), arm tracking, which is a dual action exercise ergometer, (n = 1) and seated double poling ergometry (n = 1). For conducting the 28 submaximal exercise tests, wheelchair ergometry (n = 13) and arm crank ergometry (n = 10) were used, as well as the hand cycle (n = 3), supine arm crank ergometry (n = 1) and seated double poling ergometry (n = 1).

When relating the identified aerobic fitness indications to the used exercise modes, it appears that active or trained participants were involved in 35% of the studies that used wheelchair ergometry, rehabilitants in 30% of these studies, athletes in 21% and inactive or untrained participants in 5% of these studies. The aerobic fitness indication was not reported in 9% of the studies. For the arm crank ergometry, somewhat similar results were found, but fewer rehabilitants were involved in these studies (14%) and a higher number of studies did not report on aerobic fitness indication (29%). For hand cycling studies, active participants (67%) and rehabilitants (33%) performed the exercise tests.

Test protocols

A warm-up was performed prior to the actual test protocol in 42 maximal exercise tests and six submaximal exercise tests. The warm-ups had a duration of one to five minutes and were
performed at zero or low resistance loads. The reported propulsion speed ranged from 3 to 8.5 km/h or 50-60 rpm.

For most maximal exercise test protocols, the time to exhaustion varied between six to 15 minutes. The shortest time to exhaustion was found in the study of Lasko-McCarthy & Davis (1991), in which the tests was ended after 4.51 minutes (69). The study of McLean et al. (1995) reported the longest time to exhaustion of over 20 minutes (78). This study involved an intermittent maximal test protocol in which exercise periods were alternated with 80 seconds rest periods.

Three different maximal test protocols was used, namely incremental, intermittent and constant load maximal exercise tests. These protocols will now be further described, as well as the test protocols of submaximal exercise tests.

**Incremental maximal exercise tests.** Four different test protocols were described for the incremental tests. Most of these tests (n = 68) increased activity by increasing loads or resistance. The size of these increments ranged from 3 to 15W per 1 to 3 minutes for tests conducted with a wheelchair ergometer. For the tests using arm crank ergometry and hand cycling, step sizes ranged from 2W to 30W with step duration ranging from 1 to 3 minutes. Several studies used different incremental steps, depending on the participants’ lesion level (20, 50, 57, 67-69, 74, 81, 82, 84, 94, 98). Participants were instructed to keep up with a certain speed, which was set at 2-5 km/h for the majority of wheelchair ergometry test and at 50-60 rpm for tests conducted with arm crank ergometry.

Other studies described a test protocol in which physical demands were increased by slope gradient inclination (n = 12). Most of these studies using such a protocol applied the protocol as described by Kilkens et al. (2004) (105). This protocol involves starting at a propulsion speed of 2, 3 or 4 km/h, depending on the lesion level, and increments in slope gradient of 0.36° per minute. Eight studies used a protocol similar to the protocol used in the
This protocol describes an increment in speed until a certain speed was reached. Subsequently, load was added or slope gradient was increased in order to increase the physical demands. One study used a speed-graded protocol.

Intermittent maximal exercise tests. The physical demands in all 14 intermittent test protocols were increased by increments in load per stage. The increments were mostly between 2W and 10W, but two studies reported on increments of 15W per stage. The propulsion speed was comparable to the incremental test protocols, with 3-8 km/h for tests performed in a wheelchair ergometer or hand cycle and 50-70 rpm for tests that used arm crank ergometry. In all intermittent protocols, the period of exercise was longer (2-4 min) than the period of rest (30s - 3 min). The rest period allowed for blood lactate, blood pressure and RPE measurements. Two studies applied an intermittent protocol because it prevents for arm fatigue and would therefore result in higher peak aerobic values.

Constant load maximal exercise tests. In the two studies that used wheelchair ergometry, no increments per stage were applied but participants had to propel at a maximal tolerated constant load, while keeping a speed of 4.5 or 5.5 km/h.

Submaximal exercise tests. Two types of submaximal test protocols were identified: those with increments in physical demands (20 tests) and those without increments (8 tests). The physical demands were increased by adding load (11 tests), increasing the slope gradient with 0.36° (7 tests), or increasing heart rate with 15 bpm or 20%HR_{max} (2 tests). Load increments ranged from 5 to 30W, or were set at 20%PO_{est}, 30% of Maximal Tolerated Power (MTP) or 75 kpm. The number of stages varied among the submaximal tests. The protocol of six tests consisted of one stage, 11 tests applied two stages of exercise in the test protocol, seven tests included three stages and four tests consisted of five or six stages. Stage duration ranged from 2 to 7 minutes and these stages of exercise were alternated with periods of 1 to 12 minutes rest in 11 of the 28 submaximal protocols.
Adherence to guidelines

Pre-test screening procedures were reported by 35 studies. The screening was usually performed by a physician and involved medical examination, an ECG and spirometry. Other reported procedures were conducting a health questionnaire or obtaining a medical history. Five studies referred to the ACSM guidelines and one study referred to the American Thoracic Society for pre-test screening procedures (14, 21, 32, 65, 91).

There were two reasons identified to terminate a maximal exercise test: when a patient becomes symptomatic and when the patient has reached maximum effort. Nineteen tests applied symptom-limited test termination criteria of which ten referred to the ACSM guidelines. The other nine tests used ECG abnormalities, blood pressure drop, dysreflexia, or adverse symptoms as criteria. Maximal effort was reported in 81 tests as termination reference, with volitional exhaustion (n = 32), unable to maintain speed or load (n = 21) or both the latter (n = 28) as criteria. Five studies did not report on termination guidelines.

Adverse events

Of the nine studies that reported on clinical abnormalities during maximal testing, five reported no clinical abnormalities. Three studies reported on relevant abnormalities in three patients, which included a fall in systolic blood pressure during cooling down, inability to keep up with the speed and bradycardia and hypotension after testing (37, 76, 92). For one subject, $PO_{peak}$ could not be determined due to unknown reason (27).
For submaximal testing, two studies reported on adverse events, which were the
inability to maintain 3 minutes of propulsion (2 persons) and mild muscle spasms during
cycling (4 persons) (53, 93). One study reported no adverse events (104).

**Peak outcomes**

Thirty studies described criteria for reaching a valid VO\textsubscript{2peak}. The criteria used included
attainment of the age-predicted maximal heart rate (APMHR) (n = 16), RER above a certain
level (>1.0-1.15) (n = 21), VO\textsubscript{2} plateau despite an increase in work rate (n = 17) and blood
lactate above a certain level (> 8-10 mmol/l) (n = 4). Four studies opted for a supra-maximal
protocol in order to verify the attained peak VO\textsubscript{2}. Other criteria were similar to the previously
described termination guidelines, including exhaustion or inability to maintain speed or load
(n = 5). One study referred to the ACSM guidelines (71).

Approximately half of the studies (n = 16) also reported the number of people who
met the predefined criteria. The number of participants reaching a VO\textsubscript{2} plateau was reported
by eight studies, with 60% to 100% reaching the plateau. Defined criteria related to RER,
APMHR and blood lactate were met by 80% to 100% of the participants.

Varying outcomes in VO\textsubscript{2peak} were reported in the included studies. For tests
performed in a wheelchair ergometer, the mean reported VO\textsubscript{2peak} of all included studies was
24.2 ml/kg/min with ranging values from 7.5 to 40.4 ml/kg/min. Mean value (19.21
ml/kg/min) and range (8.8-38.1 ml/kg/min) were comparable for tests using arm crank
ergometry or hand cycling. The lowest values were found in untrained participants with
cervical lesions (25, 106), whereas the highest values were found in trained participants with a
paraplegia (19, 74). Some studies reported VO\textsubscript{2peak} in l/min, with values ranging from 0.55 to
2.35 l/min (13, 24).
The majority of P_{O,peak} outcomes was expressed in Watts with a mean P_{O,peak} of 56.4W (11-210W) for wheelchair ergometry tests and 66.5W (15-159W) for tests using arm crank ergometry or hand cycling. The lowest reported value was 11W, found in a group of participants with high cervical lesions (24). The highest reported P_{O,peak} was 210W, found in the same group of participants that reported the highest V_{O,2,peak} value using wheelchair ergometry (74). Other reported outcome measures for P_{O,peak} were W/kg (0.15-1.11 W/kg), kgm/min (255-653 kgm/min) and kpm/min (141-761 kpm/min) (25, 34, 35, 61, 63, 98). The mean and ranging values for RER and H_{R,peak} were 1.19 (0.92-1.44) and 155 bpm (96-198 bpm), respectively.

Submaximal outcomes

Reported submaximal V_{O,2} means ranged from 9.3-13.1 ml/kg/min and 0.74-1.90 l/min, with overall mean values of 11.2 ml/kg/min and 1.16 l/min respectively. Mean P_{O}, RER and H_{R,peak} values were 46.0W (17.7-78.4W), 0.92 (0.88-0.96) and 116 bpm (97-166 bpm), respectively.

DISCUSSION

The aim of this systematic review was to summarize the available maximal and submaximal aerobic exercise tests for wheelchair-dependent persons with a SCI. The identified exercise tests showed a large variety in population characteristics, exercise modes, testing protocols and outcome measures. Limited studies reported on adherence to recommendations, adverse events and oxygen uptake validation. Possible useful applications of the available maximal and submaximal aerobic exercise tests for clinical SCI rehabilitation will be discussed.
Exercise mode

Arm crank ergometry and wheelchair ergometry were the most commonly used exercise modes among the included studies. $P_{\text{O}_{\text{peak}}}$ and $V_{\text{O}_{2\text{peak}}}$ comparisons between both modalities showed no difference in $V_{\text{O}_{2\text{peak}}}$, but a somewhat higher $P_{\text{O}_{\text{peak}}}$ for arm crank ergometry. This is in line with previous studies in which a group of persons with a paraplegia performed a maximal exercise test in both modes (76, 107). Additionally, two studies that only compared $V_{\text{O}_{2\text{peak}}}$ outcomes for both modes reported no differences in $V_{\text{O}_{2\text{peak}}}$ as well (44, 77).

Although no adverse events of musculoskeletal problems were reported, previous literature indicated that wheelchair ergometry was usually more straining to the musculoskeletal system than arm crank ergometry and hand cycling. Wheelchair ergometry puts the participant to a higher risk for over-use problems of the upper-extremities (29, 76, 108, 109). On the contrary, wheelchair ergometry has excellent application opportunities for submaximal testing in SCI rehabilitation, since it provides relevant data of wheelchair performance and mobility in daily life (110). Exercise modes that are more suitable for maximal exercise testing in clinical rehabilitation are arm crank ergometry and hand cycling. Both modes allow for continuous force application and no peak loads occur during propulsion. The hand cycle mode was found to be highly relevant for training and testing the peak cardiovascular capacity and fitness during rehabilitation, and it was demonstrated that exercise intensities as prescribed by the ACSM guidelines could be attained (29, 92, 111). Notwithstanding, further research is necessary on how hand cycling can be optimally used for training and testing in the SCI rehabilitation setting (112, 113).

The final choice of equipment depends on the goal of the test and of the participants’ ability. For example, when designing a test for rehabilitants, the arm crank ergometer and hand cycle are recommended for determining peak physical capacities during maximal
exercise testing, whereas the more task-specific hand-rim wheelchair propulsion has a higher relevance for submaximal testing and assessing daily life performance (110).

Test protocols

In order to attain the peak physical abilities during an aerobic maximal exercise test, it is important to determine the increments per stage carefully. This is especially true for those who are rehabilitating from a SCI, since these people are often vulnerable and sensitive to overuse problems (27, 114, 115). When large increments per stage are applied, the relationship between oxygen uptake and workload is usually weaker. Therefore, it is recommended to use small to modest individualized increments per stage, resulting in completion of the test between 8 and 12 minutes (7, 116). The results revealed two common ways of increasing the physical demands during incremental testing. One way is to add resistance each stage (5W-10W), with lower amounts of resistance increments for those with a high lesion level. Another option is to increase the slope gradient per stage (0.36°), while fixing the belt velocity at a certain speed (2 or 3 or 4 km/h) depending on the physical capacity of the patient. The duration of the stage should be between 60s and 120s. Both protocol types seem to be feasible and can be executed with any exercise mode. However, one should take into account that performing a maximal exercise test has some practical limitations for clinical rehabilitation. For example, if the slope gradient is getting too steep during testing, the patient could be forced to quit because of muscular failure rather than cardiovascular failure. A sudden termination of the test could cause the patient to roll backwards on the treadmill. When opting for increasing the resistance by using a pulley system, instead of increasing the slope gradient, these practical limitations do not apply. In fact, the posture of the patient does not change while using a pulley system to increase the physical demands and this system allows for a larger variety in increments per stage. Because
of these practical advantages, it would be preferred to opt for increasing the resistance by using a pulley system in a clinical rehabilitation setting, rather than increasing the slope gradient of the treadmill.

Adherence to guidelines

In previous review studies it was found that exercise testing in patient groups does not always comply with exercise testing guidelines (117, 118). This is line with the findings of the present review, in which only five studies referred to the common accepted ACSM guidelines for exercise testing. These guidelines recommend pre-test screening for identifying contraindications for maximal exercise and it is obvious that all participants should have a pre-test screening. A pre-test screening was, however, reported in only 35 of 95 of the included studies in the current review. In the future, inclusion- and exclusion criteria should be clearly described, pre-test screening should be performed and participants should be monitored during the test. Approval of the involved physician, responsible for the treatment, should be an additional criterion for SCI patients. Test termination criteria used in the included studies were all in accordance with ACSM guidelines.

For participants who cannot sustain incremental exercise due to safety reasons of physical limitations, it is recommended to conduct an intermittent test protocol. Such a protocol allows for the prevention of muscle fatigue, but also for monitoring blood pressure measurement (14). In case intermittent exercise is not feasible either, the maximal aerobic capacity can be estimated from submaximal testing outcomes (110).

Reporting outcomes
The reported peak values are difficult to interpret, since 30 studies described criteria for reaching a valid peak oxygen uptake. Of these 30 studies, only 16 studies reported the number of participants who satisfied these criteria. The primary criterion for $V\text{O}_{2\text{peak}}$ is the achievement of a $V\text{O}_2$ plateau despite an increase in work rate (7, 119). The use of this criterion is, however, questionable, since more than one plateau can be achieved during incremental exercise or the plateau cannot be found (119-121). In case a $V\text{O}_2$ plateau could not be determined, Edvardsen et al. (2014) recommend the use of an RER cut-off value (>1.0-1.15) as criterion for attaining $V\text{O}_{2\text{peak}}$ (119). This recommendation is in line with findings of the current review.

Several studies used the attainment of the APMHR as a criterion for maximal effort, but the use of this criterion in the SCI population is questionable, since the sympathetic innervation of the heart derives from T1 to T4. Persons with a lesion at or above T4 might show a non-linear relation between HR and $V\text{O}_2$ (84, 89). The attainment of APMHR is therefore not recommended as a criterion for attaining a valid $V\text{O}_2$.

There are currently no guidelines available for reporting outcomes of exercise testing for any clinical population (117). It is, however, recommended to report peak oxygen uptake and power output values, since these two parameters were identified as primary outcome measure in a previous literature study regarding persons with a SCI. Furthermore, it is recommended to report on $V\text{O}_2$ plateau and mean RER measures (116, 118). Additionally, in order to enhance comparability of clinical rehabilitation outcomes, the criteria and reasons for test termination should be reported and results need to be compared with norm scores for persons with a paraplegia and tetraplegia.

**Implications for rehabilitation**
• Regularly testing the cardiovascular capacity during SCI rehabilitation will enable us to monitor the impact of rehabilitation interventions on an individual level.

• The incremental arm ergometry test with small increments per stage is most relevant for the assessment of the peak cardiovascular capacity.

• For the assessment of daily life functioning, the submaximal wheelchair ergometer test is preferable.

• Hand cycling is a promising exercise mode for both testing and training.

• Systematically reporting on test termination, criteria for attaining valid peak outcomes and adverse events is necessary to enhance comparability of results.

Limitations and recommendations

A few limitations need to be taken into account when interpreting the results of the current review. First of all, it might be possible that some studies using an aerobic exercise test in the SCI population have been missed, even though a comprehensive search was conducted. We are however confident that the results and conclusions are representative, given the large number of 95 included studies. A disadvantageous effect of the broad inclusion strategy, however, is the wide diversity found regarding study methods and populations, which makes it more difficult to draw conclusion. At the same time, this latter issue is contradicted by the fact that persons with a SCI with all kinds of fitness levels, from rehabilitant to athlete, are represented in the current study.

The current review provides some guidance for creating an evidence-based standardized aerobic exercise test, but it should be noted that measuring peak cardiorespiratory abilities is only one part of the total physical capacity when referring to the ACSM definition of physical fitness. The ACSM identified several components of physical
fitness in addition to cardiorespiratory fitness, namely body composition, flexibility, muscular strength and muscular endurance (7). In order to attain a full understanding of a patients’ physical capacity, it is necessary to measure these other components as well (6).

An important factor for research in the context of using exercise testing as a means of evaluating training or active lifestyle interventions is the use of a control group in the study design. In only 12 of the 68 studies in the present review, of which two studies were a randomized controlled trial, a control group was included. Although establishing a control group is often complicated in SCI research due to the absence of an unlimited source of persons with a SCI and the existing heterogeneity in this population, it should be encouraged to establish larger subject groups, and thus statistical power, in future studies. A possibility could be conducting structured training and testing programs in able-bodied persons, since their physiological stress and strain appears to be comparable for those with a paraplegia (112). Furthermore, by introducing multicenter collaboration, outcomes of various training and testing procedures can be evaluated systematically in a homogeneous group as well (6).

Another option is to perform a multilevel analysis to compare groups of patients with SCI. This statistical analysis technique, that was applied in a recent longitudinal cohort study on physical activity behavior in patients, allows for missing values and can correct for differences at the level of rehabilitation center (122).

The current review showed various opportunities for the application of exercise testing in SCI rehabilitation. However, the findings did not enable us to describe the most preferable test protocol for maximal and submaximal testing. Future research should therefore focus on validating the different exercise modes. Furthermore, practical limitations should be considered and consensus regarding reporting outcomes needs to be achieved.

CONCLUSION
This systematic review can be seen as a first step in the development of a standardized aerobic exercise test for daily SCI rehabilitation practice. An extensive variety in population characteristics, exercise modes, testing protocols and outcome measures was revealed. Limited studies reported on adherence to recommendations, adverse events and oxygen uptake validation. An incremental test protocol with small, individualized increments per stage seems preferable, but additional validation of the exercise modes is required to draw definitive conclusions. Submaximal testing is relevant for assessing the performance at daily life intensities and for estimating VO_{2peak}. Furthermore, consensus regarding reporting test procedures and outcomes needs to be achieved to enhance comparability of rehabilitation results.

DECLARATION OF INTEREST

We can confirm that there are no known conflicts of interest associated with this publication and there has been no significant financial support for this work that could have influenced this outcome. The manuscript has been read and approved by all named authors.

REFERENCES


