Doctoral thesis

HandbikeBattle

A challenging handcycling event

A study on physical capacity testing, handcycle training and effects of participation

Ingrid Kouwijzer

Promotores
Prof. dr. L.H.V. van der Woude
Prof. dr. M.W.M. Post

Copromotor
Dr. L.J.M. Valent

Table of contents

Chapter 1

General introduction

Sport and exercise in rehabilitation

In 1944, prominent neurosurgeon Ludwig Guttmann started treating veterans from World War II at the Stoke Mandeville Spinal Injuries Unit in England. At that time, permanent hospitalization and low survival rate were accepted for people with a spinal cord injury (SCI). Guttmann was rebelling against this common practice and introduced sport and exercise into the rehabilitation program. He believed that sport and exercise played a vital role in physical and mental rehabilitation and that it was a pathway that could help people with a disability to be happier, achieve independence and gain confidence [1]. Guttmann was the founding father of the Paralympic Games as we know it today. Nowadays, science supports his believes and the role of sport and exercise in physical and mental rehabilitation [2–6].

Exercise is Medicine®

In 2007 the American College of Sports Medicine (ACSM) together with the American Medical Association launched the worldwide Exercise is Medicine® initiative with the purpose to make physical activity assessment and exercise prescription a standard part of the disease prevention and treatment paradigm for all patients [7–10]. Physical inactivity is defined as a global public health problem, with 3.2 million deaths each year attributable to physical inactivity [11]. While physical inactivity is a widespread problem in the able-bodied population, individuals with a disability often are even less physically active, with the lowest activity levels among individuals with an SCI [12]. It is shown that the majority of the SCI population does not meet the ACSM guidelines of at least 30 minutes of moderate-intensity physical activity per day and that they spend only 3.4% of time per day on dynamic activities, compared with 9.9% for able-bodied controls [12]. In addition to an inactive lifestyle, the resting metabolic rate is 14-27% lower in individuals with SCI compared with able-bodied individuals due to the reduced fat-free mass and altered sympathetic nervous system activity [13]. As a result, their total daily energy expenditure is reduced, which corresponds with a higher chance of being overweight or obese. Five years after discharge only 25% of individuals with an SCI have a body mass index (BMI) within the recommended level ($BMI < 22$), and 54% is obese ($BMI \geq 25$) [14,15]. It is not surprising that metabolic syndrome is common in wheelchair users with an SCI. Metabolic syndrome is defined as a combination of at least three of the following characteristics: abdominal obesity, high blood pressure, high triglycerides, low high-density lipoprotein cholesterol and high fasting glucose [16]. A previous large Dutch SCI cross-sectional study reported metabolic syndrome in 39% of participants with long-standing SCI ($N=223$) [17]. These physical problems are associated with a high prevalence of cardiometabolic disease, which is the leading cause of mortality in individuals with an SCI [18–21]. Physical activity and exercise are associated with a decrease in body mass, abdominal obesity, high blood pressure and cholesterol levels [14,21–24]. Therefore, exercise interventions to increase physical activity and physical capacity are recommended in this population [25,26].

Physical capacity of wheelchair users

Physical capacity is the combined outcome of muscle strength, respiratory function and cardiovascular function (figure 1) [27,28]. The gold standard to measure physical capacity is a graded exercise test (GXT) until volitional exhaustion with outcome parameters peak oxygen uptake (VO_2peak, L/min) and peak power output (POpeak, W). POpeak is dependent on the exercise mode and type of exercise test. Therefore, specificity of testing is recommended. In this thesis outcome parameters are (change in) POpeak and VO_2peak, measured as peak (handcycling) capacity. They are referred to as physical capacity or cardiorespiratory fitness, depending on the chapter.

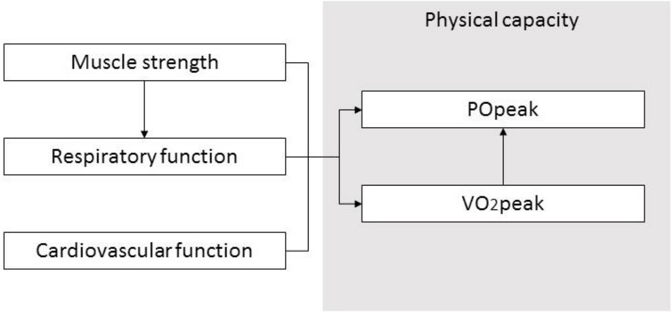


Figure 1. Components of physical capacity. Adapted from the thesis of Janneke Haisma 2008: Physical capacity and complications during and after inpatient rehabilitation for spinal cord injury [29].

In general, wheelchair users have a low physical capacity compared with able-bodied individuals. This is due to a lower muscle mass and lower efficiency of the muscles in the upper body compared with the lower body, and to an inactive lifestyle. In addition, the type and severity of the impairment affect the individual's physical capacity. Most research on physical capacity in wheelchair users is performed among people with SCI. Results showed that in particular individuals with a cervical SCI have a very low physical capacity due to extensive muscle paralysis and loss of sympathetic control under the lesion level [30–32]. A low physical capacity is not only associated with high physical strain during activities of daily living [2,3], but also with a lower chance to return to work [5,6] and a lower life satisfaction [4]. Therefore, interventions to increase physical capacity are important. One of the exercise options for wheelchair users is handcycling.

Handcycling as an exercise mode during and after rehabilitation

Handcycling is a common exercise mode for wheelchair users during and after rehabilitation in the Netherlands. This has several reasons. The first reason is that handcycling induces a cardiorespiratory strain that is sufficient to improve physical capacity. A previous study showed that during inpatient rehabilitation patients spent 77% of time at an intensity >40% heart rate reserve (HRR) during handcycling, which was the highest percentage of all therapy modalities [33]. Second, handcycling can be a good way to improve physical capacity in patients with SCI already early in rehabilitation and even in the most vulnerable patients with a tetraplegia [34]. Third, handcycling offers the possibility to cover larger distances outdoors for a longer duration and at higher speeds than handrim wheelchair propulsion due to the higher mechanical efficiency and lower energy cost during submaximal exercise [35]. Last, the mean contact force in the shoulder is up to 41% lower during handcycling with a synchronous attach-unit crank system compared with handrim wheelchair propulsion at the same workload [36]. Due to the continuous closed-chain movement, the forces are more evenly distributed over the full cycle during handcycling, in contrast to the high peak forces during the short push phase in wheelchair propulsion [36]. As a result, relative muscle forces of the rotator cuff muscles are lower during handcycling, which might reduce the development of overuse injuries in the shoulder [36].

Handcycling history and configurations

Today handcycling has become a popular sport for manual wheelchair users [37]. The first handcycle already dates, however, from as early as 1655. The German watchmaker Stephan Farfler invented the manumotive carriage, which was a wooden chair on a tricycle construction with an asynchronous crank propulsion mechanism (figure 2). The first handcycles had an asynchronous propulsion mode, more closely resembling bicycling. This is in contrast to the synchronous propulsion mode of handcycles in the present day. It is hypothesized that asynchronous propulsion influences the steering direction much more than synchronous propulsion leading to the need for stabilizing muscle activation which goes at the cost of a higher energy consumption [38–40]. Consequently, gross mechanical efficiency and POpeak are higher during synchronous handcycling [35,38,39,41,42]. In a lab-based setting with the equipment fixed to the wall and where steering is not an issue, differences between propulsion modes are less clear. In this thesis, most chapters are based on data collected in a lab-based setting with a synchronous crank propulsion mode.

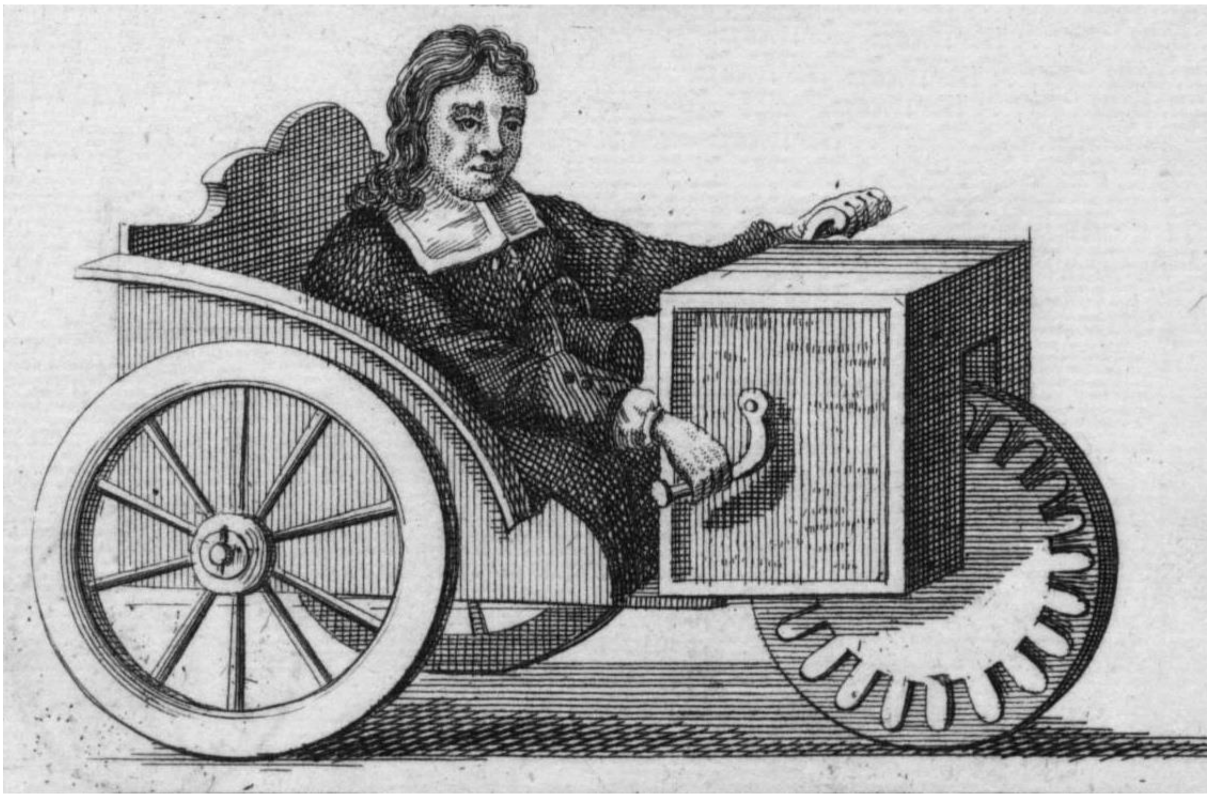


Figure 2. The wooden manumotive carriage with asynchronous arm crank propulsion designed by watchmaker Stephan Farfler (1655) [43].

For handcycling purposes outside, several configurations exist. First, the attach-unit crank systems are commonly used for activities of daily living and recreational purposes. This system consists of an extra (fifth) wheel with a synchronous cranking device that can be attached to the daily handrim wheelchair [44]. Second, the synchronous rigid-frame handcycles that serve sport purposes. Two main sport configurations can be distinguished within the rigid-frame handcycles: the recumbent (arm power (AP)) model and the kneeling (arm trunk power (ATP)) model (figure 3 A and B respectively). Competitive handcycling is conducted following a classification system based on distinctive impairment-based classes [45]. Dependent on the handcycling class (i.e., the impairment-associated functional ability), the individual will ride a recumbent or kneeling handcycle. In handcycling, five sport classes are distinguished: H1 is the class with the highest degree of impairment and H5 with the least. H1 and H2 handcyclists have limitations in arm-hand function, trunk, and pelvis/legs. H3 has no limitations in arm-hand function, but has limitations in trunk and pelvis/legs. H4 has no limitations in arm-hand function and trunk, but has no or very limited lower-limb function, whereas H5 has incomplete loss of lower-limb function. H1 – H4 athletes use a recumbent handcycle, whereas H5 athletes use a kneeling handcycle [45].

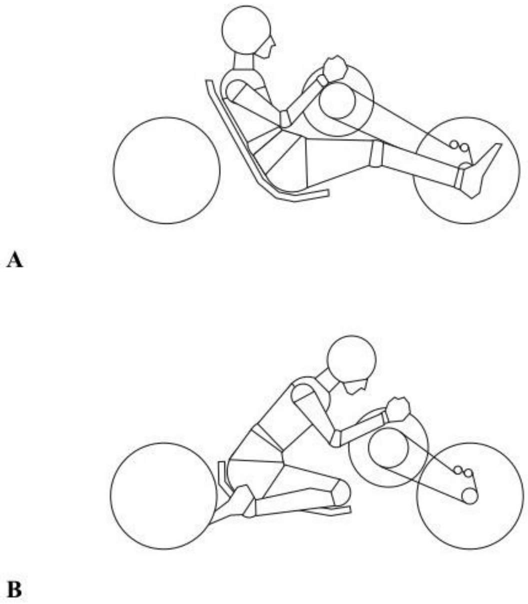


Figure 3. A. recumbent handcycle sports model. B. A kneeling handcycle sports model [46].

Handcycling in a kneeling handcycle results in a higher POpeak due to the recruitment of the trunk muscle mass, the higher range of movement of the trunk and gravitational forces [47–49]. However, handcycling in a kneeling handcycle has a higher physical strain and lower gross mechanical efficiency, and is disadvantageous with respect to air drag [47,49].

HandbikeBattle

The HandbikeBattle event is an annual uphill handcycling mountain time trial in Austria among teams of Dutch rehabilitation centers. The HandbikeBattle project involves the preparatory phase and HandbikeBattle event, which includes, amongst others: a medical screening and GXT, a free-living training period, a week in Austria with peers, and the HandbikeBattle event itself. The HandbikeBattle study is an observational cohort study that aims to monitor participants during and after the HandbikeBattle project up to one year after participation.

The origin of the HandbikeBattle event

Handcycling is a popular exercise mode for wheelchair users during and after rehabilitation in the Netherlands. This is partly due to the above-mentioned reasons, and partly to the infrastructure (flat) and strong cycling culture as part of the daily commuting and recreation in the Netherlands. Another phenomenon that contributed to the formation of the HandbikeBattle event, is the widespread and popular Dutch tradition of raising awareness and money for charity by cycling up a steep mountain road in the Alps. An example of a very popular annual event (14th edition in 2019) is the “Alpe d’HuZes” [50]. Each year 5000 participants cycle up the Alpe d’Huez (six times in a row) to raise money for cancer research in the Netherlands. Two Dutch rehabilitation centers (Heliomare 2010 – 2012 and Rijndam 2011 – 2012) participated in the Alpe d’HuZes with a handcycle team consisting of former rehabilitation patients. Around the same time, a handcycle team from another rehabilitation center (De Hoogstraat) climbed the Col de la Colombière. In 2011 the idea arose to organize a handcycling event in the Alps among teams from Dutch rehabilitation centers. Rogier Broeksteeg (Rijndam), Abel ten Hoorn (De Hoogstraat), Linda Valent and Mechteld Hagoort (Heliomare) were the original founders of the event. The mission of the HandbikeBattle is threefold: 1) to encourage wheelchair users to initiate or keep training after the rehabilitation period, 2) to learn from others and gain confidence to achieve other goals in life, 3) to show that not only elite able-bodied athletes are capable of incredible performances, but recreationally active wheelchair users as well [51].

The Kaunertaler Gletscherstraße in Austria is the stage setting for the event, with two large wheelchair-user friendly accommodations at the foot of the climb. The Kaunertaler Gletscherstraße is a steep climb (figure 4). The start is at 1287 m altitude and the finish is at 2150 m altitude. The total length is 20.2 km. In the middle of the track is a reservoir with approximately 5 km of semi-flat road. The remainder of the track is uphill with percentages around 10-12%. It is a serious climb, which is comparable to Mont Ventoux or Alpe d’Huez.

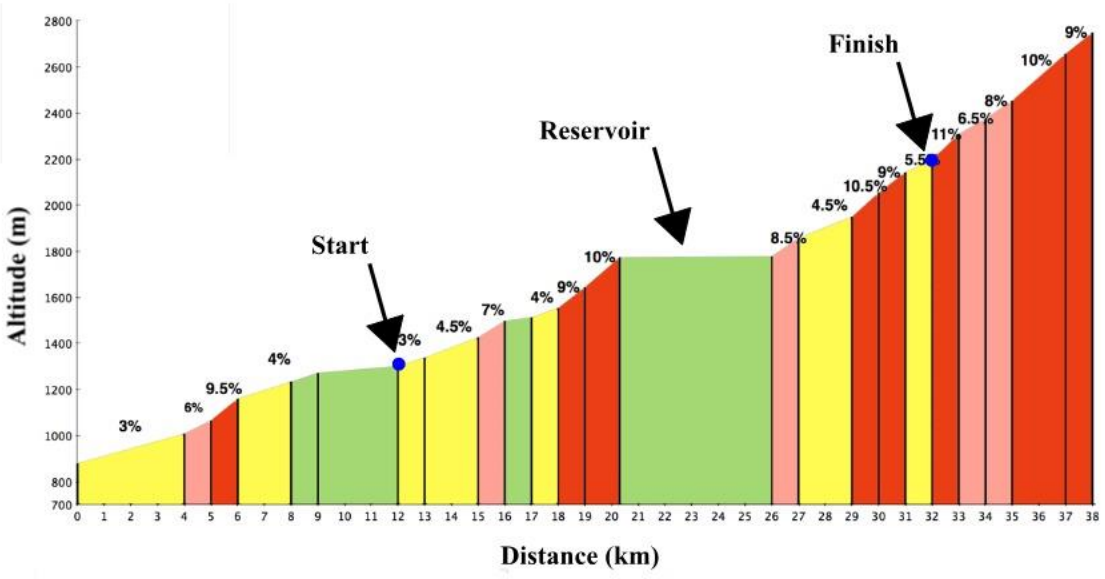


Figure 4. A global impression of the height profile of the Kaunertaler Gletscherstraße. The start of the HandbikeBattle race is at 1287 m altitude and the finish is at 2150 m altitude.

In June 2013 the first edition of the HandbikeBattle was a fact. All eight Dutch rehabilitation centers with a specialized SCI unit participated, with in total 49 participants. The HandbikeBattle became an annual event in June. In the last edition (2019) the event had grown to a total of 112 participants, of which 66 participants were part of a team of one of twelve participating Dutch rehabilitation centers, and 46 participants were individual participants. Individual participants are persons that participated in previous editions of the event and are determined to experience it once more, or (non-elite) international handcyclists. In previous editions, handcyclists from Norway, Switzerland, Belgium and the United States of America participated individually.

Apart from all the volunteers in the organization, it is the large base of support from all rehabilitation professionals from the Dutch SCI rehabilitation network, later supplemented with four other rehabilitation centers, that made the event a consistent success*.

*****The Netherlands has a rich history in SCI rehabilitation research. The Dutch SCI rehabilitation clinical research network (SCIONN) is a multicenter and multi-disciplinary collaboration among all eight Dutch rehabilitation centers with a specialized SCI unit, academic research groups, the Dutch Flemish Spinal Cord Society (DUFSCoS) and the Dutch SCI patient organization (Dwarslaesie Organisatie Nederland (DON)). The research collaboration started in 1999 with the research program "Restoration of mobility in spinal cord injury rehabilitation" (among which the multi-center Umbrella Project)[52] funded by ZonMW. The Umbrella project was a large prospective cohort study in which patients with SCI were followed during inpatient rehabilitation up to five years after discharge. In 2010 a second multi-center research program started called "Active LifestyLe Interventions in aging Spinal Cord injury" (ALLRISC) [53] funded by ZonMW and Fonds NutsOhra. The four projects in this ALLRISC program focused on follow-up care and long-term effects of SCI on active lifestyle, fitness, health, participation and quality of life. This strong clinical research network among the eight specialized SCI rehabilitation centers in the Netherlands (which is still present today) was the foundation for the start of the HandbikeBattle study in 2013. This also explains the large percentage of HandbikeBattle participants with an SCI.

Participating handcycle teams from Dutch rehabilitation centers in 2019

Adelante	Beatrixoord	Heliomare
De Hoogstraat	Libra Revalidatie en Audiologie	Reade
Revant	Rijndam	Roessingh
Sint Maartenskliniek	Tolbrug	Vogellanden

The HandbikeBattle project

Although the HandbikeBattle has a competition element, this is not its main focus. Initially, the HandbikeBattle event was organized to give wheelchair users a mutual goal to start handcycle training (with peers). In this way, the founders hoped to encourage wheelchair users to start or continue training after the rehabilitation period and initiate an active lifestyle. The therapists from the rehabilitation centers ask potential participants during inpatient or outpatient rehabilitation whether they would like to participate in the event in the next year, or former patients contact their rehabilitation center themselves, asking about possibilities to join. Reasons for selection of potential participants are diverse. Some participants were active in endurance sports before their injury and could use some guidance during training, whereas most participants have a very limited sporting background and are new in the handcycling sport. As a result, most participants are relatively untrained handcyclists at the start of the training period. Next to training sessions with peers, participants are supposed (to learn) to take initiative and train independently at home or with family and friends. They should make arrangements with the municipality to obtain funding to buy or borrow their own rigid-frame handcycle, and find sponsors to finance their travel costs to and stay in Austria. The therapists from the rehabilitation centers facilitate this process by giving team members assistance and organizing weekly or monthly (training) meetings. In this way, participants meet peers and learn a lot from the whole project (and from their peers). In addition, during training weekends and the week in Austria they encounter new situations during activities of daily living and learn from peers how to overcome these barriers. They do not only push physical and mental boundaries during the race, but the idea is that during the whole project they improve their social network, learn from peers, see what perseverance and determination gives them, and find confidence to actively participate and pursue even higher goals in life. The main goal of the HandbikeBattle is, therefore, not to win, but to participate in the whole project and eventually reach the finish line at the day of the race. Participants may need up to seven hours to complete the climb (figure 5). During the project and event elite handcyclists Jetze Plat and Mischa Hielkema serve as role models, giving advice and providing training clinics.

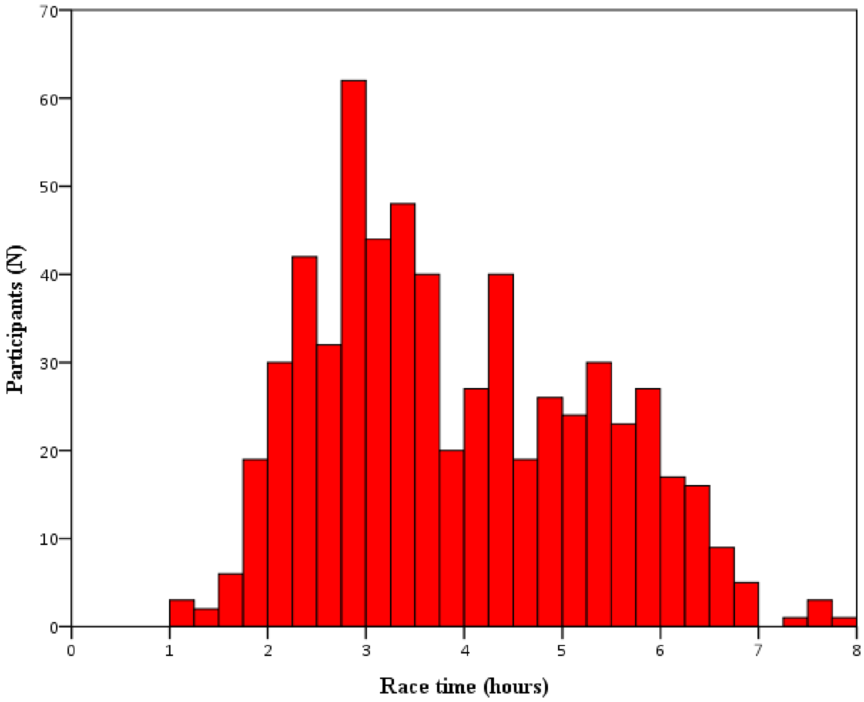


Figure 5. Histogram with race times of all participants of the HandbikeBattle event 2013 – 2019 (N=616). The fastest handcyclists finish in just over the hour (e.g., elite handcyclist Jetze Plat). Mean race time is: 3:56 ± 1:24 (hh:mm).

The start of the HandbikeBattle study

Given the large (research) network that formed the foundation of the HandbikeBattle event, it is not surprising that an observational cohort study was started to investigate the effects of participation in the HandbikeBattle with the first edition in 2013. Sonja de Groot (Reade) and Linda Valent (Heliomare) initiated the study. Teams from all rehabilitation centers participated in the study. The study became a large collaboration among (rehabilitation / sports) physicians, researchers and therapists of all centers, together with the DON, the Dutch cycling federation (KNWU) and handcycle and wheelchair company Double Performance®. In 2016 research funding became available from Revalidatiefonds (currently: HandicapNL), that together with funding from Stichting Mitialto, Heliomare and University Medical Center Groningen of the University of Groningen, resulted in a PhD trajectory of four years with this thesis as final outcome.

Overarching aim of the study

The overarching aim of the study was to carefully monitor all participants physically and mentally. The results of this monitoring process could be used to: 1) inform the rehabilitation centers about the results and use these results to optimize training prescription; 2) gain knowledge about the effects of handcycle training on general, physical and mental health, and understand the associations among these health outcomes; 3) implement this knowledge into rehabilitation practice.

Participants

For each annual cohort, all HandbikeBattle *team* participants from the rehabilitation centers were asked to join the study. All participants voluntarily signed an informed consent form. The study was approved by the Local Ethics Committee of the Center for Human Movement Sciences, University Medical Center Groningen, University of Groningen, the Netherlands (ECB/2012_12.04_l_rev/Ml). Given the fact that most *individual* participants already had participated in a previous year and due to logistics and possible extra costs for these participants, the *individual* HandbikeBattle participants were *not* asked to join the study. This resulted in 48 – 72 new study participants with different impairment types each year (figure 6).

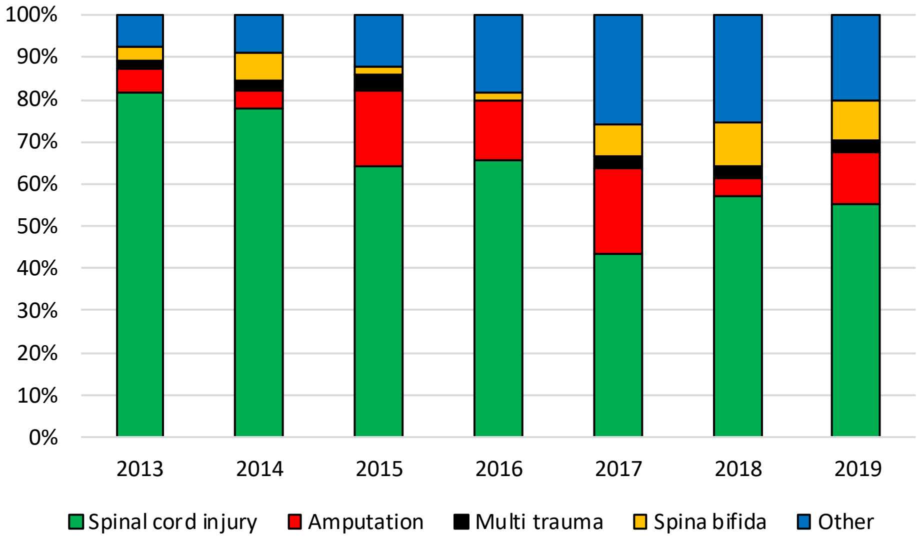
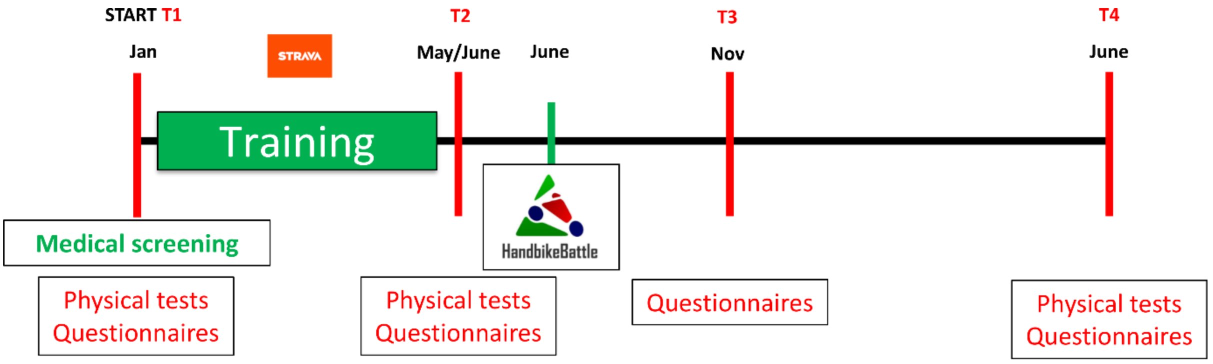


Figure 6. Distribution of impairment types among HandbikeBattle study participants in each year. “Other” includes, amongst others: cerebral palsy, neuromuscular disease, stroke and chronic pain syndrome.

The study set-up

The study had a prospective cohort design, starting from the start of the training period (T1), up to one year after participating in the HandbikeBattle event (T4) (figure 7).



HBB **event**: the mountain time trial in Austria
HBB **project**: medical screening, free-living training (with peers, guided by rehabilitation center), week in Austria among peers, participation in the HBB event
HBB **study**: 4 test occasions with physical tests and questionnaires, monitoring training sessions with Strava. Aim: to study the effects of participating in the HandbikeBattle project for new participants

Figure 7. The set-up of the HandbikeBattle (HBB) study for each cohort. The study had four main testing moments: T1 to T4. In 2020 the HandbikeBattle event did not take place due to COVID-19. Measurements and study data of the HandbikeBattle study cover years 2013-2019.

Testing moment 1 (T1): start of the training period

Medical screening. Each participant (team and individual) was submitted to a medical screening and performed a recent valid GXT to allow participation in the event. If there appeared to be a medical contra-indication to exercise (apparent from the screening or GXT), the participant was excluded from the HandbikeBattle project (and study) by the rehabilitation center. The medical screening was performed by a rehabilitation physician or sports physician at the rehabilitation center. The screening comprised a medical anamnesis and physical examination, including blood pressure, body mass, waist circumference measurements and ECG. Team participants were included by their respective rehabilitation center around January.

Graded exercise test. Respiratory function was measured and a handcycling / arm crank GXT until volitional exhaustion was performed at the rehabilitation center. Depending on the rehabilitation center, the GXT was performed with the use of an arm ergometer (Lode Angio, Groningen, the Netherlands), or a sport handcycle attached to the Tacx roller (Tacx, Terneuzen, the Netherlands) or Cyclus 2 ergometer (RBM eletronik-automation GmbH, Leipzig, Germany). All tests were performed in synchronous mode of cranking. A guideline and instructions were provided to the test assistants of all centers to make the tests as uniform as possible. Either a 1-min stepwise protocol or continuous ramp protocol was used, depending on the preference and practice of the test assistant in the different rehabilitation centers. The testing mode and protocol were consistent within participants over time.

Multilevel regression analyses were performed during all statistical procedures to make corrections for possible differences among rehabilitation centers.

Data from the team participants were collected for the study. In addition, for study purposes, team participants were asked to take part in the following measurements:

Questionnaires. A set of questionnaires was sent to the participants around the same time as the GXT took place at T1. The questionnaires took participants 30 – 45 minutes to complete in total. Participants were invited by email with a link and filled out the questionnaires in their own time at home. The set of questionnaires comprised: questions on sporting history and training aims, Musculoskeletal pain [54], Exercise self-efficacy [55,56], Life satisfaction [57,58], Mental health [59,60], Illness cognition questionnaire [61], Purpose in life test [62], Stage of change [63], the adjusted Physical Activity Scale for Individuals with Physical Disabilities (PASIPD) [64], Disability management [65,66], Body satisfaction [67,68], and questions on activities of daily living.

Training period

Guidance during the training period was provided by therapists from the respective rehabilitation centers, but otherwise the training period was free-living, i.e., no specific training program was provided by the researchers. After the GXT at T1, participants started to train indoors and outdoors. The training was done individually or together with HandbikeBattle participants from the same rehabilitation center. *Training diary*: All participants were asked to monitor all their sporting activity with an online app (Strava) or a training diary on paper. Participants were asked for each training session to write down: the type of training (e.g. handcycling or strength and conditioning), the duration of the training session in minutes, and the intensity of the training with Rating of Perceived Exertion (RPE) on a scale from 0 – 10 [69]. In addition, during the training period, each month participants received a short questionnaire with questions about musculoskeletal pain and overuse injuries.

Testing moment 2 (T2): after the training period, just before the event

In May / June blood pressure, body mass, waist circumference and respiratory function were measured again. In addition, a GXT with the same equipment and protocol as T1 was performed at the rehabilitation center of the participant. Around the same time, the same set of questionnaires was sent to the participants by email, supplemented with questions on evaluation of the training period.

Testing moment 3 (T3): four months after the event

In October / November, a (short term) follow-up measurement took place. The same set of questionnaires was sent to the participants by email, supplemented with questions on losses or benefits of participating in the project.

Testing moment 4 (T4): one year after the event

In June, one year after the event, a (long term) follow-up measurement took place. The same set of questionnaires was sent to the participants by email. In addition, rehabilitation centers were asked to perform a final GXT with the same equipment and protocol. This was only possible in a subset of the rehabilitation centers due to logistics (T4 GXT was in the same time period as T2 GXT of the next cohort) and financial reasons (there was no extra financial compensation for these GXTs available).

Aims and outline of this thesis

The main aims of the studies in this thesis were to investigate the effects of participation in the HandbikeBattle project and event on physical capacity and quality of life, and to answer questions from clinical rehabilitation practice regarding physical capacity testing and handcycle training. In order to study the main aims, this thesis follows three themes: 1) Physical capacity testing, 2) Handcycle training, 3) Effects of participation.

Physical capacity testing

The outcome of the GXT at T1 forms the basis for individualized training. The preferred GXT duration is 8 – 12 minutes [70] in which the anticipated POpeak defines the step size or ramp slope of the protocol. However, it is hard to estimate each individual's POpeak prior to testing, due to the highly diverse individuals participating in the HandbikeBattle. As such, many determinants could play a role, for example: sex, age and lesion level [31,32,71]. It is, however, essential to select the right individualized protocol as the protocol itself affects actual peak performance [70,72,73]. When the step size or ramp slope is too small or too large and, consecutively, test time is too long or too short, it will be unclear whether the "true" peak physical capacity is reached [70,72,74,75]. Moreover, training guidelines based on these peak values will be non-optimal [70,72,74,75]. Therefore, a predictive model for POpeak based on participant characteristics could be helpful in the selection of the right individualized protocol. The specific aims of **chapter 2** were, therefore: 1) To develop and validate predictive models for peak power output (POpeak (W and W/kg)) in a synchronous handcycling GXT for individuals with SCI; 2) To define reference values for absolute and relative POpeak and peak oxygen uptake (VO_2peak) in handcycling based on lesion level and sex.

After the GXT has been performed, individualized training schemes can be based on the results of the GXT. Training prescription with intensity zones based on ventilatory thresholds (VT) is very common in elite able-bodied training literature. The first ventilatory threshold (VT1) is a physiological point during exercise at which a nonlinear increase in carbon dioxide (CO_2) production occurs, coinciding with the first increase in lactate production [76]. The second ventilatory threshold (VT2) represents the onset of exercise induced hyperventilation with respect to VCO_2 as a reaction to metabolic acidosis, which coincides with the maximal lactate steady state [76,77]. These VTs provide boundaries that allow to set individualized training zones: zone 1 at low intensity (below VT1), zone 2 at moderate intensity (between VT1 and VT2), and zone 3 at high intensity (above VT2) [77,78]. A subset of the rehabilitation centers uses these VTs to determine the intensity zones for their participants. However, this training principle has been developed in studies on lower-body exercise with able-bodied participants and athletes, and little or no research has been done regarding the reliability of VT determination in upper-body GXT in recreationally active individuals with SCI. The specific aims of **chapter 3** were therefore: 1) To examine whether it is possible to detect both VTs in recreationally active individuals with tetraplegia or paraplegia; 2) To examine the interrater and intrarater reliability of VT determination.

In **Chapter 4**, the influence of protocol design on the attainment of both VTs and peak physical capacity was investigated. The three commonly used protocols within the HandbikeBattle study were compared among a group of able-bodied participants: ramp protocol, versus 1-min stepwise protocol, versus 3-min stepwise protocol. The specific aim of **chapter 4** was: To examine the effects of stage duration with a ramp protocol, 1-min stepwise protocol, and 3-min stepwise protocol on PO, VO_2, and HR at both peak level and at VT1 and VT2 during synchronous arm crank ergometry.

Handcycle training

A previous HandbikeBattle study showed that on group level there was an increase in POpeak of 17% and in VO_2peak of 7% between T1 and T2 [79]. It is, however, unknown which training regimes led to these improvements. To gain knowledge about training regimes, monitoring training load is essential. In essence, training load can be divided in two categories: external training load and internal training load [80]. In (hand)cycling, external training load is often represented by the training stress score (TSS) based on PO [81]. Internal training load measures are, for example, the training impulse (TRIMP) based on heart rate reserve (HRR) [82] or the session rating of perceived exertion (sRPE) [69]. The TSS is an objective measure of training load, but it can only be determined in handcycle training sessions, and not during, for example, strength and conditioning. Moreover, power meters should be installed in the hub or crank, which is expensive for such a large group of handcyclists. $TRIMP_{HR}$ is less expensive, but not feasible for participants with a tetraplegia as training intensity based on HR is not applicable due to the altered sympathetic response to exercise [83]. $TRIMP_{sRPE}$ would, therefore, be the preferred method to monitor training load in a large cohort of HandbikeBattle participants. In a previous HandbikeBattle study, training load monitoring based on TSS, $TRIMP_{HR}$, and $TRIMP_{sRPE}$, were compared in a subgroup during handcycle training [84]. Partial correlation coefficients were very large (r = 0.77 - 0.81) between $TRIMP_{sRPE}$ and $TRIMP_{HR}$, and $TRIMP_{sRPE}$ and TSS. The dose-response relationship between $TRIMP_{sRPE}$ and change in physical capacity is, however, unclear. Therefore, the specific aims of **chapter**

5 were: 1) To analyze training characteristics of the HandbikeBattle participants; 2) To examine the associations between training load and the change in physical capacity.

Effects of participation

Previous cross-sectional studies showed that physical activity and participating in exercise are associated with quality of life [85–92]. Previous studies also showed that, for example in people with SCI, life satisfaction is reduced and mental health problems are more common compared with the general population [93–95]. Most of these studies, however, reported cross-sectional associations, and longitudinal studies are, unfortunately, scarce. A previous longitudinal study showed that an increase in physical capacity was associated with an increase in life satisfaction in individuals with SCI [4]. It was hypothesized that participating in the HandbikeBattle project would result in an increase in life satisfaction and mental health, and that the increase in physical capacity would be longitudinally associated with this increase in life satisfaction and mental health. The specific aims of **chapter 6** were, therefore: 1) To examine changes in life satisfaction and mental health during five months of training prior to the HandbikeBattle and at four months of follow-up; 2) To examine the associations among changes in handcycling physical capacity and changes in life satisfaction and mental health during the training period.

With respect to physical capacity, several studies showed positive effects of exercise on upper body physical capacity in wheelchair users [79,85,96]. Exercise maintenance on the long term is, however, a challenge. Long-term follow-up studies among wheelchair users are scarce, which is unfortunate as long-term follow-up data are essential to gain knowledge on determinants of maintenance and relapse in physical activity behavior. It was hypothesized that participants who completed the HandbikeBattle would maintain an active lifestyle because the training was not lab-based but self-organized in their own environment, and because they experience less barriers as they overcame certain barriers during the training period. The maintenance of this active lifestyle would result in stable levels of physical capacity at long-term follow-up. Therefore, the specific aims of **chapter 7** were: 1) To compare physical capacity one year after the HandbikeBattle event with physical capacity before and after the training period; 2) To identify determinants that influence the course of physical capacity during follow-up.

In **chapter 8**, the main findings and conclusions of each chapter are summarized and discussed.

References

1. Gold JR, Gold MM. Access for all: The rise of the Paralympic Games. *J R Soc Promot Health*. 2007;127(3):133-141.

2. Janssen TWJ, van Oers CAJM, Veeger HEJ, Hollander AP, van der Woude LHV, Rozendal RH. Relationship between physical strain during standardised adl tasks and physical capacity in men with spinal cord injuries. *Paraplegia*. 1994;32(12):844-859.

3. Dallmeijer AJ, Hopman MT, van As HH, van der Woude LHV. Physical capacity and physical strain in persons with tetraplegia; the role of sport activity. *Spinal Cord*. 1996;34(11):1173-1176.

4. Van Koppenhagen CF, Post M, de Groot S, 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(3):328-337.

5. Van Velzen JM, de Groot S, Post MWM, Slootman JR, van Bennekom CAM, van der Woude LHV. Return to work after spinal cord injury. *Am J Phys Med Rehabil*. 2009;88(1):47-56.

6. Noreau L, Shephard RJ. Return to work after spinal cord injury: The potential contribution of physical fitness. *Paraplegia*. 1992;30(8):563-572.

7. Sallis RE. Exercise is medicine and physicians need to prescribe it! *Br J Sports Med*. 2009;43(1):3-4.

8. Sallis R. Exercise is medicine: A call to action for physicians to assess and prescribe exercise. *Phys Sportsmed*. 2014;43(1):22-26.

9. Exercise is Medicine. www.exerciseismedicine.org. Accessed August 14, 2020.

10. Cowan RE. Exercise Is Medicine Initiative: Physical activity as a vital sign and prescription in adult rehabilitation practice. *Arch Phys Med Rehabil*. 2016;97(9):S232-S237.

11. World Health Organizanization: Global Strategy on Diet, Physical Activity and Health. www.who.int/dietphysicalactivity/factsheet_inactivity/en. Accessed August 14, 2020.

12. Van den Berg-Emons RJ, Bussmann JBJ, Haisma JA, et al. A Prospective study on physical activity levels after spinal cord injury during inpatient rehabilitation and the year after discharge. *Arch Phys Med Rehabil*. 2008;89(11):2094-2101.

13. Buchholz AC, Pencharz PB. Energy expenditure in chronic spinal injury. *Curr Opin Clin Nutr Metab Care*. 2004;7(6):635-639.

14. De Groot S, Post MWM, Snoek GJ, Schuitemaker M, van der Woude LHV. Longitudinal association between lifestyle and coronary heart disease risk factors among individuals with spinal cord injury. *Spinal Cord*. 2013;51(4):314-318.

15. Laughton GE, Buchholz AC, Martin Ginis KA, Goy RE, Group TSSR. Lowering body mass index cutoffs better identifies obese persons with spinal cord injury. *Spinal Cord*. 2009;47(10):757-762.

16. Saklayen MG. The Global Epidemic of the Metabolic Syndrome. *Curr Hypertens Rep*. 2018;20(2):1-8.

17. De Groot S, Adriaansen JJ, Tepper M, Snoek GJ, van der Woude LHV, Post MWM. Metabolic syndrome in people with a long-standing spinal cord injury: Associations with physical activity and capacity. *Appl Physiol Nutr Metab*. 2016;41(11):1190-1196.

18. Garshick E, Kelley A, Cohen SA, et al. A prospective assessment of mortality in chronic spinal cord injury. *Spinal Cord*. 2005;43(7):408-416.

19. Chopra AS, Miyatani M, Craven BC. Cardiovascular disease risk in individuals with chronic spinal cord injury: Prevalence of untreated risk factors and poor adherence to treatment guidelines. *J Spinal Cord*

Med. 2018;41(1):2-9.

20. Myers J, Lee M, Kiratli J. Cardiovascular disease in spinal cord injury: An overview of prevalence, risk, evaluation, and management. *Am J Phys Med Rehabil*. 2007;86(2):142-152.

21. Nightingale TE, Metcalfe RS, Vollaard NB, Bilzon JL. Exercise guidelines to promote cardiometabolic health in spinal cord injured humans: Time to raise the intensity ? *Arch Phys Med Rehabil*. 2017;98(8):1693-1704.

22. Manns PJ, McCubbin JA, Williams DP. Fitness, inflammation, and the metabolic syndrome in men with paraplegia. *Arch Phys Med Rehabil*. 2005;86(6):1176-1181.

23. Nooijen CFJ, de Groot S, Postma K, et al. A more active lifestyle in persons with a recent spinal cord injury benefits physical fitness and health. *Spinal Cord*. 2012;50(4):320-323.

24. Bakkum AJT, Paulson TAW, Bishop NC, et al. Effects of hybrid cycle and handcycle exercise on cardiovascular disease risk factors in people with spinal cord injury: A randomized controlled trial. *J Rehabil Med*. 2015;47(6):523-530.

25. Van der Scheer JW, Ginis KAM, Ditor DS, et al. Effects of exercise on fitness and health of adults with spinal cord injury: A systematic review. *Neurology*. 2017;89(7):736-745.

26. Martin Ginis KA, van der Scheer JW, Latimer-Cheung AE, et al. Evidence-based scientific exercise guidelines for adults with spinal cord injury: An update and a new guideline. *Spinal Cord*. 2018;56(4):308-321.

27. Stewart MW, Melton-Rogers SL, Morrison S, Figoni SF. The measurement properties of fitness measures and health status for persons with spinal cord injuries. *Arch Phys Med Rehabil*. 2000;81(4):394-400.

28. Van der Woude L, de Groot S, van Drongelen S, et al. Evaluation of manual wheelchair performance in everyday life. *Top Spinal Cord Inj Rehabil*. 2009;15(2):1-15.

29. Haisma JA. Physical capacity and complications during and after inpatient rehabilitation for spinal cord injury. Published online 2008. repub.eur.nl/publications/dissertations/mh_diss/index/947169971/%5CnC:%5CEMH%5CScannede artikler referanser%5CRefMan2301.pdf

30. 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(11):642-652.

31. Janssen TWJ, Dallmeijer AJ, Veeger HEJ, van der Woude LHV. Normative values and determinants of physical capacity in individuals with spinal cord injury. *J Rehabil Res Dev*. 2002;39(1):29-39.

32. Simmons OL, Kressler J, Nash MS. Reference fitness values in the untrained spinal cord injury population. *Arch Phys Med Rehabil*. 2014;95(12):2272-2278.

33. Koopman ADM, Eken MM, van Bezeij T, Valent LJM, Houdijk H. Does clinical rehabilitation impose sufficient cardiorespiratory strain to improve aerobic fitness? *J Rehabil Med*. 2013;45(1):92-98.

34. Valent LJM, Dallmeijer AJ, Houdijk H, et al. Effects of hand cycle training on physical capacity in individuals with tetraplegia: a clinical trial. *Phys Ther*. 2009;89(10):1051-1060.

35. Dallmeijer AJ, Ottjes L, de Waardt E, van der Woude LHV. A physiological comparison of synchronous and asynchronous hand cycling. *Int J Sports Med*. 2004;25(8):622-626.

36. Arnet U, van Drongelen S, Scheel-Sailer A, van der Woude LHV, Veeger HEJ. Shoulder load during synchronous handcycling and handrim wheelchair propulsion in persons with paraplegia. *J Rehabil Med*. 2012;44(3):222-228.

37. Abel T, Vanlandewijck Y, Verellen J. Handcycling. In: Goosey-Tolfrey V, editor. *Wheelchair Sport*. Human Kinetics; 2010:187-197.

38. Van der Woude LHV, Bosmans I, Bervoets B, Veeger HEJ. Handcycling: different modes and gear ratios. *J Med Eng Technol*. 2000;24(6):242-249.

39. Van der Woude LHV, Horstman A, Faas P, Mechielsen S, Bafghi HA, de Koning JJ. Power output and metabolic cost of synchronous and asynchronous submaximal and peak level hand cycling on a motor driven treadmill in able-bodied male subjects. *Med Eng Phys*. 2008;30(5):574-580.

40. Hettinga FJ, Valent L, Groen W, van Drongelen S, de Groot S, van der Woude LHV. Hand-cycling: An active form of wheeled mobility, recreation, and sports. *Phys Med Rehabil Clin N Am*. 2010;21(1):127-140.

41. Kraaijenbrink C, Vegter RJK, Hensen AHR, Wagner H, van der Woude LHV. Biomechanical and physiological differences between synchronous and asynchronous low intensity handcycling during practice-based learning in able-bodied men. *J Neuroeng Rehabil*. 2020;17(1):1-13.

42. Bafghi HA, de Haan A, Horstman A, van der Woude LHV. Biophysical aspects of submaximal hand cycling. *Int J Sports Med*. 2008;29(8):630-638.

43. Wikipedia. https://en.wikipedia.org/wiki/Stephan_Farffler. Accessed November 25, 2019.

44. Van der Woude LHV, Dallmeijer AJ, Janssen TWJ, Veeger HEJ. Alternative modes of manual wheelchair ambulation: an overview. *Am J Phys Med Rehabil*. 2001;80(10):765-777.

45. Union Cycliste Internationale Cycling Regulations, part 16 Para-cycling. Available at: https://www.uci.org/docs/default-source/rules-and-regulations/16-par-20200211-e.pdf. Accessed November 25, 2019.

46. Van Breukelen K. *Rolstoel Performance*. Double Performance; 2014.

47. Verellen J, Meyer C, Janssens L, Vanlandewijck Y, Lacour JR. Peak and submaximal steady-state metabolic and cardiorespiratory responses during arm-powered and arm-trunk-powered handbike ergometry in able-bodied participants. *Eur J Appl Physiol*. 2012;112(3):983-989.

48. Faupin A, Gorce P, Meyer C, Thevenon A. Effects of backrest positioning and gear ratio on nondisabled subjects' handcycling sprinting performance and kinematics. *J Rehabil Res Dev*. 2008;45(1):109-116.

49. Kouwijzer I, Nooijen CFJ, van Breukelen K, Janssen TWJ, De Groot S. Effects of push-off ability and handcycle type on handcycling performance in able-bodied participants. *J Rehabil Med*. 2018;50(6):563-568.

50. Alpe d'HuZes: a life changing experience. www.opgevenisgeenoptie.nl. Accessed August 14, 2020.

51. HandbikeBattle. www.handbikebattle.nl. Accessed August 14, 2020.

52. De Groot S, Dallmeijer AJ, Post MWM, et al. Demographics of the Dutch multicenter prospective cohort study "Restoration of mobility in spinal cord injury rehabilitation." *Spinal Cord*. 2006;44(11):668-675.

53. Van der Woude LHV, de Groot S, Postema K, et al. Active LifestyLe Rehabilitation Interventions in aging Spinal Cord injury (ALLRISC): a multicentre research program. *Disabil Rehabil*. 2013;35(13):1097-1103.

54. Eriks-Hoogland IE, Hoekstra T, de Groot S, Stucki G, Post MWM, van der Woude LHV. Trajectories of musculoskeletal shoulder pain after spinal cord injury: Identification and predictors. *J Spinal Cord Med*. 2014;37(3):288-298.

55. Kroll T, Kehn M, Ho P, Groah S. The SCI exercise self-efficacy scale (ESES): development and psychometric properties. *Int J Behav Nutr Phys Act*. 2007;4(34):2-7.

56. Nooijen CFJ, Post MWM, Spijkerman DCM, Bergen MP, Stam HJ, van den Berg-emons RJG. Exercise self-

efficacy in persons with spinal cord injury: psychometric properties of the Dutch translation of the exercise self-efficacy scale. *J Rehabil Med*. 2013;45(4):347-350.

57. Van Koppenhagen CF, Post MWM, van der Woude LHV, et al. Recovery of life satisfaction in persons with spinal cord injury during inpatient rehabilitation. *Am J Phys Med Rehabil*. 2009;88(11):887-895.

58. Post MWM, van Leeuwen CMC, van Koppenhagen CF, de Groot S. Validity of the Life Satisfaction Questions , the Life Satisfaction Questionnaire , and the Satisfaction With Life Scale in persons with spinal cord injury. *YAPMR*. 2012;93(10):1832-1837.

59. Ware JE, Snow KK, Kosinski M, Gandek B. *SF-36 Health Survey Manual and Interpretation Guide*. The Health Institute, New England Medical Center; 1993.

60. Van Leeuwen CMC, van der Woude LHV, Post MWM. Validity of the mental health subscale of the SF-36 in persons with spinal cord injury. *Spinal Cord*. 2012;50(9):707-710.

61. Evers AWM, Kraaimaat FW, van Lankveld W, Jongen PJH, Jacobs JWG, Bijlsma JWJ. Beyond unfavorable thinking: The Illness Cognition Questionnaire for chronic diseases. *J Consult Clin Psychol*. 2001;69(6):1026-1036.

62. Crumbaugh JC. Cross-validation of Purpose in Life test based on Frankl's concepts. *J Individ Psychol*. 1968;24(1):74-81.

63. Kosma M, Ellis R, Cardinal BJ, Bauer JJ, McCubbin JA. The mediating role of intention and stages of change in physical activity among adults with physical disabilities: an integrative framework. *J Sport Exerc Psychol*. 2007;29(1):21-38.

64. Washburn RA, Zhu W, McAuley E, Frogley M, Figoni SF. The physical activity scale for individuals with physical disabilities: Development and evaluation. *Arch Phys Med Rehabil*. 2002;83(2):193-200.

65. Amtmann D, Bamer AM, Cook KF, Askew RL, Noonan VK, Brockway JA. University of Washington Self-Efficacy Scale: A new self-efficacy scale for people with disabilities. *Arch Phys Med Rehabil*. 2012;93(10):1757-1765.

66. Cijsouw A, Adriaansen JJE, Tepper M, et al. Associations between disability-management self-efficacy, participation and life satisfaction in people with long-standing spinal cord injury. *Spinal Cord*. 2017;55(1):47-51.

67. Bassett RL, Ginis KAM, Buchholz AC, Group the SSR. A pilot study examining correlates of body image among women living with SCI. *Spinal Cord*. 2009;47(6):496-498.

68. Reboussin BA, Rejeski WJ, Martin KA, et al. Correlates of satisfaction with body function and body appearance in middle- and older aged adults : The activity counseling trial (ACT). *Psychol Health*. 2000;15:239-254.

69. Foster C, Florhaug JA, Franklin J, et al. A New approach to monitoring exercise training. *J Strength Cond Res*. 2001;15(1):109-115.

70. Buchfuhrer MJ, Hansen JE, Robinson TE, Sue DY, Wasserman K, Whipp BJ. Optimizing the exercise protocol for cardiopulmonary assessment. *J Appl Physiol Respir Environ Exerc Physiol*. 1983;55(5):1558-1564.

71. Muraki S, Tsunawake N, Tahara Y, Hiramatsu S, Yamasaki M. Multivariate analysis of factors influencing physical work capacity in wheelchair-dependent paraplegics with spinal cord injury. *Eur J Appl Physiol*. 2000;81(1-2):28-32.

72. Smith PM, Amaral I, Doherty M, Price MJ, Jones AM. The influence of ramp rate on VO_2peak and "excess" $\dot{V}O_2$ during arm crank ergometry. *Int J Sports Med*. 2006;27(8):610-616.

73. Eerden S, Dekker R, Hettinga 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(5):497-521.

74. Bentley DJ, Newell J, Bishop D. Incremental exercise test design and analysis: Implications for performance diagnostics in endurance athletes. *Sport Med*. 2007;37(7):575-586.

75. Myers J, Bellin D. Ramp exercise protocols for clinical and cardiopulmonary exercise testing. *Sports Med*. 2000;30(1):23-29.

76. Kindermann W, Simon G, Keul J. The significance of the aerobic-anaerobic transition for the determination of work load intensities during endurance training. *Eur J Appl Physiol*. 1979;42(1):25-34.

77. Meyer T, Lucía A, Earnest CP, Kindermann W. A conceptual framework for performance diagnosis and training prescription from submaximal gas exchange parameters - Theory and application. *Int J Sport Med*. 2005;26(Suppl 1):38-48.

78. Seiler KS, Kjerland GØ. Quantifying training intensity distribution in elite endurance athletes: Is there evidence for an “optimal” distribution? *Scand J Med Sci Sport*. 2006;16(1):49-56.

79. Hoekstra SP, Valent LJM, Gobets D, van der Woude LHV, de Groot S. Effects of four-month handbike training under free-living conditions on physical fitness and health in wheelchair users. *Disabil Rehabil*. 2017;39(16):1581-1588.

80. Halson SL. Monitoring training load to understand fatigue in athletes. *Sport Med*. 2014;44(Suppl 2):S139-S147.

81. Hunter A, Coggan A, McGregor S. Beyond average power. In: *Training and Racing with a Power Meter*. third edition. Velopress; 2019:105-138.

82. Banister E, Calvert T. Planning for future performance: implications for long term training. *Can J Appl Sport Sci*. 1980;5(3):170-176.

83. Valent LJM, Dallmeijer AJ, Houdijk H, et al. The individual relationship between heart rate and oxygen uptake in people with a tetraplegia during exercise. *Spinal Cord*. 2007;45(1):104-111.

84. De Groot S, Hoekstra SP, Grandjean Perrenod Comtesse P, Kouwijzer I, Valent LJM. Relationship between internal and external handcycle training load in people with spinal cord injury training for the HandbikeBattle. *J Rehabil Med*. 2018;50(3):261-268.

85. Hicks AL, Martin KA, Ditor DS, et al. Long-term exercise training in persons with spinal cord injury: effects on strength, arm ergometry performance and psychological well-being. *Spinal Cord*. 2003;41(1):34-43.

86. Sweet SN, Ginis KAM, Tomasone JR. Investigating intermediary variables in the physical activity and quality of life relationship in persons with spinal cord injury. *Health Psychology.* 2013;32(8):877-885.

87. Motl RW, McAuley E. Pathways between physical activity and quality of life in adults with multiple sclerosis. *Health Psychology.* 2009;28(6):682-689.

88. Krops LA, Jaarsma EA, Dijkstra PU, Geertzen JHB, Dekker R. Health related quality of life in a Dutch rehabilitation population: Reference values and the effect of physical activity. PLoS One. 2017;12(1):e0169169.

89. Slaman J, van den Berg-Emons HJG, van Meeteren J, et al. A lifestyle intervention improves fatigue, mental health and social support among adolescents and young adults with cerebral palsy: focus on mediating effects. Clin Rehabil. 2015;29(7):717-727.

90. Anneken V, Hirschfeld S, Scheuer T, Thietje R. Influence of physical exercise on quality of life in individuals with spinal cord injury. *Spinal Cord*. 2009;48(5):393-399.

91. Ginis KAM, Jetha A, Mack DE, Hetz S. Physical activity and subjective well-being among people with spinal

cord injury: a meta-analysis. *Spinal Cord*. 2009;48(1):65-72.

92. Bragaru M, Dekker R, Geertzen JHB, Dijkstra PU. Amputees and Sports: a systematic review. *Sports Med*. 2011;41(9):721-740.

93. Post MWM, van Leeuwen CMC. Psychosocial issues in spinal cord injury: a review. *Spinal Cord*. 2012;50(5):382-389.

94. Van Koppenhagen CF, Post MWM, van der Woude LHV, et al. Changes and determinants of life satisfaction after spinal cord injury: A cohort study in the Netherlands. Arch Phys Med Rehabil. 2008;89(9):1733-1740.

95. Dijkers MPJM. Quality of life of individuals with spinal cord injury: A review of conceptualization, measurement, and research findings. 2005;42(3):87-110.

96. Jacobs PL, Nash MS. Exercise recommendations for individuals with spinal cord injury. *Sport Med*. 2004;34(11):727-751.

Chapter 2

Peak power output in handcycling of individuals with a chronic spinal cord injury: predictive modeling, validation and reference values

Kouwijzer I, Valent LJM, Osterthun R, van der Woude LHV, de Groot S, HandbikeBattle Group. Peak power output in handcycling of individuals with a chronic spinal cord injury: predictive modeling, validation and reference values. Disability & Rehabilitation 2020;42(3):400-409.

Abstract

Purpose: To develop and validate predictive models for peak power output to provide guidelines for individualized handcycling graded exercise test protocols for people with spinal cord injury (SCI); and to define reference values.

Materials and methods: Power output was measured in 128 handcyclists with SCI during a synchronous handcycling exercise test. Eighty percent of the data was used to develop four linear regression models: two theoretical and two statistical models with peak power output (in W and W/kg) as dependent variable. The other 20% of the data was used to determine agreement between predicted versus measured power output. Reference values were based on percentiles for the whole group.

Results: Lesion level, handcycling training hours and sex or body mass index were significant determinants of peak power output. Theoretical models (R^2=42%) were superior to statistical models (R^2=39% for power output in W, R^2=30% for power output in W/kg). The intraclass correlation coefficients varied between 0.35 and 0.60, depending on the model. Absolute agreement was low.

Conclusions: Both models and reference values provide insight in physical capacity of people with SCI in handcycling. However, due to the large part of unexplained variance and low absolute agreement, they should be used with caution.

Keywords

Arm ergometry, graded exercise test, physical capacity, normative values, post-rehabilitation

Introduction

Today synchronous handcycling has become a popular sport for wheelchair users [1]. This is not surprising since handcycling is a relatively easy mode to cover large distances at a high speed compared to handrim wheelchair propulsion [1]. Benefits of handcycling include its higher efficiency and lower strain compared to wheelchair propulsion, possibly reducing the risk of upper body overuse injuries [2–4]. Moreover, it has been shown that handcycling can be a good way to improve physical capacity in, for example, individuals with a spinal cord injury (SCI) already early in rehabilitation [5]. This is an important result, as the physical capacity in this population is generally low due to muscle paralysis and loss of sympathetic control under the lesion level, as well as a sedentary lifestyle [6–9]. In previous studies, the benefits of an improvement in physical capacity for wheelchair users with an SCI have already been shown, such as a more favorable lipid profile [10], a higher life satisfaction [11,12] and a higher chance to return to work [13,14].

Above mentioned results are predominantly based on studies that focused on wheelchair capacity, which is different from handcycling, as demonstrated by the lower submaximal strain and higher peak power output (POpeak) during handcycling [3,4]. Next to wheelchair ergometry, asynchronous arm ergometry is studied in individuals with SCI [15–18]. However, several studies highlighted differences in physiological responses between the asynchronous and synchronous propulsion mode [15,19]. For example, a higher net and gross efficiency, and a higher POpeak were found during asynchronous arm cranking compared to synchronous arm cranking [15,19]. Therefore, results of these studies investigating asynchronous arm ergometry cannot be applied to the synchronous handcycling propulsion mode investigated in the present study. This emphasizes the importance of specificity in testing when studying submaximal and peak physiological responses.

In order to stimulate an improvement in physical capacity by means of handcycling in wheelchair users with SCI, the HandbikeBattle is organized as an annual event since 2013. The HandbikeBattle is an uphill handcycling mountain race in Austria in which currently 11 Dutch rehabilitation centers participate with approximately six participants each [20]. All participants are chronic wheelchair users and relatively inexperienced handcyclists who train between four and six months prior to the event. Prior to participation, medical screening including a peak handcycle or synchronous arm crank aerobic exercise test (GXT) is obligatory. The GXT is part of the cardiopulmonary check-up and forms the basis for an individualized training guideline. When using a typical one-minute protocol and preferred GXT duration of 8 – 12 minutes [21], the anticipated POpeak (W) defines step size of the protocol. As many factors play a role in determining the potential physical capacity of these highly diverse individuals with SCI [9], it is hard to estimate each individual's POpeak prior to testing. As such it is difficult to select an optimal GXT protocol. It is, however, essential to select the right individualized protocol for an individual with SCI as the protocol itself affects actual peak performance [21–23]. When the step size or ramp slope is too small or too large and, consecutively, test time is too long or too short, it will be unclear whether the "true" peak physical capacity is reached [21,22,24,25]. Moreover, training guidelines based on these peak values will be non-optimal [21,22,24,25].

To select an optimal individual handcycling GXT protocol for individuals with SCI and, consecutively, improve the development of individualized training guidelines, a POpeak prediction model could be valuable. In such models, POpeak is estimated based on known participant characteristics. Moreover, development of a model could give a theoretical background in the underlying factors influencing physical capacity in individuals with SCI during handcycling and insight in which factors should be influenced to increase physical capacity. In addition to merely statistics-driven modeling, theory-driven statistical models could be useful to further clarify and explain the associations of underlying determinants with physical capacity for this specific mode of exercise.

Based on previous literature investigating wheelchair ergometry or asynchronous arm ergometry in individuals with SCI, several participant characteristics were identified to be of influence on POpeak. Sex, for example, showed to be an important characteristic, as women generally produce a lower POpeak than men [26], which might be explained by the smaller upper-body muscle mass [27]. Moreover, lesion level and completeness are inversely related to POpeak [9,17,18,26,28–30]. POpeak also declines with age [17,26,29,31] and increases with activity level [9,17,29,32,33]. Time since injury (TSI) could be a determinant as physical capacity shows an increase in the first years after SCI [9,34,35] but thereafter seems to decrease [9,36] Janssen et al. (N=166) performed a statistical stepwise (forward) multiple regression analysis for POpeak in wheelchair ergometry and found lesion level, hours of sport, age, body mass, TSI and completeness to be significant determinants (with a cumulative explained variance (R^2) of 80%) [29]. Simmons et al. (N=179) found functional classification, BMI and motor level of injury to be significant determinants for relative POpeak (W/kg) in (asynchronous) arm ergometry (cumulative R^2 of 57%) and motor level of injury, functional classification and sex for absolute POpeak (W) (cumulative R^2 of 57%), performing a forward multiple regression analysis [18]. To date, in synchronous handcycling it is, however, still unknown which factors determine physical capacity. Moreover, previously described models have never been validated. Therefore, the validity of these models for use in clinical practice remains uncertain. Next to the missing knowledge about underlying factors influencing physical capacity in handcycling and uncertainty about the validity of predictive modeling, comparison to group level is lacking, as handcycling reference values for physical capacity for individuals with a SCI are scarce. The aims of this study were, therefore:

1) To develop four predictive models: two theory-driven and two statistically-driven models for POpeak (W and W/kg) in a synchronous handcycling GXT for people with SCI.
2) To validate the four predictive models for POpeak.

3) To define reference values for absolute and relative POpeak and peak oxygen uptake (VO_2peak) in handcycling based on lesion level and sex.

Materials and methods

Participants

Participants were retrospectively selected from the HandbikeBattle 2013, 2014, 2015, 2016 and 2017 cohorts. Every year was a unique cohort. Selection criteria for this study were having an SCI or spina bifida and the availability of comprehensive testing results. A total of 168 participants with SCI or spina bifida were selected. Forty participants were excluded due to missing data in either outcome variables or determinants. This led to 128 recreational handcyclists with SCI or spina bifida being included in this study. Participant characteristics are listed in table 1. The study was approved by the Local Ethical Committee of the Center for Human Movement Sciences, University Medical Center Groningen, the Netherlands (ECB/2012_12.04_l_rev/Ml). All participants voluntarily signed an informed consent form after they were given information about the testing procedures. The study was registered in the Dutch Trial Register (www.trialregister.nl, NTR6586).

Outcomes

In this cross-sectional study, participants underwent a medical screening including a medical history and a physical examination obtained by a physician. Moreover, all participants performed a GXT as part of the medical screening. As the GXT took place before the training period, participants were relatively untrained handcyclists. Depending on the rehabilitation center the pre-training GXT was performed with the use of an arm ergometer (Lode Angio, Groningen, The Netherlands) or a recumbent sport handcycle attached to the Tacx roller (Tacx, Terneuzen, The Netherlands) or Cyclus 2 ergometer (RBM elektronik-automation GmbH, Leipzig, Germany). Comparable peak physiological responses are to be expected between these ergometers (ICC 0.87 Lode vs Tacx, ICC 0.88 Lode vs Cyclus2) [37]. All tests were performed in synchronous mode of cranking. A testing guideline and instructions were provided to the test assistants of all centers to make the tests as uniform as possible. Either a one-minute step protocol or continuous ramp protocol was used, depending on the preference and practice of the test assistant in the different rehabilitation centers. There was no systematic difference in VO_2peak and POpeak to be expected between these protocols [38]. For the one-minute protocol, the test started at 20-30 W with increments of 5-15 W/min.

Table 1. Participant characteristics of the total group (N=128), the model group (80% of data, N=104), and the validation group (20% of data, N=24).

	Total group (N = 128)		**Model group** (N = 104)	**Validation Group** (N = 24)
	M ± SD or N	*N total*	*M ± SD or N*	*M ± SD or N*
SCI/spina bifida	118/10	128	96/8	22/2
Lesion level (>Th6/≤Th6)	37/86	123	32/68	5/18
Completeness (motor compl/incompl)	77/41	118	61/35	16/6
Sex (male/female)	106/22	128	85/19	21/3
Age (years)	39 ± 12	128	39 ± 12	39 ± 12
TSI (years)	10 ± 10	119	10 ± 10	10 ± 9
Height (m)	1.80 ± 0.10	127	1.79 ± 0.10	1.80 ± 0.11
Body Mass (kg)	78 ± 17	127	78 ± 16	79 ± 18
BMI (kg/m^2)	24 ± 4	126	24 ± 4	24 ± 4
Waist circumference (cm)	91 ± 15	116	91 ± 15	88 ± 17
Handcycling training (h)	3.39 ± 3.70	121	3.51 ± 3.84	2.84 ± 2.99
Handcycling classification (H1-H3/H4-H5)	67/57	124	55/46	12/11
POpeak (W)	119 ± 34	128	119 ± 33	121 ± 40
POpeak/kg (W/kg)	1.54 ± 0.47	127	1.54 ± 0.46	1.56 ± 0.51
VO_2peak (L/min)	1.91 ± 0.58	126	1.88 ± 0.56	2.05 ± 0.66
VO_2peak/kg (ml/kg/min)	24.93 ± 7.91	125	24.58 ± 7.60	26.51 ± 9.17
HRpeak (bpm)	171 ± 22	124	171 ± 22	174 ± 23
RERpeak	1.21 ± 0.12	115	1.21 ± 0.12	1.22 ± 0.11
Cyclus 2/Tacx/arm ergometer	35/24/69	128	29/22/53	6/2/16
1 min/ramp	79/49	128	66/38	13/11

SCI: spinal cord injury; TSI: time since injury; BMI: body mass index; POpeak: peak power output; VO_2peak: peak oxygen uptake; HRpeak: peak heart rate; RERpeak: peak respiratory exchange ratio. Lesion level: two categories: (1) above Th6 and (2) equal to or below Th6. Completeness: AIS (two categories: (1) motor complete (AIS A+B) and (2) motor incomplete (AIS C+D)), handcycling training: average handcycling weekly training hours in the last 3 months, handcycling classification: two categories: (1) H1-H3 and (2) H4-H5. Measurement device: cyclus 2, Tacx, or arm ergometer. Protocol type: 1 min step protocol or ramp protocol.

For the ramp protocol, the test started at 0 W with a slope of 1 W / 12 sec (5 W/min), 1 W / 6 sec (10 W/min), 1 W / 4 sec (15 W/min) or 1 W / 3 sec (20 W/min). The selection of the appropriate protocol per individual participant was based on expert opinion of the test assistant. Criteria to stop the test were volitional exhaustion or failure in keeping a constant cadence above the preset value. PO (W) was measured during the test. POpeak was defined as the highest PO attained during this specific synchronous GXT. For the one-minute protocol POpeak (W) was defined as the highest PO that was maintained for at least 30 s. For the ramp protocol the highest PO achieved during the test was considered POpeak. Apart from PO, gas exchange was measured using the Cosmed (Cosmed, Roma, Italy), Cortex (Cortex, CORTEX Biophysik GmbH, Germany) or Oxycon (Erich Jaeger, Viasys Healthcare, Germany). The equipment was calibrated before each test. VO_2peak (l/min) and the peak respiratory exchange ratio (RERpeak) were defined as the highest 30-s average for VO_2 (l/min) and RER, respectively. HRpeak (bpm) was defined as the highest heart rate achieved during the test.

Determinants

During the medical screening, age (years), sex, height (m), TSI (years), lesion level, completeness of the lesion (using the ASIA Impairment Scale (AIS, category A, B, C, D) [39]) and average handcycling weekly training hours in the last 3 months (hours) were obtained anamnestically. As all individual lesion levels would create too many dummy variables for the analyses, and only 12 individuals with a tetraplegia (of 128 participants) could be included, lesion level was split in two categories: (1) above Th6 and (2) equal to or below Th6 to investigate the effect of loss of sympathetic cardiac innervation (lesion level above Th6) and preserved sympathetic cardiac innervation (lesion level equal to or below Th6) on POpeak [40]. Body mass (kg) was measured on a wheelchair scale with the wheelchair included. Afterwards the mass of the wheelchair was weighted separately and subtracted from the total mass to obtain the body mass of the participant. Body Mass Index (BMI, in kg/m^2) was calculated by dividing the body mass by the squared height. Waist circumference (cm) was measured three times at the level of the umbilicus in supine position. The average of the three measurements was used for analysis. Handcycling classification was determined by an UCI certified Paracycling classifier, following the UCI Para-cycling Regulations: ranging from H1 to H5, in which H1 is the most impaired class and H5 the least impaired class [41].

Statistical Analysis

The analyses were performed using SPSS (IBM SPSS Statistics 20, SPSS, Inc, Chicago, IL, USA) and MLWin software [42].

Descriptives

Means and standard deviations (M ± SD) were calculated for outcome measures and determinants, and data was tested for normality by means of the Kolmogorov–Smirnov test with Lilliefors Significance Correction and the Shapiro–Wilk test. In addition, *z*-scores for skewness and kurtosis were calculated.

Splitting the data

In order to validate the models, the group of 128 participants was randomly split into two samples, using random sample of cases in SPSS: (1) one sample to develop the predicted models (80% of the data; model group) and (2) one sample to cross-validate the models (20% of the data; validation group). This is based on the statement that the ratio of number of independent variables to the number of participants should be at least 1:10 in a multiple linear regression analysis [43]. In this study, 10 possible independent variables were identified; therefore, around 100 participants deemed necessary for the development of the model. First, the two sample groups were checked for systematic differences in baseline values to ensure equality between groups. Thereafter, the predictive model was developed using a multi-level regression analysis to correct for rehabilitation center (i.e., to correct for possible differences in test setting/testers/protocols between the 11 rehabilitation centers). A two-level model was created with participant as first level and center as second level.

Outcome measures and determinants

The dependent variables of the analyses were POpeak (W) and POpeak/kg (W/kg). POpeak/kg was chosen to compare the results of the present study with previous literature [18], and because of the importance of values in W/kg for the HandbikeBattle population as they are participating in an uphill mountain race. The independent variables were: age (years), sex (0=male, 1=female), body mass (kg), BMI in kg/m^2, waist circumference (cm), TSI (years), lesion level (two categories: (1) above Th6 and (2) equal to or below Th6), handcycling classification (two categories: (1) H1-H3 and (2) H4-H5), completeness of the lesion (two categories: (1) motor complete (AIS A+B) and (2) motor incomplete (AIS C+D)) and average handcycling weekly training hours in the last 3 months (h).

Predictive models

First, all variables were checked for multicollinearity as described by Field [44]. Thereafter, all applicable independent variables were used in each of the two theoretical models. For the two statistical models, first, a series of univariate regression models was used within the model group to determine significant associations per variable ($p<0.10$). Thereafter a multi-level regression analysis was performed with all significant variables from the univariate analysis, using a backward elimination technique to develop a model with significant variables only ($p<0.05$). Only simple main effects of determinants were evaluated. For all four models the proportion of explained variance (R^2) was calculated.

Validation of the models

With the use of the developed models, the estimated POpeak was calculated in the validation group (N=24). Thereafter, these estimated scores for POpeak were compared to the (actual) measured POpeak (N=24). Systematic differences between these values were investigated with the paired-samples t-test. The intraclass correlation coefficient was used to measure relative agreement (ICC, two-way random, absolute agreement, single measures) and the Bland-Altman plots with 95% limits of agreement (LoA) to measure absolute agreement [45,46]. The following interpretation was used for the ICC: < 0.40 “poor”, 0.40 - 0.59 “fair”, 0.60 - 0.74 “good”, ≥ 0.75 “excellent” [47].

Reference values

Reference values for POpeak, POpeak/kg, VO_2peak and VO_2peak/kg based on lesion level and sex were developed with the data of all 128 participants. Quintiles were defined based on percentiles: Poor (below 20%), Fair (20% to 40%), Average (40% to 60%), Good (60% to 80%), and Excellent (above 80%), as described by Janssen et al [29].

Results

Descriptives

Means and standard deviations of outcome measures and determinants are depicted in table 1. Main outcome measures were normally distributed.

Splitting the data

No systematic differences in personal and fitness characteristics were observed between the model group and validation group (table 1).

Predictive models

For both models of POpeak and POpeak/kg, a two-level model was created with participant as first level and center as second level. For both models the -2log likelihood did not significantly change after adding center as a level to the constant, i.e., rehabilitation center did not have a substantial effect on the outcome.

Of the possible determinants, lesion level and handcycling classification showed a significant correlation ($r = 0.46$, $p < 0.001$, tolerance = 0.79, variance inflation factor (VIF) = 1.27). Body mass, BMI and waist circumference showed a significant correlation as well ($r \geq 0.78$, $p < 0.001$, tolerance ≤ 0.33, VIF ≥ 3.07 for all correlations). This indicates multicollinearity and, therefore, these variables were not analyzed in combination with each other in the models. Separate models were developed for these variables: BMI and lesion level were used as determinants in the final four models based on significance and proportion of explained variance.

Theory-driven models

In the theoretical model for POpeak, sex, lesion level, handcycling training hours and age were significant determinants. In the theoretical model for POpeak/kg, sex, lesion level, handcycling training hours, BMI and age were significant determinants. R^2 was 42% for both models (table 2).

Statistically-driven models

In the statistical model for POpeak, sex, lesion level, handcycling classification, body mass, BMI and handcycling training hours were significant determinants based on the univariate analysis. In the backward analysis sex, lesion level and handcycling training hours remained significant and formed the final statistical model for POpeak (R^2 = 39%) (table 2).

In the statistical model for POpeak/kg, age, lesion level, body mass, BMI, waist circumference and handcycling training hours were significant determinants based on the univariate analysis. In the backward analysis, lesion level, handcycling training hours and BMI remained significant and formed the final statistical model for POpeak/kg (R^2 = 30%) (table 2).

Validation of the models

For all four models, no systematic differences were found between the predicted POpeak and the measured POpeak. Validation of the models showed varying results, depending on the model (table 3). A fair relative agreement (ICC = 0.43) for the theoretical POpeak model was found, while the Bland-Altman plot showed a large variation (95% LoA -69 to 54 W) indicating a low absolute agreement (figure 1A). The theoretical POpeak/kg model showed a good relative agreement (ICC = 0.60), however, the Bland-Altman plot showed a large variation (95% LoA -0.78 to 0.57 W/kg) for this model as well (figure 1B). A poor relative agreement (ICC = 0.35) for the statistical POpeak model was found, which was supported by the large variation observed in the Bland-Altman plot (95% LoA -64 to 57 W) (figure 1C). Lastly, the statistical POpeak/kg model showed a fair relative agreement (ICC = 0.43), with a large variation (95% LoA -0.92 to 0.68 W/kg) in the Bland-Altman plot (figure 1D).

Reference values

Table 4 and table 5 show reference values for POpeak, POpeak/kg, VO_2peak and VO_2peak/kg based on lesion level and sex, developed with the data of all 128 participants.

Table 2. Results for both theoretical models (with all potential determinants) and for both statistical models (after backward regression analyses) to predict absolute and relative POpeak.

	Theoretical models						**Statistical models**					
	POpeak (N=84)			*POpeak/kg (N=84)*			*POpeak (N=95)*			*POpeak/kg (N=94)*		
	β (SE)	*95%CI*	*p*-value	β (SE)	*95%CI*	*p*-value	β (SE)	*95%CI*	*p*-value	β (SE)	*95%CI*	*p*-value
Intercept	107.05 (18.54)	70.7 143.4	< 0.01	2.94 (0.26)	2.43 3.44	< 0.01	99.97 (5.14)	89.9 110.0	< 0.01	2.36 (0.23)	1.91 2.81	< 0.01
Sex	-41.13 (7.88)	-56.6 -25.7	< 0.01	-0.38 (0.11)	-0.60 -0.16	< 0.01	-41.29 (6.96)	-54.9 -27.6	< 0.01	ns	NA	NA
Lesion level	26.67 (5.90)	15.1 38.2	< 0.01	0.33 (0.08)	0.17 0.49	< 0.01	28.88 (5.69)	17.7 40.0	< 0.01	0.31 (0.09)	0.13 0.49	< 0.01
Handcycling training (h)	1.82 (0.75)	0.35 3.29	0.02	0.03 (0.01)	-0.01 0.05	< 0.01	1.77 (0.71)	0.38 3.16	0.01	0.03 (0.01)	0.01 0.05	0.01
BMI (kg/m^2)	0.52 (0.84)	-1.13 2.17	0.54	-0.06 (0.01)	-0.08 -0.04	< 0.01	ns	NA	NA	-0.05 (0.01)	-0.07 -0.03	< 0.01
TSI (years)	0.18 (0.33)	-0.47 0.83	0.59	0.01 (0.01)	-0.01 0.03	0.23	ns	NA	NA	ns	NA	NA
Completeness	10.92 (6.24)	-1.31 23.15	0.08	0.10 (0.09)	-0.08 0.28	0.24	ns	NA	NA	ns	NA	NA
Age (years)	-0.59 (0.30)	-1.18 -0.002	0.05	-0.01 (0.004)	-0.02 -0.002	< 0.01	ns	NA	NA	ns	NA	NA
R^2	42%			42%			39%			30%		

β (SE): beta with standard error; 95%CI: 95% confidence interval; R^2: proportion of explained variance; ns: non-significant; NA: not applicable. Independent variables: sex (0=male, 1=female), lesion level (two categories: (1) above Th6 and (2) equal to or below Th6), average handcycling weekly training hours in the last 3 months (hours), body mass index (BMI) in kg/m^2, time since injury (TSI, years), completeness following AIS (two categories: (1) motor complete (AIS A+B) and (2) motor incomplete (AIS C+D)), age (years).

Table 3. Validation of the models. Results of comparison between measured and predicted POpeak with intraclass correlation coefficient (N=24).

	Measured	**Theoretical model**	**ICC (95% CI)**	**Statistical model**	**ICC (95% CI)**
	M ± SD	*M ± SD*		*M ± SD*	
POpeak (W)	121 ± 40	123 ± 17	0.43 (-0.03 to 0.74)*	126 ± 14	0.35 (-0.09 to 0.68)
POpeak/kg (W/kg)	1.56 ± 0.51	1.50 ± 0.31	0.60 (0.21 to 0.82)*	1.52 ± 0.23	0.43 (0.01 to 0.72)*

M ± SD: indicates mean ± standard deviation; 95% CI = 95% confidence interval. * Significant correlation ($p < 0.05$).

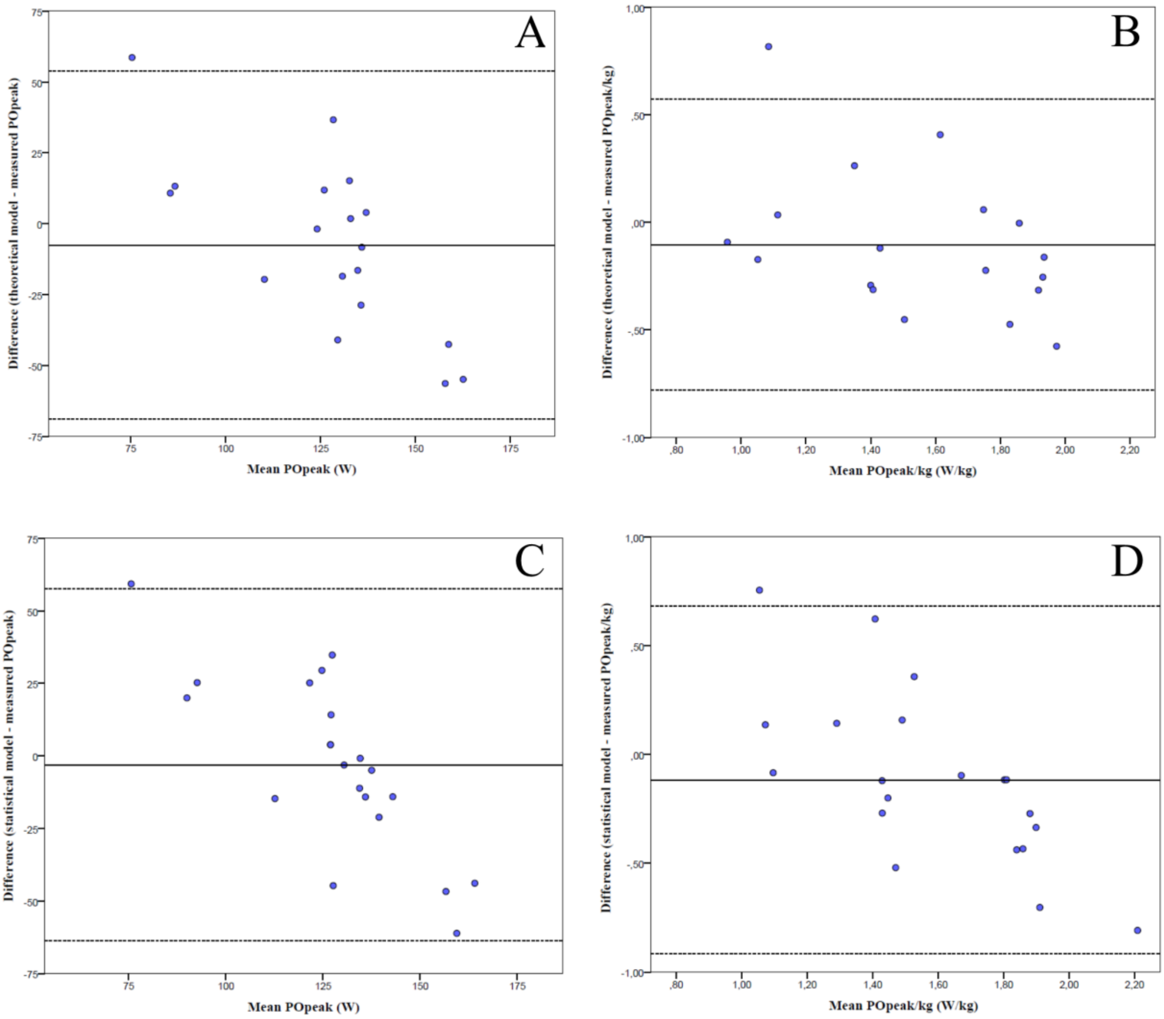


Figure 1. Bland-Altman plots representing the absolute agreement between the predicted POpeak and the measured POpeak. Solid line represents the mean, dotted lines represent mean ± 2SD (95% LoA). Each circle represents a participant of the validation group. A: The difference in POpeak between the POpeak predicted with the theoretical model and the measured POpeak. B: The difference in POpeak/kg between the POpeak/kg predicted with the theoretical model and the measured POpeak/kg. C: The difference in POpeak between the POpeak predicted with the statistical model and the measured POpeak. D: The difference in POpeak/kg between the POpeak/kg predicted with the statistical model and the measured POpeak/kg.

Table 4. Reference values for POpeak, POpeak/kg, VO_2peak and VO_2peak/kg, for participants with (1) lesion level above Th6 (>Th6) and (2) equal to or below Th6 (≤Th6).

Variable	Level	N	Poor	Fair	Average	Good	Excellent
POpeak (W)	>Th6	37	< 63	63 - 96	96 - 117	117 – 137	> 137
	≤Th6	86	< 101	101 - 120	120 - 136	136 – 154	> 154
POpeak/kg (W/kg)	>Th6	37	< 0.81	0.81 – 1.16	1.16 – 1.47	1.47 – 1.79	> 1.79
	≤Th6	85	< 1.31	1.31 – 1.52	1.52 – 1.70	1.70 – 2.01	> 2.01
VO_2peak (L/min)	>Th6	37	< 1.11	1.11 – 1.47	1.47 – 1.72	1.72 – 2.02	> 2.02
	≤Th6	84	< 1.65	1.65 – 1.95	1.95 – 2.20	2.20 – 2.49	> 2.49
VO_2peak/kg (ml/kg/min)	>Th6	37	< 15.53	15.53 – 17.57	17.57 – 21.90	21.90 – 26.63	> 26.63
	≤Th6	83	< 21.18	21.18 – 24.61	24.61 – 27.42	27.42 – 31.58	> 31.58

Poor (<20%), Fair (20-40%), Average (40-60%), Good (60-80%) and Excellent (>80%) (N=128).

Table 5. Reference values for POpeak, POpeak/kg, VO_2peak and VO_2peak/kg, for male (M) and female (F) participants.

Variable	Sex	N	Poor	Fair	Average	Good	Excellent
POpeak (W)	M	106	< 104	104 - 120	120 - 135	135 - 150	> 150
	F	22	< 69	69 - 81	81 - 92	92 – 107	> 107
POpeak/kg (W/kg)	M	105	< 1.18	1.18 – 1.47	1.47 – 1.65	1.65 – 2.05	> 2.05
	F	22	< 1.10	1.10 – 1.32	1.32 – 1.53	1.53 – 1.64	> 1.64
VO_2peak (L/min)	M	105	< 1.53	1.53 – 1.80	1.80 – 2.08	2.08 – 2.43	> 2.43
	F	21	< 1.09	1.09 – 1.33	1.33 – 1.66	1.66 – 1.82	> 1.82
VO_2peak/kg (mL/kg/min)	M	104	< 18.08	18.08 – 22.68	22.68 – 26.69	26.69 – 30.76	> 30.76
	F	21	< 17.89	17.89 – 22.11	22.11 – 24.45	24.45 – 27.77	> 27.77

Poor (<20%), Fair (20-40%), Average (40-60%), Good (60-80%) and Excellent (>80%) (N=128).

Discussion

This study is the first to have developed and validated predictive models and reference values for synchronous handcycling. Four predictive models on POpeak (W and W/kg) were developed in a group of recreational handcyclists: two theory-driven models and two statistically-driven models. The theoretical models showed a somewhat higher explained variance than the statistical models, although overall the explained variance was low for all four models (R^2 ranged from 30% to 42%). Validation of the models showed a poor to good relative agreement, depending on the model, with a low absolute agreement for all models. In accordance with the third aim, reference values for POpeak, POpeak/kg, VO_2peak and VO_2peak/kg based on lesion level and sex were developed.

Predictive models

Due to missing data, both theoretical models were based on fewer participants (N=84) than the statistical models (N=94-95) (table 2). However, these models showed more statistically significant determinants and a higher explained variance than the statistical models. This might be due to a different interdependent association between the determinants in these models; in the theoretical models all determinants were included simultaneously (forced entry) based on our understanding of interdependency, whereas in the statistical models first an univariate analysis was performed. In this univariate analysis, some determinants were excluded from the model based on their individual association with POpeak, obviously without considering their possible indirect association with POpeak through their interactions with other determinants. Compared to theory-driven modeling, this is a disadvantage of stepwise statistical modeling as only mathematical criteria are used to select determinants [44]. In future studies, it could be interesting to focus on these possible interactions between determinants when modeling physical capacity in individuals with SCI.

Theory-driven models

In this study, two theory-driven models for POpeak were developed using multi-level regression analysis. The selection of determinants was based on theoretical constructs, investigated in previous wheelchair and arm ergometry literature concerning individuals with an SCI. The aim was to gain more insight in the underlying determinants influencing physical capacity in individuals with SCI during handcycling. The results showed that sex, lesion level, handcycling training hours and age are significant determinants for POpeak (table 2). Of these determinants handcycling training hours is the only determinant that can be influenced. Therefore, in order to increase physical capacity in individuals with an SCI during handcycling, individually optimized training intensity and volume should be encouraged. Another modifiable determinant, BMI, was positively related to POpeak, although not significant, and inversely related to POpeak/kg, which indicates a decrease in physical capacity with every increase in BMI. This can partly be explained by the shared term for mass in the outcome measure (POpeak/kg) and the determinant (BMI). Comparable relationships were previously described by Janssen et al. [29] and Simmons et al. [18] in wheelchair ergometry and asynchronous arm ergometry, respectively. They explain that an elevated BMI in this population is, therefore, probably related to overweight due to adipose tissue and a low physical activity, instead of a large muscle mass. BMI was chosen in this study (instead of bio impedance analysis or DXA) due to its wide use in literature and clinical practice, inexpensiveness, applicability, and in order to compare our results with previous literature about predictive models in wheelchair exercise and asynchronous arm ergometry.

Statistically-driven models

Next to the theory-driven models, two statistically-driven models were developed. The aim was to use multi-level regression analyses with a backward elimination technique to accurately predict POpeak during handcycling based on statistically significant determinants. Results showed that only three determinants appeared to be statistically significant determinants (sex, lesion level and handcycling training for POpeak, and lesion level, handcycling training and BMI for POpeak/kg) following the current statistical selection criteria and backward approach. In previous literature, only statistical models were developed to investigate the association between POpeak and participant characteristics, based on wheelchair testing and asynchronous arm ergometry. Simmons et al. [18] developed a model for POpeak during asynchronous arm ergometry in untrained individuals with an SCI based on motor level of injury, functional classification and sex (R^2 = 0.57) and a model for POpeak/kg based on functional classification, BMI and motor level of injury (R^2 = 0.57) using (forward) stepwise regression. Other possible factors such as age, TSI and completeness were not significantly correlated to POpeak in the study of Simmons et al., [18] comparable to the results in the present study. An important difference between the study by Simmons et al. and the present study is the determinant handcycling training (hours). This determinant was significant in both statistical models in the present study, however, was not analyzed in the study by Simmons et al. Janssen et al. [29] found a comparable determinant, activity level, to be significantly related to POpeak in wheelchair exercise testing. Moreover, several other studies highlighted the relationship between activity level or sports participation and physical capacity in individuals with a SCI during wheelchair testing [32,35] and asynchronous arm ergometry [9,17].

Despite the significant determinants that were found, a large part of the variance in the present study remained unexplained (58-70%). This might have several reasons. First, due to the multicenter character of the study, different test assistants performed the tests and different test equipment and protocols were used. This causes inevitable variability in test results. Although, in the present study, no significant differences were found between rehabilitation centers, test equipment and protocols, it would be optimal to standardize these measures in order to pursue homogeneity. However, the reader should be aware that in order to achieve a large number of participants in rehabilitation related research, homogeneity is only possible to a certain extent. In this study, a correction was made for the possible (non-significant) differences between rehabilitation centers by multi-level regression analysis. Second, we need to critically evaluate the way determinants are reported and consider other possible determinants. For example, handcycling training was reported; however, other activities of daily living and lifestyle factors as well as other types of training (e.g., swimming, wheelchair rugby, but also strength training) were not taken into account as the response rate on these separate questions and the validity of the answers were considered too low to be representative. This is unfortunate, as these factors might explain a larger part of the variance than handcycling training alone. Moreover, training hours do not take the actual intensity level into account. Therefore, an overall, easy to use measure of training load should be considered such as Training Impulse based on session ratings of perceived exertion (sRPE) [48,49], to increase the proportion of explained variance.

As emphasized by Van Der Woude et al. [50], POpeak is associated with several factors, including the factors that were taken into account in the present study. POpeak is, however, also directly related to the mode of exercise (e.g., handrim wheelchair or handbike propulsion), including notions of efficiency, skill and talent, as well as aerobic exercise (cardiorespiratory) and anaerobic capacity. POpeak is, therefore, a general measure of handcycling physical capacity. This is in contrast to VO_2peak, as VO_2peak is a general measure of cardiorespiratory function only [50,51]. Therefore, more factors associated with POpeak should be taken into account. For example, in a previous study by Janssen et al. [30] a strong association was found between anaerobic POpeak and aerobic POpeak (R^2 = 81%) in individuals with an SCI on a wheelchair ergometer. Future studies could focus on this association in handcycling with, for example, a Wingate Test, which might lead to a higher explained variance and, subsequently, better estimation of POpeak.

Validation of the models

To the authors' knowledge, this is the first study that investigated validity of a POpeak prediction model in arm exercise. Despite a good relative agreement for the theoretical POpeak/kg model, all models showed a low absolute agreement as represented by the high variation in the Bland-Altman plots (figure 1). Although a high relative and absolute agreement are desirable, it must be emphasized that these models were not designed to replace the GXT. It is, therefore, not necessarily needed to predict the exact POpeak, a certain valid range, however, is a prerequisite. It has been suggested that a test duration of 8 – 12 minutes would be optimal to achieve peak physiological responses during a GXT [21,25], although it is important to mention that the optimal test duration for arm exercise is not known [52]. This test duration is important, as it is inherent to the number of steps and the step size of the protocol. Studies have shown that when the step size is too large, and consequently the test is too short, peak physical capacity tends to be overestimated and studying the effect of certain therapy or training is less reliable [25]. However, when the test is too long due to the small step size or long step duration, peak physical capacity tends to be underestimated [21,24]. As an average test duration of 10 minutes ± 20% is said to be optimal, it could be argued that a predicted POpeak within a range of ± 20% is a valid value to use in the selection of an individualized GXT protocol. In this study, depending on the model, 52 – 67% of the predicted POpeak values fell within this range. This indicates that the validity of the models is not high enough to solely base GXT protocol selection on. Therefore, future research should focus on improving the validity of these models and diminishing the large proportion of unexplained variance.

Reference values

To date, this is the first study that describes reference values for (synchronous) handcycling based on a large group of handcycle users with SCI. Comparing the results to previous literature, it has to be emphasized that our group was heterogeneous and that not all participants were completely untrained. In the study by Lovell et al. [53], a mean POpeak of 121 W was found for untrained handcyclists with paraplegia, which is comparable to the results in the present study (120 – 136W). It must be emphasized that it is unclear whether synchronous or asynchronous arm cranking was performed in the study by Lovell et al. Due to the heterogeneity of the population in the present study, the reference values will give a good reflection of the diversity in the SCI population. However, individuals with a very low physical capacity or absent training motivation are probably not represented in this study, as these individuals are not motivated to participate in a mountain race. Moreover, elite handcyclists did not participate in our study, as a POpeak of 210 W as described by Lovell et al. [53] for “trained” handcyclists with an SCI was reached by none of the participants in the present study. This has to be considered when interpreting the predictive models and reference values.

Next to training status, other factors need to be kept in mind comparing the results of the present study to previous research. For example, test device (wheelchair ergometry versus arm ergometry versus handcycling), propulsion mode (asynchronous versus synchronous), test protocol and other participant characteristics. Overall, the reference values of the present study were higher compared to values found in previous studies focusing on asynchronous arm ergometry. Simmons et al. [18] found an average POpeak of 62 – 78 W and 0.85 – 0.98 W/kg during (asynchronous) arm ergometry for men with paraplegia, compared to 120 – 136 W and 1.52 – 1.70 W/kg, respectively, for the group with low paraplegia in the present study. Next to POpeak, VO_2peak showed higher values in the present study: Simmons et al. [18] found an average VO_2peak of 1.28 – 1.41 L/min and 15.31 – 17.69 ml/kg/min during arm ergometry for men with paraplegia, compared to an average VO_2peak of 1.95 – 2.20 L/min and 24.61 – 27.42 ml/kg/min, respectively, in the present study. Earlier reviews by Haisma et al. [7] and Valent et al. [54] studying reference values for individuals with paraplegia during asynchronous arm ergometry support the finding of Simmons et al. The reviews showed a POpeak of 66 – 117 W [7] and a VO_2peak of 1.06 – 2.34 L/min [7] and 1.33 – 1.90 L/min [54].

The reference values found in the present study are comparable to a previous study investigating synchronous handcycling [55]. Janssen et al. performed a descriptive study with 16 male handcycle users, measuring physical capacity by means of a GXT in an add-on handcycle on a treadmill [55]. Although not exclusively individuals with an SCI were studied, they found similar values for the group with lower-limb disabilities: 129 ± 26 W and 1.64 ± 0.32 W/kg, comparable to results of the present study. Dallmeijer et al. [3] studied physical capacity by means of a GXT in an add-on handcycle on a treadmill in nine men with a paraplegia and found a POpeak of 117 ± 32 W and a VO_2peak of 1.88 ± 0.44 L/min. These results are slightly lower than in the present study.

Implications

The theoretical POpeak/kg model was the best predicting model to assess POpeak, with an explained variance of 42% and ICC of 0.60. However, a large part of the variance still remained unexplained and the Bland-Altmann plot showed a low absolute agreement. Moreover, the finding that only 67% of the predicted POpeak values fell into the range of ± 20% indicates that the validity of this model is not high enough to solely base GXT protocol selection on. Therefore, the models should be used with caution and only in addition to expert opinion of the practitioner when there is indecisiveness in what protocol to choose. It must be explicitly emphasized that the models should not be used to replace a GXT. In future studies standardization of test setting and protocol is necessary.

The same large part of unexplained variance is reflected on the reference values. Nevertheless, this is the first study to describe reference values for (synchronous) handcycling in individuals with an SCI. Although the values should be used with caution, they give a global overview of the physical capacity of individuals with an SCI during handcycling. As these values are based on a large heterogeneous group, they give an indication of the normal variation in the SCI population, for both men and women, and only applicable to synchronous handcycling.

Study limitations

There was variation in the measurement set-up due to the fact that tests were performed in 11 different rehabilitation centers. Although, in the present study, no significant effect of rehabilitation centers was found, it would be optimal to standardize these measures in order to pursue homogeneity. Second, due to the low number of individuals with a tetraplegia (N=12), it was not possible to divide the group in people with tetraplegia and paraplegia. The results of this study are, therefore, not applicable to individuals with a tetraplegia. Moreover, due to the relatively low number of female participants (N=22) it was not possible to define reference values based on sex and lesion level together. Therefore, separate reference values were defined; 1) for lesion level, and 2) for sex. Lastly, possible important determinants such as training load were not taken into account. This might be interesting for future research.

Conclusion

This study is the first to have developed and validated predictive models and reference values for synchronous handcycling. Lesion level, handcycling training hours and sex or BMI appeared to be significant determinants of POpeak in handcyclists with SCI in all four models. The theoretical models showed the highest proportion of explained variance. Validation showed varying relative agreement, and a low absolute agreement. Moreover, a large part of the variance remained unexplained in all models. Therefore, these models and reference values might be useful in clinical practice, but should not replace a GXT. Both models and reference values provide insight in physical capacity of the diverse SCI population, based on a relatively large sample performing synchronous handcycling GXT.

Disclosure statement

No potential conflict of interest was reported by the authors.

Acknowledgements

*HandbikeBattle group name: Paul Grandjean Perrenod Comtesse, Adelante Zorggroep, Hoensbroek, The Netherlands. Eric Helmantel, University Medical Center Groningen, Center for Rehabilitation Beatrixoord, Groningen, The Netherlands. Mark van de Mijll Dekker, Heliomare Rehabilitation Center, Wijk aan Zee, The Netherlands. Maremka Zwinkels, Rehabilitation Center De Hoogstraat, Utrecht, The Netherlands. Misha Metsaars, Libra Rehabilitation and Audiology, Eindhoven, The Netherlands. Lise Wilders, Sint Maartenskliniek, Nijmegen, The Netherlands. Linda van Vliet, Amsterdam Rehabilitation Research Center | Reade, Amsterdam, The Netherlands. Karin Postma, Rijndam Rehabilitation Center, Rotterdam. Bram van Gemeren, Roessingh Rehabilitation Center, Enschede, The Netherlands. Selma Overbeek, Jeroen Bosch Hospital, Tolbrug Rehabilitation Centre, ‘s-Hertogenbosch, The Netherlands. Alinda Gjaltema, Vogellanden, Zwolle, The Netherlands.

References

1. Abel T, Vanlandewijck Y, Verellen J. Handcycling. In: Goosey-Tolfrey V, editor. *Wheelchair Sport*. Champaign, IL: Human Kinetics; 2010. p. 187-197.

2. Arnet U, van Drongelen S, Scheel-Sailer A, van der Woude LHV, Veeger HEJ. Shoulder load during synchronous handcycling and handrim wheelchair propulsion in persons with paraplegia. *J Rehabil Med*. 2012;44(3):222-228.

3. Dallmeijer AJ, Zentgraaff IDB, Zijp NI, van der Woude LHV. Submaximal physical strain and peak performance in handcycling versus handrim wheelchair propulsion. *Spinal Cord*. 2004;42(2):91-98.

4. Van der Woude LHV, Dallmeijer AJ, Janssen TWJ, Veeger HEJ. Alternative modes of manual wheelchair ambulation: an overview. *Am J Phys Med Rehabil*. 2001;80(10):765-777.

5. Valent LJM, Dallmeijer AJ, Houdijk H, et al. Effects of hand cycle training on physical capacity in individuals with tetraplegia: a clinical trial. *Phys Ther*. 2009;89(10):1051-1060.

6. Van den Berg-Emons RJ, Bussmann JBJ, Haisma JA, et al. A Prospective study on physical activity levels after spinal cord injury during inpatient rehabilitation and the year after discharge. *Arch Phys Med Rehabil*. 2008;89(11):2094-2101.

7. 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(11):642-652.

8. Jacobs PL, Nash MS. Exercise recommendations for individuals with spinal cord injury. *Sport Med*. 2004;34(11):727-751.

9. Muraki S, Tsunawake N, Tahara Y, Hiramatsu S, Yamasaki M. Multivariate analysis of factors influencing physical work capacity in wheelchair-dependent paraplegics with spinal cord injury. *Eur J Appl Physiol*. 2000;81(1-2):28-32.

10. De Groot S, Dallmeijer AJ, Post MWM, Angenot ELD, van der Woude LHV. The longitudinal relationship between lipid profile and physical capacity in persons with a recent spinal cord injury. *Spinal Cord*. 2008;46(5):344-351.

11. Van Koppenhagen CF, Post MWM, de Groot S, 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(3):328-337.

12. Manns PJ, Chad KE. Determining the relation between quality of life, handicap, fitness, and physical activity for persons with spinal cord injury. *Arch Phys Med Rehabil*. 1999;80(12):1566-1571.

13. Van Velzen JM, de Groot S, Post MWM, Slootman JR, van Bennekom CAM, van der Woude LHV. Return to work after spinal cord injury. *Am J Phys Med Rehabil*. 2009;88(1):47-56.

14. Noreau L, Shephard RJ. Return to work after spinal cord injury: The potential contribution of physical fitness. *Paraplegia*. 1992;30(8):563-572.

15. Goosey-Tolfrey VL, Sindall P. The effects of arm crank strategy on physiological responses and mechanical efficiency during submaximal exercise. *J Sports Sci*. 2007;25(4):453-460.

16. Hooker SP, Wells CL. Aerobic power of competitive paraplegic road racers. *Paraplegia*. 1992;30(6):428-436.

17. Hutzler Y, Ochana S, Bolotin R, Kalina E. Aerobic and anaerobic arm-cranking power outputs of males with lower limb impairments: relationship with sport participation intensity, age, impairment and functional classification. *Spinal Cord*. 1998;36(3):205-212.

18. Simmons OL, Kressler J, Nash MS. Reference fitness values in the untrained spinal cord injury population. *Arch Phys Med Rehabil*. 2014;95(12):2272-2278.

19. Mossberg K, Willman C, Topor MA, Crook H, Patak S. Comparison of asynchronous versus synchronous arm crank ergometry. *Spinal Cord.* 1999;37:569-574.

20. De Groot S, Postma K, van Vliet L, Timmermans R, Valent LJM. Mountain time trial in handcycling: Exercise intensity and predictors of race time in people with spinal cord injury. *Spinal Cord*. 2014;52(6):455-461.

21. Buchfuhrer MJ, Hansen JE, Robinson TE, Sue DY, Wasserman K, Whipp BJ. Optimizing the exercise protocol for cardiopulmonary assessment. *J Appl Physiol Respir Environ Exerc Physiol*. 1983;55(5):1558-1564.

22. Smith PM, Amaral I, Doherty M, Price MJ, Jones AM. The influence of ramp rate on VO_2peak and “excess” VO_2 during arm crank ergometry. *Int J Sports Med*. 2006;27(8):610-616.

23. Eerden S, Dekker R, Hettinga 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(5):497-521.

24. Bentley DJ, Newell J, Bishop D. Incremental exercise test design and analysis: Implications for performance diagnostics in endurance athletes. *Sport Med*. 2007;37(7):575-586.

25. Myers J, Bellin D. Ramp exercise protocols for clinical and cardiopulmonary exercise testing. *Sports Med*. 2000;30(1):23-29.

26. Dallmeijer AJ, Kilkens OJE, Post MWM, et al. Hand-rim wheelchair propulsion capacity during rehabilitation of persons with spinal cord injury. *J Rehabil Res Dev*. 2005;42(3 Suppl 1):55-64.

27. Falkel JE, Sawka MN, Levine L, Pimental NA, Pandolf KB. Upper-body exercise performance: Comparison between women and men. *Ergonomics*. 1986;29(1):145-154.

28. Coutts KD, Rhodes EC, McKenzie DC. Maximal exercise responses of tetraplegics and paraplegics. *J Appl Physiol Respir Environ Exerc Physiol*. 1983;55(2):479-482.

29. Janssen TWJ, Dallmeijer AJ, Veeger HEJ, van der Woude LHV. Normative values and determinants of physical capacity in individuals with spinal cord injury. *J Rehabil Res Dev*. 2002;39(1):29-39.

30. Janssen TWJ, van Oers CA, Hollander AP, Veeger HEJ, van der Woude LHV. Isometric strength, sprint power, and aerobic power in individuals with a spinal cord injury. *Med Sci Sports Exerc*. 1993;25(7):863-870.

31. Sawka MN, Glaser RM, Laubach LL, Al-Samkari O, Suryaprasad AG. Wheelchair exercise performance of the young, middle-aged, and elderly. *J Appl Physiol Respir Environ Exerc Physiol*. 1981;50(4):824-828.

32. Dallmeijer AJ, Hopman MT, van As HH, van der Woude LHV. Physical capacity and physical strain in persons with tetraplegia; the role of sport activity. *Spinal Cord*. 1996;34(12):729-735.

33. Hicks AL, Martin Ginis KA, Pelletier CA, Ditor DS, Foulon B, Wolfe DL. The effects of exercise training on physical capacity, strength, body composition and functional performance among adults with spinal cord injury: A systematic review. *Spinal Cord*. 2011;49(11):1103-1127.

34. Janssen TWJ, van Oers CA, Rozendaal EP, Willemsen EM, Hollander AP, van der Woude LHV. Changes in physical strain and physical capacity in men with spinal cord injuries. *Med Sci Sports Exerc*. 1996;28(5):551-559.

35. Dallmeijer AJ, van der Woude LHV, Hollander AP, Angenot EL. Physical performance in persons with spinal cord injuries after discharge from rehabilitation. *Med Sci Sport Exerc*. 1999;31(8):1111-1117.

36. de Groot S, van der Scheer JW, Bakkum AJT, et al. Wheelchair-specific fitness of persons with a long-term spinal cord injury: cross-sectional study on effects of time since injury and physical activity level. *Disabil Rehabil*. 2016;38(12):1180-1186.

37. Hoekstra SP, Valent LJM, Janssen TWJ, Paulich M, van der Woude LHV, de Groot S. Transferability of exercise test results among three arm ergometers. Poster session presented at: Biannual conference of

the international Paralympic Committee, VISTA; 2015; Girona, Spain.

38. Smith PM, Doherty M, Drake D, Price MJ. The influence of step and ramp type protocols on the attainment of peak physiological responses during arm crank ergometry. *Int J Sports Med*. 2004;25(8):616-621.

39. Kirshblum SC, Burns SP, Biering-Sorensen F, et al. International standards for neurological classification of spinal cord injury (Revised 2011). *J Spinal Cord Med*. 2011;34(6):535-546.

40. Krassioukov A, West C. The role of autonomic function on sport performance in athletes with spinal cord injury. *PM&R*. 2014;6(8):S58-S65.

41. UCI Cycling Regulations, Part 16. Para-cycling; 2017. p1-75. Available from: UCI Para-cycling. https://www.uci.org/docs/default-source/rules-and-regulations/16-par-20200211-e.pdf

42. Rasbash J, Charlton C, Browne W, Healy M, Cameron B. MLwiN Version 2.02. Centre for Multilevel Modelling. Bristol (UK): University of Bristol; 2005.

43. Altman DG, Royston P. What do we mean by validating a prognostic model? *Stat Med*. 2000;19:453-473.

44. Field A. Chapter 8 Regression. In: Carmichael M, editor. *Discovering Statistics Using IBM SPSS Statistics*. 4th ed. London (UK): SAGE; 2013. p 321-326.

45. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*. 1986;327:307-310.

46. Mantha S, Roizen MF, Fleisher LA, Thisted R, Foss J. Comparing methods of clinical measurement: reporting standards for bland and altman analysis. *Anesth Analg*. 2000;90(3):593-602.

47. Cicchetti DV. Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychol Assess*. 1994;6(4):284-290.

48. Foster C, Florhaug JA, Franklin J, et al. A New approach to monitoring exercise training. *J Strength Cond Res*. 2001;15(1):109-115.

49. De Groot S, Hoekstra SP, Grandjean Perrenod Comtesse P, Kouwijzer I, Valent LJM. Relationships between internal and external handcycle training load in people with spinal cord injury training for the HandbikeBattle. *J Rehabil Med*. 2018;50(3):261-268.

50. Van der Woude LHV, de Groot S, van Drongelen S, et al. Evaluation of manual wheelchair performance in everyday life. *Top Spinal Cord Inj Rehabil*. 2009;15(2): 1-15.

51. Kilkens OJ, Dallmeijer AJ, Nene AV, Post MW, van der Woude LHV. The longitudinal relation between physical capacity and wheelchair skill performance during inpatient rehabilitation of people with spinal cord injury. *Arch Phys Med Rehabil*. 2005;86(8):1575-1581.

52. Maher JL, Cowan RE. Comparison of 1- versus 3-minute stage duration during arm ergometry in individuals with spinal cord injury. *Arch Phys Med Rehabil*. 2016;97(11):1895-1900.

53. Lovell D, Shields D, Beck B, Cuneo R, McLellan C. The aerobic performance of trained and untrained handcyclists with spinal cord injury. *Eur J Appl Physiol*. 2012;112(9):3431-3437.

54. Valent LJM, Dallmeijer AJ, Houdijk H, Talsma E, van der Woude LHV. The effects of upper body exercise on the physical capacity of people with a spinal cord injury: a systematic review. *Clin Rehabil*. 2007;21(4):315-330.

55. Janssen TWJ, Dallmeijer AJ, van der Woude LHV. Physical capacity and race performance of handcycle users. *J Rehabil Res Dev*. 2001;38(1):33-40.

Chapter 3

Interrater and intrarater reliability of ventilatory thresholds determined in individuals with spinal cord injury

Kouwijzer I, Cowan RE, Maher JL, Groot FP, Riedstra F, Valent LJM, van der Woude LHV, de Groot S. Interrater and intrarater reliability of ventilatory thresholds determined in individuals with spinal cord injury. Spinal Cord 2019;57(8):669-678.

Abstract

Study design: Cross-sectional.

Objectives: Individualized training regimes are often based on ventilatory thresholds (VTs). The objectives were to study: 1) whether VTs during arm ergometry could be determined in individuals with spinal cord injury (SCI), 2) the intra- and interrater reliability of VT determination.

Setting: University research laboratory.

Methods: Thirty graded arm crank ergometry exercise tests with 1-min increments of recreationally active individuals (tetraplegia (N=11), paraplegia (N=19)) were assessed. Two sports physicians assessed all tests blinded, randomly, in two sessions, for VT1 and VT2, resulting in 240 possible VTs. Power output (PO), heart rate (HR) and oxygen uptake (VO_2) at each VT were compared between sessions or raters using paired samples t-tests, Wilcoxon signed-rank tests, intraclass correlation coefficients (ICC, relative agreement) and Bland Altman plots (random error, absolute agreement).

Results: Of the 240 VTs, 217 (90%) could be determined. Of the 23 undetermined VTs, 2 (9%) were VT1 and 21 (91%) were VT2; 7 (30%) among individuals with paraplegia and 16 (70%) among individuals with tetraplegia. For the successfully determined VTs, there were no systematic differences between sessions or raters. Intrarater and interrater ICCs for PO, HR, and VO_2 at each VT were high to very high (0.82 – 1.00). Random error was small to large within raters, and large between raters.

Conclusions: For VTs that could be determined, relative agreement was high to very high, absolute agreement varied. For some individuals, often with tetraplegia, VT determination was not possible, thus other methods should be considered to prescribe exercise intensity.

Introduction

In wheelchair users with spinal cord injury (SCI) cardiorespiratory fitness is generally reduced [1]. Low cardiorespiratory fitness and low levels of physical activity are shown to be associated with high prevalence of cardiometabolic disease, which is the leading cause of mortality in this population [2,3]. To increase cardiorespiratory fitness, exercise interventions such as handcycling may be introduced during or after rehabilitation [4–6]. To promote handcycling in the Netherlands and to increase cardiorespiratory fitness after rehabilitation, an annual handcycle race called The HandbikeBattle [7,8] has been held since 2013 in Austria. To optimally train for events like this, but also in or after rehabilitation in general, individualized training schemes are required.

Individualized training schemes can be based on results of a graded exercise test (GXT). Training prescriptions based on maximum values, such as percentage maximum heart rate (HR) or power output (PO) are common, as well as prescriptions based on percentage HR reserve or ventilatory thresholds (VTs) [9]. Training intensity prescription based on relative percentages is shown to have downsides in able-bodied individuals. It seems not to take into account the individual's metabolic response to exercise, and has shown less improvements in maximum oxygen uptake (VO_2max) after training compared with training intensity based on VTs [9,10]. Therefore, prescribing training intensities based on VTs may more reliably achieve fitness gains. The first ventilatory threshold (VT1) is a physiological point during exercise at which a nonlinear increase in carbon dioxide (CO_2) production occurs, coinciding with the first increase in lactate production [11]. The second ventilatory threshold (VT2) represents the onset of exercise-induced hyperventilation with respect to VCO_2 as a reaction to metabolic acidosis, which coincides with the maximal lactate steady state [11,12]. These VTs provide boundaries that allow to set individualized training zones: zone 1 at low intensity (below VT1), zone 2 at moderate intensity (between VT1 and VT2) and zone 3 at high intensity (above VT2) [12,13]. This training principle has been developed in studies on lower-body exercise with able-bodied participants and athletes, and little or no research has been done regarding the reliability of VT determination in upper-body GXT in individuals with SCI. Therefore, the question arises whether the reliability of determination of both VTs is sufficient to set training schemes for individuals with SCI.

For able-bodied leg exercise, VT1 is normally positioned at 50-60% peak oxygen uptake (VO_2peak) and VT2 at 70–80% VO_2peak [14]. This is, however, dependent on cardiorespiratory fitness as values for VT1 and VT2 could increase to 75% and 90% VO_2peak for elite endurance athletes [12]. Studies in able-bodied cycling showed that experienced raters are able to identify VT1 in 90–94% of participants [15,16]. Intrarater reliability of VT1 determination was high (intraclass correlation coefficient (ICC) 0.97) in one study [17], whereas interrater reliability varied (ICC 0.21–0.98) within and between studies [16,17]. The identification rate and reliability of VT2 identification are largely unknown; only one study reports on this topic with an intrarater reliability (ICC) of 0.94–0.96 and an interrater reliability of 0.81–0.91 [18].

However, few studies reported on VTs during upper-body exercise in individuals with SCI. In two studies 89 – 96% of VT1 and 74% of VT2 could be determined in wheelchair athletes with SCI [19,20]. In both studies almost all undetermined VTs appeared to involve athletes with tetraplegia. Leicht et al. [19] explained that for athletes with tetraplegia the percentage of identifiable VT2s might be lower compared with able-bodied athletes, as the absolute ventilatory responses are generally low, resulting in a narrower range of ventilatory values compared with able-bodied athletes. A very recent study supports these findings as VT1 was only identified in 68% of untrained individuals with tetraplegia [21]. For the VT1, Coutts et al. reported a (Pearson) correlation of 0.95 between two raters for athletes with paraplegia and tetraplegia [20], and Bhambani et al. reported a Pearson correlation of 0.90 between two raters for trained and untrained individuals with tetraplegia [22]. However, although ICCs are more appropriate to asses intrarater and interrater reliability than Pearson correlations, they were not reported.

Unfortunately, no studies reported on reliability of VT determination for both thresholds, investigated in a non-athlete population with SCI. Therefore, it remains unclear whether VTs can be used to set individualized training schemes in this less fit population. The aims were, therefore:

1) To examine whether it is possible to detect both VTs in recreationally active individuals with tetraplegia or paraplegia.
2) To examine the inter- and intrarater reliability of VT determination.

Methods

The present study was a retrospective study: the data of the GXTs with 1-min increments of a previous study by Maher et al. [23] were re-analyzed to answer the research questions. Two sports physicians independently evaluated the tests twice during two separate sessions.

Participants

Thirty-three recreationally active individuals with SCI were recruited to participate in the study: 19 individuals with paraplegia and 14 with tetraplegia, 28 men, age: 38±10 years, time since injury (TSI): 12±9 years, body mass: 76±19 kg, height: 1.75±0.08 m. They were recruited through the Miami Project to Cure Paralysis database, and voluntarily trained at the Miami Project gym at least once a week. Inclusion criteria: age ≥ 18 years, non-progressive SCI, TSI of at least six months and self-reported inability to use lower extremity contractions to assist in transfers. Exclusion criteria: angina or myocardial infarction within the last month or pain in the upper extremities [23]. Informed consent was obtained from all participants included in the study. The study was approved by the University of Miami Institutional Review Board, Miami, United States of America.

Test procedure

All GXTs were performed with an (asynchronous) arm crank ergometer (Lode Angio, Groningen, the Netherlands). Participants performed the tests in their own wheelchair; positioned with arms slightly flexed in the furthest horizontal position; participants with tetraplegia used hand wraps to ensure a tight grip on the cranks; and wedges were used to minimize movement of the wheelchair. As individualized protocols are preferred for individuals with SCI [24,25], the starting work load and step size of every participant were individualized based on questions regarding activity level, current fitness program and the ability to perform a floor-to-chair transfer [23]. The aim was to develop an individualized 1-min stepwise protocol with a duration between 8 and 12 min [26]. This resulted in an individualized starting workload of 5–90 W and step size of 10 W for participants with paraplegia, and start workload of 5–30 W and step size of 3-10 W for participants with tetraplegia. The prescribed cadence was between 60 and 65 rpm. Criteria to stop the test were volitional exhaustion or failure in keeping a constant cadence above 55 RPM. During the test, PO (W) was continuously measured. Gas exchange was measured breath-by-breath (Vmax Encore metabolic cart, Carefusion, Vyaire Medical, Mettawa, IL, USA) and HR was measured by standard 12-lead electrocardiography. The metabolic cart was calibrated before each test. All raw data, except for PO, were processed using a moving average over a 15-breath window [27]. VO_2peak and HRpeak were defined as the highest 15-breath average of VO_2 and HR, respectively. POpeak and the PO at each VT were defined as the last completed work rate step, plus half times the work rate increment for any 30-s block in the non-completed work rate step [28].

Determination of ventilatory thresholds

All data of the GXTs were represented in plots as described by Wasserman et al. [29] via a custom-made Matlab-script according to the preferences of both raters [Matlab R2012b, Mathworks Inc., Natick, MA, USA]. Three plots were presented to the raters: 1) VCO_2 versus VO_2, 2) the ventilatory equivalents of oxygen (Ve/VO_2) and carbon dioxide (Ve/VCO_2) versus time, and 3) respiratory exchange ratio (RER) versus time. VT1 was defined as an increase in slope of more than one in the first plot (V-slope method) [12,15,19,30], and as the first sustained rise in Ve/VO_2 without a concomitant increase in Ve/VCO_2, in the second plot (ventilatory equivalents method) [15,19,30,31]. The RER plot was used as extra reference [12,30,31].

VT2 was defined as the first sustained increase in Ve/VCO$_2$ (ventilatory equivalents method), in the plot with Ve/VO$_2$ and Ve/VCO$_2$ versus time [12,14,31], and as second increase in slope in the plot with VCO$_2$ versus VO2 [12,18]. Again, the RER plot was used as extra reference; for example for the raters to be certain that RER at VT2 was higher than RER at VT1 [12,30,31]. The raters assessed all three plots for each VT and made their final decision based on the V-slope or the ventilatory equivalents; depending on which plot most clearly showed that particular VT.

Two experienced sports physicians independently and randomly assessed the sets of graphs. They had at least 4 years of experience with VT determination in able-bodied athletes and in upper-body exercise in individuals with a disability. They were blinded to participant ID and injury level. For each determined VT, the Matlab script calculated the corresponding PO, HR, VO$_2$ and RER at that threshold. When a rater thought a VT was indeterminate, the test data for that VT were rejected. To calculate intrarater reliability, both raters assessed all tests twice (in different random order) with at least 1 week in between.

Statistical analysis

Statistical analyses were performed using SPSS (IBM SPSS Statistics 20, SPSS, Inc, Chicago, IL, USA). The data were tested for normality using Kolmogorov-Smirnov tests with Lilliefors Significance Correction and Shapiro-Wilk tests. Additionally, *z*-scores for skewness and kurtosis were calculated. To assess intrarater reliability for each VT, PO, HR and VO$_2$ at that VT were compared between the first and second session. To assess interrater reliability for each VT, PO, HR and VO$_2$ at that VT were compared between rater one and two for the first session. Systematic differences were investigated with paired-samples t-tests for the total group and Wilcoxon signed-rank tests and Mann-Whitney tests within subgroups (tetraplegia and paraplegia) as data within subgroups were not normally distributed. ICCs with 95% confidence intervals (CI) were used to measure relative agreement on group level (ICC, two-way random, absolute agreement, single measures). For clinical/training purposes, Bland-Altman plots with 95% limits of agreement (LoA) were used to measure absolute agreement on an individual level [32]. The following interpretation was used for the ICC: 0.00–0.25, little to no correlation; 0.26–0.49, low correlation; 0.50–0.69, moderate correlation; 0.70–0.89, high correlation; and 0.90–1.00, very high correlation [33]. Values were considered significant at $p < 0.05$.

Results

Due to technical problems and short periods of stopping during testing, a total of three tests were excluded. This resulted in 30 tests to be assessed (tetraplegia N=11, paraplegia N=19). These 30 tests, with two possible VTs each, were assessed during two sessions by two independent raters, resulting in a total of 240 VTs to be analyzed (30 tests x 2 VTs x 2 sessions x 2 raters). For two tests, HR data were excluded due to problems with the HR monitoring system. The test peak values are shown in table 1.

Table 1. Arm crank test peak values (N=30).

	Total group		Paraplegia		Tetraplegia	
	N	M ± SD	N	M ± SD	N	M ± SD
POpeak (W)	30	73 ± 41	19	92 ± 38	11	40 ± 20
VO_2peak (L/min)	30	1.23 ± 0.65	19	1.50 ± 0.64	11	0.76 ± 0.32
RERpeak	30	1.28 ± 0.12	19	1.30 ± 0.12	11	1.25 ± 0.12
HRpeak (bpm)	28	140 ± 30	17	158 ± 21	11	112 ± 17
Test duration (min)	30	6.5 ± 2.2	19	7.1 ± 2.0	11	5.4 ± 2.1

POpeak = peak power output, VO_2peak = peak oxygen uptake, RERpeak = peak respiratory exchange ratio, HRpeak = peak heart rate.

Determination of ventilatory thresholds

Of the 240 VTs to be analyzed, 217 VTs (90%) could be determined. Of the 23 undetermined VTs, 2 (9%) were VT1 and 21 (91%) were VT2; and 7 (30%) related to tests in individuals with paraplegia and 16 (70%) to tests in individuals with tetraplegia (figure 1). In 18 out of the 30 tests (60%), both VTs could be determined during both sessions by both raters. Fourteen of these tests were related to individuals with paraplegia. Among individuals with paraplegia, there were no differences in peak test physiological values between tests where all VTs could (N=14) and could not (N=5) be determined (*Median* (*Mdn*) ± *standard error* (*SE*): VO_2peak 1.50±0.17 L/min vs. 1.11±0.25 L/min, *p*=0.19; POpeak 98±10 W vs. 70±15 W, *p*=0.11; HRpeak 161±6.8 bpm vs. 156±5 bpm, *p*=0.20; RERpeak 1.29±0.02 vs. 1.43±0.08, *p*=0.39), However, test duration was significantly lower in tests where one or more VTs could not be determined (*Mdn* 5.1±0.6 min), compared with tests where all VTs could be determined (*Mdn* 7.6±0.5 min, *U* = 11, *z* = -2.22, *p*=0.026). Four out of five individuals, of whom one or both VTs could not be determined by one or both raters, were individuals with a high paraplegia (thoracic level 1-5).

Among individuals with tetraplegia, there were no differences in peak test physiological values and test duration between tests where all VTs could (N=4) and could not (N=7) be determined (VO_2peak 0.79±0.09 L/min vs. 0.77±0.15 L/min, *p*=0.79; POpeak 44±11 W vs. 35±8 W, *p*=0.65; HRpeak

118±14 bpm vs. 113±3 bpm, *p*=0.79; RERpeak 1.30±0.04 vs. 1.22±0.06, *p*=0.53; test duration 5.6±1.4 min vs. 4.8±0.6 min, *p*=0.53).

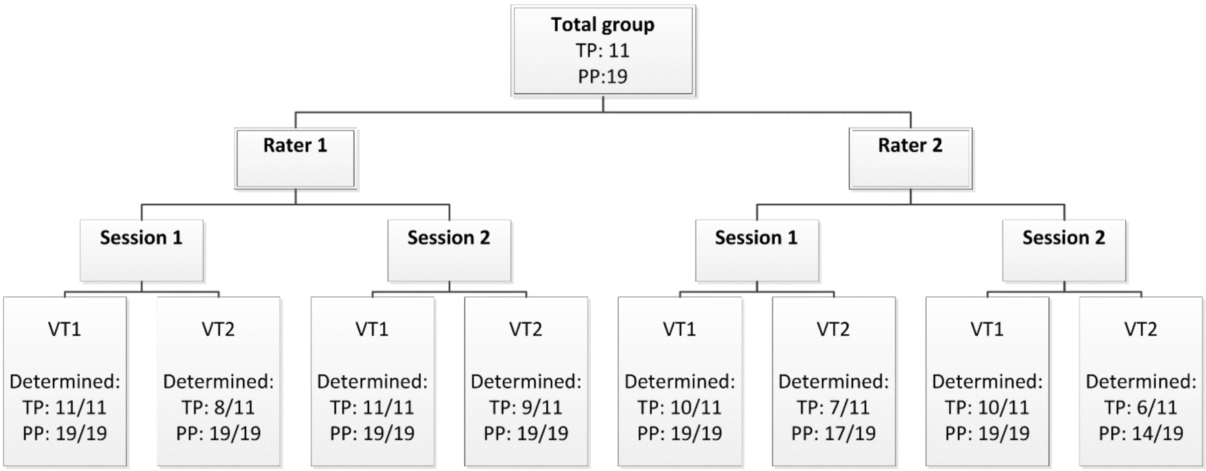


Figure 1. Flowchart of the thresholds that could be determined by two experienced raters in 30 individuals with spinal cord injury during arm crank ergometry. TP = group with tetraplegia (N=11), PP = group with paraplegia (N=19), VT = ventilatory threshold.

Intrarater reliability

For the total group and injury subgroups no systematic differences were found between session 1 and 2, except for the VO_2 at VT2 for the group with paraplegia in rater 1 (Δ Median: 0.00 L/min, Δ Mean: 0.06 L/min, T 7.0, SE 12.7, *p*=0.01). Table 2-4 show the intrarater reliability for the total group and subgroups. The relative agreement between rating sessions was very high for both raters. In subgroups, the relative agreement varied between high to very high for both raters, although small sample size and unidentifiable VTs have reduced the statistical power. This can especially be seen in table 4, where 95% CI were wide despite the high to very high ICC. Bland Altman plots showed small systematic error as represented by small mean differences. Random error was small to large as represented by the small to wide 95% LoA in figure 2 and 3. Figure 2 A, B, D, E and figure 3 A, B, D, E show the absolute agreement within raters for PO and HR, respectively.

Interrater reliability

There were no systematic differences between rater 1 and rater 2. The relative agreement between both raters was high to very high for the total group as well as for the subgroups (table 2, 3, 4). Again, due to small sample sizes and the number of excluded undetermined VTs, the number of tests in the subgroups was small. Bland-Altman plots showed small systematic error as represented by small mean differences. Random error was generally large as represented by wide 95% LoA in figure 2 and 3. Figure 2 C, F and figure 3 C, F show the absolute agreement between raters for PO and HR, respectively.

Table 2. Threshold characteristics rater 1 and rater 2 for the total group of participants during arm crank testing.

	Total group (N=30)													
	Rater 1						Rater 2							
	Session 1		Session 2		Intra		Session 1		Session 2		Intra		Inter	
	N	M ± SD	N	M ± SD	N	ICC (95%CI)	N	M ± SD	N	M ± SD	N	ICC (95%CI)	N	ICC (95%CI)
PO at VT1 (W)	30	29 ± 20	30	30 ± 21	30	0.94 (0.88 – 0.97)	29	31± 21	29	29 ± 21	29	0.98 (0.96 – 0.99)	29	0.96 (0.92 – 0.98)
% of POpeak		38 ± 15		39 ± 16				38 ± 16		36 ± 15				
PO at VT2 (W)	27	50 ± 32	28	49 ± 29	27	0.99 (0.98 – 1.00)	24	58 ± 28	20	65 ± 29	20	0.94 (0.86 – 0.98)	22	0.96 (0.90 – 0.98)
% of POpeak		64 ± 15		63 ± 15				74 ± 11		78 ± 9				
VO_2 at VT1 (L/min)	30	0.65 ± 0.25	30	0.65 ± 0.25	30	0.96 (0.92 – 0.98)	29	0.68 ± 0.24	29	0.66 ± 0.23	29	0.98 (0.95 – 0.99)	29	0.95 (0.90 – 0.98)
% of VO_2peak		59 ± 18		59 ± 17				59 ± 17		58 ± 15				
VO_2 at VT2 (L/min)	27	0.92 ± 0.50	28	0.88 ± 0.43	27	0.97 (0.93 – 0.99)	24	1.03 ± 0.43	20	1.11 ± 0.43	20	0.96 (0.91 – 0.99)	22	0.97 (0.94 – 0.99)
% of VO_2peak		74 ± 16		73 ± 16				82 ± 14		84 ± 13				
HR at VT1 (bpm)	28	105 ± 18	28	105 ± 18	28	0.94 (0.87 – 0.97)	27	104 ± 16	27	104 ± 17	27	1.00 (0.99 – 1.00)	27	0.95 (0.89 – 0.98)
% of HRpeak		77 ± 11		77 ± 11				75 ± 10		75 ± 10				
HR at VT2 (bpm)	25	119 ± 22	26	118 ± 22	25	0.99 (0.97 – 1.00)	22	123 ± 24	18	127 ± 27	18	0.94 (0.84 – 0.98)	20	0.89 (0.74 – 0.95)
% of HRpeak		84 ± 10		84 ± 10				88 ± 9		90 ± 9				

PO = power output, VO_2 = oxygen uptake, HR = heart rate, VT1 = first ventilatory threshold, VT2 = second ventilatory threshold. M = mean, SD = standard deviation, N = number of tests, ICC = intraclass correlation coefficient, 95%CI = 95% confidence intervals. Intra = the intrarater reliability, inter = the interrater reliability. The interrater reliability is based on session 1 of both raters. All ICCs were significant.

Table 3. Threshold characteristics rater 1 and rater 2 for the group with paraplegia during arm crank testing.

	Paraplegia group (N=19)													
	Rater 1						Rater 2							
	Session 1		Session 2		Intra		Session 1		Session 2		Intra		Inter	
	N	M ± SD	N	M ± SD	N	ICC (95%CI)	N	M ± SD	N	M ± SD	N	ICC (95%CI)	N	ICC (95%CI)
PO at VT1 (W)	19	37 ± 20	19	40 ± 20	19	0.91 (0.79 – 0.97)	19	38 ± 21	19	36 ± 21	19	0.97 (0.93 – 0.99)	19	0.95 (0.89 – 0.98)
% of POpeak		38 ± 13		42 ± 14				39 ± 15		37 ± 15				
PO at VT2 (W)	19	61 ± 30	19	59 ± 27	19	0.98 (0.95 – 0.99)	17	68 ± 26	14	77 ± 26	14	0.91 (0.74 – 0.97)	17	0.95 (0.87 – 0.98)
% of POpeak		64 ± 12		63 ± 13				73 ± 11		77 ± 9				
VO_2 at VT1 (L/min)	19	0.75 ± 0.24	19	0.76 ± 0.24	19	0.94 (0.86 – 0.98)	19	0.75 ± 0.25	19	0.74 ± 0.23	19	0.97 (0.92 – 0.99)	19	0.95 (0.88 – 0.98)
% of VO_2peak		53 ± 13		54 ± 14				53 ± 14		52 ± 13				
VO_2 at VT2 (L/min)	19	1.09 ± 0.48	19	1.04 ± 0.41	19	0.95 (0.88 – 0.98)*	17	1.18 ± 0.42	14	1.27 ± 0.41	14	0.94 (0.84 – 0.98)	17	0.97 (0.91 – 0.99)
% of VO_2peak		73 ± 14		71 ± 16				80 ± 14		82 ± 13				
HR at VT1 (bpm)	17	111 ± 17	17	112 ± 18	17	0.92 (0.79 – 0.97)	17	110 ± 15	17	110 ± 16	17	1.00 (0.99 – 1.00)	17	0.95 (0.88 – 0.98)
% of HRpeak		71 ± 8		72 ± 10				70 ± 7		70 ± 7				
HR at VT2 (bpm)	17	127 ± 20	17	126 ± 20	17	0.98 (0.95 – 0.99)	15	134 ± 20	12	139 ± 23	12	0.88 (0.66 – 0.97)	15	0.82 (0.55 – 0.94)
% of HRpeak		81 ± 9		80 ± 10				85 ± 10		87 ± 10				

PO = power output, VO_2 = oxygen uptake, HR = heart rate, VT1 = first ventilatory threshold, VT2 = second ventilatory threshold. M = mean, SD = standard deviation, N = number of tests, ICC = intraclass correlation coefficient, 95%CI = 95% confidence intervals. Intra = the intrarater reliability, inter = the interrater reliability. The interrater reliability is based on session 1 of both raters. All ICCs were significant. * = outcome of the Wilcoxon signed-rank test for systematic differences, significant at $p < 0.05$.

Table 4. Threshold characteristics rater 1 and rater 2 for the group with tetraplegia during arm crank testing.

Tetraplegia group (N=11)														
	Rater 1						Rater 2							
	Session 1		Session 2		Intra		Session 1		Session 2		Intra		Inter	
	N	M ± SD	N	M ± SD	N	ICC (95%CI)	N	M ± SD	N	M ± SD	N	ICC (95%CI)	N	ICC (95%CI)
PO at VT1 (W)	11	16 ± 13	11	15 ± 12	11	0.96 (0.86 – 0.99)	10	17 ± 13	10	17 ± 13	10	0.99 (0.97 – 1.00)	10	0.93 (0.73 – 0.98)
% of POpeak		37 ± 19		32 ± 18				36 ± 19		35 ± 17				
PO at VT2 (W)	8	25 ± 17	9	28 ± 18	8	1.00 (0.99 – 1.00)	7	33 ± 11	6	37 ± 14	6	0.89 (0.49 – 0.98)	5	0.85 (0.15 – 0.98)
% of POpeak		63 ± 20		64 ± 19				74 ± 11		79 ± 8				
VO_2 at VT1 (L/min)	11	0.48 ± 0.19	11	0.47 ± 0.16	11	0.97 (0.90 – 0.99)	10	0.53 ± 0.15	10	0.52 ± 0.16	10	0.99 (0.97 – 1.00)	10	0.89 (0.61 – 0.97)
% of VO_2peak		69 ± 20		67 ± 19				69 ± 18		67 ± 16				
VO_2 at VT2 (L/min)	8	0.51 ± 0.21	9	0.55 ± 0.25	8	0.98 (0.89 – 1.00)	7	0.68 ± 0.18	6	0.73 ± 0.18	6	0.98 (0.88 – 1.00)	5	0.88 (0.21 – 0.99)
% of VO_2peak		77 ± 21		76 ± 18				87 ± 12		89 ± 14				
HR at VT1 (bpm)	11	96 ± 14	11	95 ± 13	11	0.93 (0.78 – 0.98)	10	94 ± 13	10	94 ± 13	10	0.99 (0.97 – 1.00)	10	0.90 (0.66 – 0.98)
% of HRpeak		86 ± 7		85 ± 8				84 ± 7		84 ± 7				
HR at VT2 (bpm)	8	102 ± 16	9	102 ± 16	8	1.00 (0.98 – 1.00)	7	100 ± 15	6	102 ± 19	6	0.95 (0.75 – 0.99)	5	0.93 (0.54 – 0.99)
% of HRpeak		91 ± 8		91 ± 7				93 ± 5		95 ± 3				

PO = power output, VO_2 = oxygen uptake, HR = heart rate, VT1 = first ventilatory threshold, VT2 = second ventilatory threshold. M = mean, SD = standard deviation, N = number of tests, ICC = intraclass correlation coefficient, 95%CI = 95% confidence intervals. Intra = the intrarater reliability, inter = the interrater reliability. The interrater reliability is based on session 1 of both raters. All ICCs were significant.

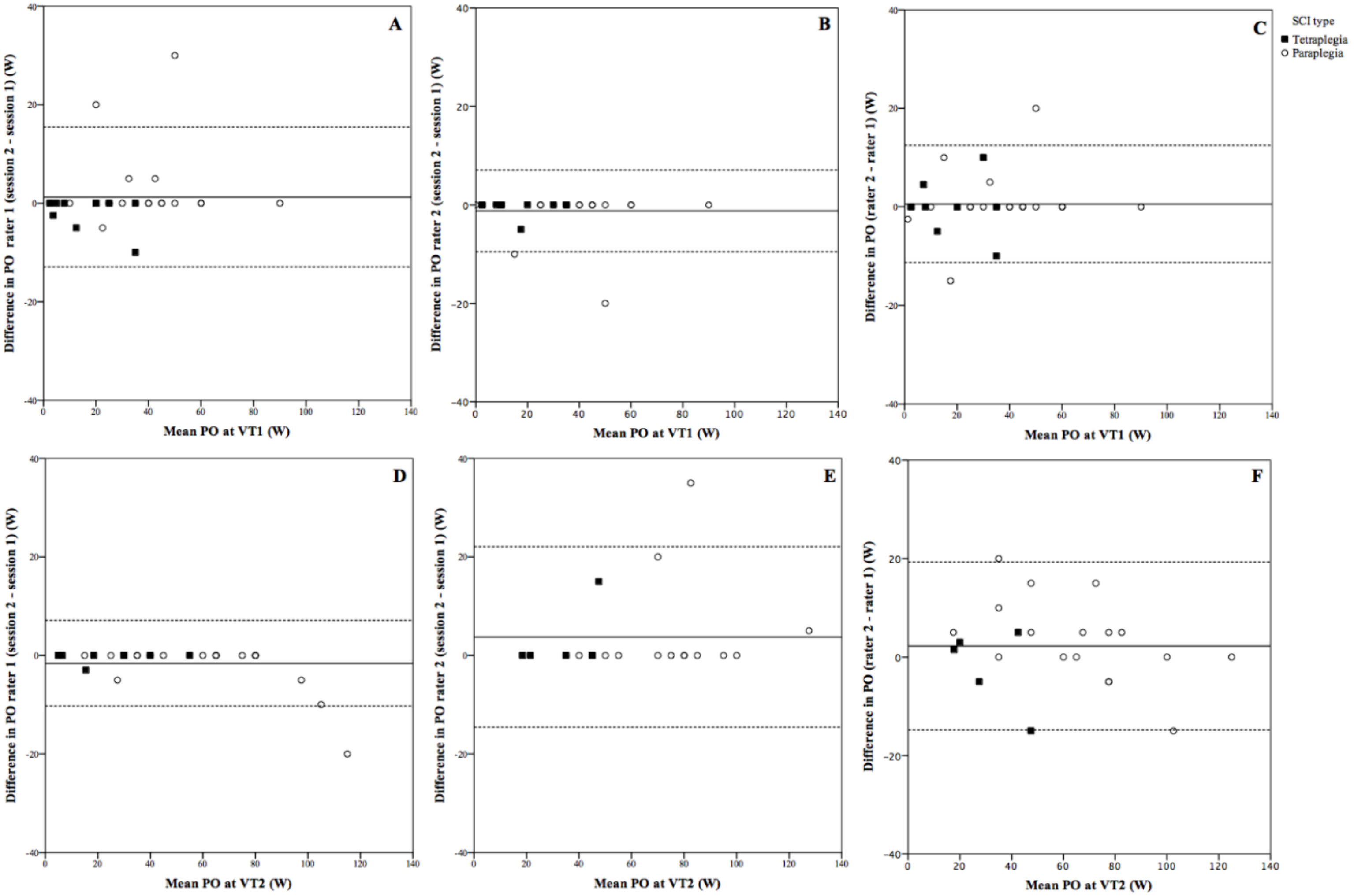


Figure 2. Bland-Altman plot representing the absolute agreement of the power output (PO) within raters and between raters. Solid line represents the mean bias (systematic error), dotted lines represent mean ± 2SD (95% LoA; random error). Circles and squares represent individuals with paraplegia and tetraplegia, respectively A. Intrarater reliability rater one at VT1. B. Intrarater reliability rater two at VT1. C. Interrater reliability at VT1. D. Intrarater reliability rater one at VT2. E. Intrarater reliability rater two at VT2. F. Interrater reliability at VT2.

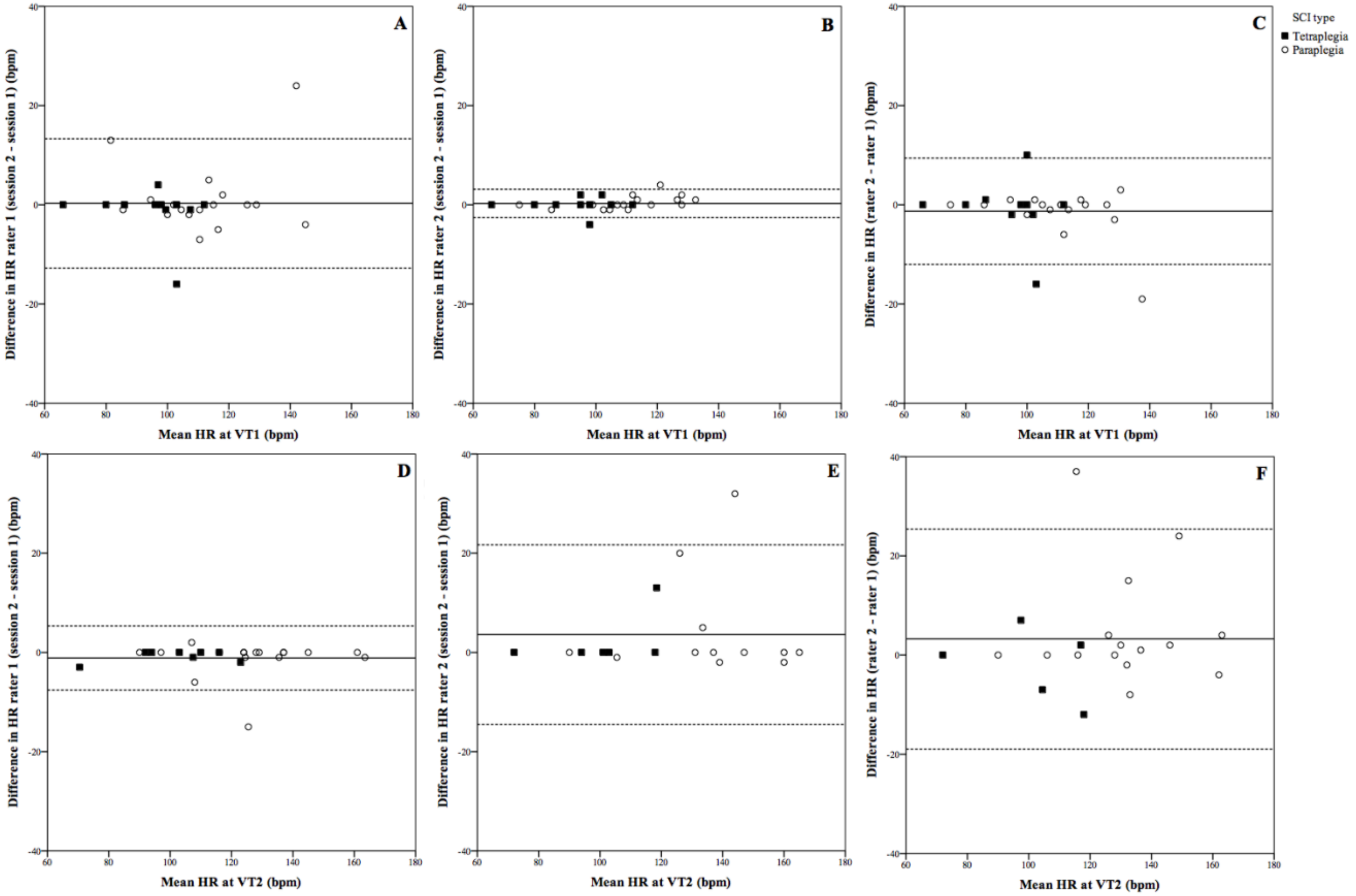


Figure 3. Bland-Altman plot representing the absolute agreement of the heart rate (HR) within raters and between raters. Solid line represents the mean bias (systematic error), dotted lines represent mean ± 2SD (95% LoA; random error). Circles and squares represent individuals with paraplegia and tetraplegia, respectively A. Intrarater reliability rater one at VT1. B. Intrarater reliability rater two at VT1. C. Interrater reliability at VT1. D. Intrarater reliability rater one at VT2. E. Intrarater reliability rater two at VT2. F. Interrater reliability at VT2.

Discussion

Of all VTs to be analyzed, 90% could be determined. Of the undetermined VTs, most were VT2 and related to individuals with tetraplegia. In 60% of the tests, both thresholds could be determined during both sessions by both raters. For the successfully determined VTs, the relative intrarater reliability was very high whereas random error ranged from small to large within raters and among outcome measures. The relative interrater reliability was high to very high with a low absolute agreement due to large random error.

The participants of the present study were recreationally active individuals with SCI. For physical fitness, the participants with paraplegia scored "good" for VO_2peak compared with the average untrained population with paraplegia, based on the study by Simmons et al. [34]. The participants with tetraplegia scored "average" for VO_2peak compared with the average untrained population with tetraplegia [34]. We might conclude that the population in the present study has a physical fitness somewhat above average, compared with the untrained population with SCI.

For the individuals with paraplegia, VO_2 at VT1 and VT2 was on average 53% and 76% of VO_2peak, respectively. In previous literature across all test modes and fitness levels, VT1 has been reported as occurring at between 56% and 77% of VO_2peak in individuals with paraplegia [19–21,35,36], whereas VT2 has been reported at 78% of VO_2peak [19]. Possible small differences between the present study and previous literature might be explained by mode of exercise and training status of the participants; physical fitness of the studied population in previous literature is generally higher than in the present study (VO_2peak on average 1.9 L/min in previous literature vs 1.5 L/min in present study) [19,35,36].

For individuals with tetraplegia, VO_2 at VT1 and VT2 was on average 68% and 81% of VO_2peak, respectively. In previous literature across all test modes and fitness levels, VT1 has been reported as occurring at between 63% and 87% of VO_2peak in individuals with tetraplegia [19–22]. Whereas VT2 has been reported at 75% of VO_2peak [19]. Overall, VTs of this subgroup are comparable with those reported in literature.

Ninety percent of VTs could be determined in the present study. This is comparable with literature with able-bodied participants [15,16] and athletes with SCI [19,20]. Most of the VTs that could not be determined were VT2s and related to tests in individuals with tetraplegia. Leicht et al. [19] found comparable results; two out of 19 VT1s (11%) could not be determined, both in athletes with tetraplegia, and 5 out of 19 VT2s (26%) could not be determined, of which 3 belonged to athletes with tetraplegia. Leicht et al. [19] explained their findings by lower absolute ventilatory responses in individuals with SCI, and tetraplegia specifically, resulting in a narrower range of ventilatory values compared with able-bodied athletes. In the present study there was no significant difference in VO_2peak between individuals with tetraplegia whose VTs could be determined compared to those whose VTs could not be determined. However, although not significant, which is potentially due to small sample sizes, it can be seen that for both persons with tetraplegia and paraplegia POpeak and test duration were generally lower in tests where one or both VTs could not be determined. This also might explain the finding that the proportion of undetermined VTs was higher in individuals with tetraplegia compared with individuals with paraplegia. This is supported by a recent study, where in 32% of tests, VT1 could not be determined in individuals with tetraplegia [21]. They explain their findings by lower peak cardiorespiratory responses and lower test duration for those individuals, compared with tests where VT1 could be determined. Another reason for not being able to determine VTs in untrained individuals with SCI, especially at higher intensity (VT2), might be premature termination of the test due to peripheral fatigue. In the present study three out of twelve individuals where one or both VTs could not be determined, stated that fatigue in the arms was the reason to stop the test.

For the VTs that could be determined, relative agreement for the total group within and between raters was high to very high. The SCI subgroups results might be hard to interpret, as these groups were small. The results are comparable to previous literature with able-bodied participants and wheelchair athletes, where an intrarater reliability of 0.94–0.97 was reported [17,18] and an interrater reliability of 0.81–0.95 [18,20,22]. The absolute agreement varied between outcome measures. For some measures, such as HR at VT2 between raters, the random error was large, as depicted in figure 3F. This figure also shows a certain degree of heteroscedasticity: random error appears to be larger for individuals with a higher HR at VT2, i.e., those with a paraplegia.

On group level the agreement is high to very high, but on individual level there might be large differences between rating sessions or raters, which has large implications for the correct prescription of exercise intensity of that individual. This suggests that relative agreement of VT determination should be interpreted with caution, not only in the present study, but also in previous literature, as the absolute agreement was unfortunately often not reported.

Practical Applications

On group level the results of the present study are positive. For the majority of tests, the VTs could be determined and relative agreement within and between raters was high to very high. Nevertheless, for 7 out of 11 tests of individuals with tetraplegia, one or both raters could not determine one or both VTs. This seemed to coincide with short test duration. Despite the extensive experience of the testers with testing in individuals with SCI, it was difficult to select a protocol resulting in test duration between 8 and 12 min. It must be emphasized that individualized protocol selection is important for individuals with SCI. However, optimal protocol selection is comprehensive as cardiorespiratory fitness in individuals with SCI is based on a lot of factors, such as lesion level, sex, BMI and training status [24,25,34]. As such, tests with a duration less than 8 min are common in clinical practice and are not specific for the present study [21].

As known, training intensity based on HRpeak or HR reserve might not be applicable to individuals with a lesion level above thoracic spinal nerve 6 due to the altered sympathetic response to exercise [37], this is also shown in the present study, as HRpeak was low in individuals with tetraplegia (table 1). The present study shows that it is sometimes impossible to determine VTs in this group, which makes training based on training zones challenging as well. Other methods to prescribe exercise intensity might provide better precision, such as training based on ratings of perceived exertion and/or %POpeak [38]. In the present study it was not investigated whether exercise intensity prescription based on VTs is favorable to prescription based on RPE, %HRR, or %POpeak in terms of improvements in cardiorespiratory fitness. This should be further investigated in future research. Moreover, as the large random error within and between raters suggests, training schemes based on VTs should be clinically evaluated on individual level. For example, a talk test may be used to evaluate whether the intensity is either too high or too low [39]. If this appears to be the case, VT determination should be critically re-evaluated by one or more experts in order to prevent over- or undertraining in that individual. In addition, the low absolute agreement between raters suggest that during a longitudinal follow-up with several GXTs within an individual, it would be advised to identify the VTs by the same rater.

Study Limitations

Although the sample size of the present study was equal to or higher than the sample size in comparable studies [17,18,20,22], the sample size of the subgroups, especially for individuals with tetraplegia, was small. Therefore the statistical power was reduced, which makes interpretation of the ICCs for subgroups less reliable. It must be noted, however, that large sample sizes in rehabilitation populations are difficult to obtain. Another aspect that was not investigated in the present study, is the test-retest reliability across days of the GXT itself. It might be interesting for future studies to investigate reliability of VTs during repeated GXTs, as the variability of VTs between tests within individuals is unknown for this population.

Conclusions

Ninety percent of VTs could be determined. Most of the VTs that could not be determined were VT2s and related to tests in individuals with tetraplegia. For the VTs that could be determined, the relative intrarater reliability was very high with small to large random error. The relative interrater reliability was high to very high with large random error. Although these results are positive on group level and show that determination of VTs might be a promising method to define training intensity for the majority of the tested recreationally active individuals with SCI, it should be noted that a critical evaluation of the VTs is necessary and other exercise intensity prescription methods should be considered when one or both VTs cannot be determined.

Conflicts of interest

The authors have no conflicts of interest to declare.

Acknowledgments

The authors would like to thank Marcel Post, Center of Excellence in Rehabilitation Medicine, De Hoogstraat Rehabilitation, Utrecht, the Netherlands for his support with the manuscript, and Ilse Blokland, Research and Development, Heliomare Rehabilitation Center, Wijk aan Zee, the Netherlands for her support with Matlab programming.

Funding

This study was funded by the National Institutes of Health National Institute of Neurological Disorders and Stroke (grant no. NS083064), Miami Project to Cure Paralysis, HandicapNL, Stichting Mitialto, Stichting Beatrixoord Noord-Nederland, University Medical Center Groningen, Heliomare Rehabilitation Center and Stichting Handbike Events.

References

1. 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(11):642-652.

2. Garshick E, Kelley A, Cohen SA, et al. A prospective assessment of mortality in chronic spinal cord injury. *Spinal Cord*. 2005;43(7):408-416.

3. Nightingale TE, Metcalfe RS, Vollaard NB, Bilzon JL. Exercise guidelines to promote cardiometabolic health in spinal cord injured humans : Time to raise the intensity? *Arch Phys Med Rehabil*. 2017;98(8):1693-1704.

4. Valent LJM, Dallmeijer AJ, Houdijk JHP, Slootman JR, Janssen TWJ, van der Woude LHV. Effects of hand cycle training on wheelchair capacity during clinical rehabilitation in persons with a spinal cord injury. *Disabil Rehabil.* 2010;32(26):2191-2200.

5. Valent LJM, Dallmeijer AJ, Houdijk H, et al. Effects of hand cycle training on physical capacity in individuals with tetraplegia: a clinical trial. *Phys Ther*. 2009;89(10):1051-1060.

6. Valent LJM, Dallmeijer AJ, Houdijk H, Slootman HJ, Post MWM, van der Woude LHV. Influence of hand cycling on physical capacity in the rehabilitation of persons with a spinal cord injury: A longitudinal cohort study. *Arch Phys Med Rehabil*. 2008;89(6):1016-1022.

7. De Groot S, Postma K, van Vliet L, Timmermans R, Valent LJM. Mountain time trial in handcycling: Exercise intensity and predictors of race time in people with spinal cord injury. *Spinal Cord*. 2014;52(6):455-461.

8. Hoekstra SP, Valent LJM, Gobets D, van der Woude LHV, de Groot S. Effects of four-month handbike training under free-living conditions on physical fitness and health in wheelchair users. *Disabil Rehabil*. 2017;39(16):1581-1588.

9. Wolpern AE, Burgos DJ, Janot JM, Dalleck LC. Is a threshold-based model a superior method to the relative percent concept for establishing individual exercise intensity? A randomized controlled trial. *BMC Sports Sci Med Rehabil*. 2015;7(1):16.

10. Meyer T, Gabriel HHW, Kindermann W. Is determination of exercise intensities as percentage of VO_2max or HRmax adequate? *Med Sci Sport Exerc*. 1999;31(9):1342-1345.

11. Kindermann W, Simon G, Keul J. The significance of the aerobic-anaerobic transition for the determination of work load intensities during endurance training. *Eur J Appl Physiol*. 1979;34(42):25-34.

12. Meyer T, Lucía A, Earnest CP, Kindermann W. A conceptual framework for performance diagnosis and training prescription from submaximal gas exchange parameters - Theory and application. *Int J Sport Med*. 2005;26(1):38-48.

13. Seiler KS, Kjerland GØ. Quantifying training intensity distribution in elite endurance athletes: Is there evidence for an “optimal” distribution? *Scand J Med Sci Sport*. 2006;16(1):49-56.

14. Mezzani A, Hamm LF, Jones AM, et al. Aerobic exercise intensity assessment and prescription in cardiac rehabilitation. *J Cardiopulm Rehabil Prev*. 2012;32(6):327-350.

15. Gaskill SE, Ruby BC, Walker AJ, Sanchez OA, Serfass RC, Leon AS. Validity and reliability of combining three methods to determine ventilatory threshold. *Med Sci Sport Exerc*. 2001;33(11):1841-1848.

16. Shimizu M, Myers J, Buchanan N, et al. The ventilatory threshold: Method, protocol, and evaluator agreement. *Am Heart J*. 1991;122(2):509-516.

17. Gladden LB, Yates JW, Stremel RW, Stamford BA. Gas exchange and lactate anaerobic thresholds: inter- and intraevaluator agreement. *J Appl Physiol*. 1985;58:2082-2089.

18. Aunola S, Rusko H. Reproducibility of aerobic and anaerobic thresholds in 20-50 year old men. *Eur J Appl Physiol Occup Physiol*. 1984;53(3):260-266.

19. Leicht CA, Griggs KE, Lavin J, Tolfrey K, Goosey-Tolfrey VL. Blood lactate and ventilatory thresholds in wheelchair athletes with tetraplegia and paraplegia. *Eur J Appl Physiol*. 2014;114(8):1635-1643.

20. Coutts KD, McKenzie DC. Ventilatory thresholds during wheelchair exercise in individuals with spinal cord injuries. *Paraplegia*. 1995;33(7):419-422

21. Au JS, Sithamparapillai A, Currie KD, Krassioukov AV, MacDonald MJ, Hicks AL. Assessing ventilatory threshold in individuals with motor-complete spinal cord injury. *Arch Phys Med Rehabil*. 2018;99(10):1991-1997.

22. Bhambhani YN, Burnham RS, Wheeler GD, Eriksson P, Holland LJ, Steadward RD. Ventilatory threshold during wheelchair exercise in untrained and endurance-trained subjects with quadriplegia. *Adapt Phys Act Q*. 1995;12(4):333-343.

23. Maher JL, Cowan RE. Comparison of 1- versus 3-minute stage duration during arm ergometry in individuals with spinal cord injury. *Arch Phys Med Rehabil*. 2016;97(11):1895-1900.

24. Kouwijzer I, Valent LJM, Osterthun R, van der Woude LHV, de Groot S, HandbikeBattle Group. Peak power output in handcycling of individuals with a chronic spinal cord injury: predictive modeling, validation and reference values. *Disabil Rehabil*. 2020;42(3):400-409.

25. Janssen TWJ, van Oers CAJ, Hollander PA, Veeger DHEJ, van der Woude LHV. Isometric strength, sprint power, and aerobic power in individuals with a spinal cord injury. *Med Sci Sports Exerc*. 1993;25(7):863-870.

26. Buchfuhrer MJ, Hansen JE, Robinson TE, Sue DY, Wasserman K, Whipp BJ. Optimizing the exercise protocol for cardiopulmonary assessment. *J Appl Physiol Respir Environ Exerc Physiol*. 1983;55(5):1558-1564.

27. Robergs RA, Dwyer D, Astorino T. Recommendations for improved data processing from expired gas analysis indirect calorimetry. *Sport Med*. 2010;40(2):95-111.

28. Kuipers H, Verstappen FTJ, Keizer P, Geurten P, van Kranenburg G. Variability of aerobic performance in the laboratory and its physiologic correlates. *Int J Sports Med*. 1985;6(4):197-201.

29. Wasserman K, Hansen JE, Sue DY, Stringer WW, Sietsema, KE, Sun X-G, Whipp BJ. *Principles of Exercise Testing and Interpretation 5th Edition*. Philadelphia, USA. Lippincott Williams & Wilkins; 2012.

30. Beaver WL, Wasserman K, Whipp BJ. A new method for detecting anaerobic threshold by gas exchange. *J Appl Physiol*. 1986;60(6):2020-2027.

31. Binder RK, Wonisch M, Corra U, et al. Methodological approach to the first and second lactate threshold in incremental cardiopulmonary exercise testing. *Eur J Cardiovasc Prev Rehabil*. 2008;15(6):726-734.

32. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*. 1986;1(8476):307-310.

33. Munro BH. *Statistical methods for health care research*. 7th edition. Philadelphia, USA. Lippincott Williams & Wilkins; 2004. p. 239 - 258.

34. Simmons OL, Kressler J, Nash MS. Reference fitness values in the untrained spinal cord injury population. *Arch Phys Med Rehabil*. 2014;95(12):2272-2278.

35. Lovell D, Shields D, Beck B, Cuneo R, McLellan C. The aerobic performance of trained and untrained handcyclists with spinal cord injury. *Eur J Appl Physiol*. 2012;112(9):3431-3437.

36. Schneider DA, Sedlock DA, Gass E, Gass G. VO_2peak and the gas-exchange anaerobic threshold during

incremental arm cranking in able-bodied and paraplegic men. *Eur J Appl Physiol Occup Physiol*. 1999;80(4):292-297.

37. Valent LJM, Dallmeijer AJ, Houdijk H, et al. The individual relationship between heart rate and oxygen uptake in people with a tetraplegia during exercise. *Spinal Cord*. 2007;45(1):104-111.

38. Van der Scheer JW, Hutchinson MJ, Paulson T, Martin Ginis KA, Goosey-Tolfrey VL. Reliability and validity of subjective measures of aerobic intensity in adults with spinal cord injury: A systematic review. *PM R*. 2018;10(2):194-207.

39. Cowan R, Ginnity K, Kressler J, Nash M. Assessment of the talk test and rating of perceived exertion for exercise intensity prescription in persons with paraplegia. *Top Spinal Cord Inj Rehabil*. 2012;18(3):212-219.

Chapter 4

The influence of protocol design on the identification of ventilatory thresholds and the attainment of peak physiological responses during synchronous arm crank ergometry in able-bodied participants

Kouwijzer I*, Valize M*, Valent LJM, Grandjean Perrenod Comtesse P, van der Woude LHV, de Groot S. The influence of protocol design on the identification of ventilatory thresholds and the attainment of peak physiological responses during synchronous arm crank ergometry in able-bodied participants. European Journal of Applied Physiology 2019; 119(10):2275-2286.

*Both authors contributed equally.

Abstract

Purpose: To examine the effects of stage duration on power output (PO), oxygen uptake (VO_2) and heart rate (HR) at peak level and ventilatory thresholds during synchronous arm crank ergometry.

Methods: Nineteen healthy participants completed a ramp, 1-min stepwise, and 3-min stepwise graded arm crank exercise test. PO, VO_2 and HR at first and second ventilatory threshold (VT1, VT2), and peak level were compared among the protocols: a repeated measures analysis of variance was performed to test for systematic differences, while intraclass correlation coefficients (ICC) and Bland-Altman plots were calculated to determine relative and absolute agreement.

Results: Systematic differences among the protocols were found for PO at VT1, VT2 and peak level. At peak level, PO differed significantly among all protocols (ramp: 115±37 W; 1-min stepwise: 108±34 W; 3-min stepwise: 94±31 W, $p \leq 0.01$). No systematic differences for HR or VO_2 were found among the protocols. VT1 and VT2 were identified at 52% and 74% of VO_2peak, respectively. The relative agreement among protocols varied (ICC: 0.02 – 0.97) while absolute agreement was low with small-to-large systematic error and large random error.

Conclusions: PO at VTs and peak level was significantly higher in short stage protocols compared with the 3-min stepwise protocol, whereas HR and VO_2 showed no differences. Therefore, training zones based on PO determined in short stage protocols might give an overestimation. Moreover, due to large random error in HR at VTs between the protocols, it is recommended that different protocols should not be used interchangeably within individuals.

Keywords

Protocol type, graded exercise test, physical capacity, ventilatory thresholds, training

Introduction

Handcycling is a rapidly growing sport for people with disabilities worldwide, especially in persons with spinal cord injury (SCI), muscular disease or leg amputation [1]. People have turned to handcycling as a means to improve their physical capacity during or after rehabilitation [2,3]. To promote handcycling as an exercise mode in the Netherlands, an annual handcycle event called the HandbikeBattle is organised in Austria since 2013 [4].

To become physically fit or to prepare optimally for such an event, a valid and reliable individualised training scheme is necessary. The intensity of the aerobic endurance training in this training scheme is based on the cardiorespiratory fitness of the individual, measured during a graded exercise test (GXT). Results of the GXT are used to develop individualised schemes based on percentages of peak power output (POpeak) or peak heart rate (HRpeak), or based on training zones delineated by power output (PO) or heart rate (HR) at ventilatory thresholds (VTs) [5–8]. These VTs provide boundaries to set individualized training zones: zone 1 at low intensity (below the first ventilatory threshold (VT1)), zone 2 at moderate intensity (between VT1 and the second ventilatory threshold (VT2)), and zone 3 at high intensity (above VT2) [6,8]. Over the years, several GXT protocol designs with varying stage durations have been employed. For example, workload increases at set intensities following a defined interval of time as a series of “steps” in a stepwise protocol, or workload increases in a smooth linear way in a ramp protocol [9–17]. It is, however, not entirely known what the effects are of these different types of protocols and stage durations on outcome measures such as oxygen uptake (VO_2), PO and HR, at both peak exercise and VTs during synchronous arm exercise.

In asynchronous arm cranking, two studies investigated effects of stage duration on peak physiological responses. Smith et al. compared a 2-min stepwise protocol with a ramp protocol in able-bodied participants (N=14), whereas Maher et al. compared a 1-min with a 3-min stepwise protocol in individuals with SCI (N=38) [9,11]. The protocols were designed in such a way that patterns of work rate increase, external work, and test duration were comparable between protocols. Both studies found no significant differences in VO_2peak, HRpeak and POpeak between the different protocols. Additionally, Smith et al. compared two ramp protocols with different ramp slopes. While VO_2peak and HRpeak were unaffected, they found a significantly higher POpeak and shorter test duration in the protocol with a steeper ramp slope [10].

In able-bodied cycling, effects from protocols with different stage duration are studied widely. In general, POpeak was higher in ramp protocols or protocols with short-stage duration, compared with protocols with longer stage duration [12–18]. In the tests with longer stage duration, total test duration was also longer. Peak oxygen uptake (VO_2peak) was not significantly different between protocols [12,13,15–17], or was higher in protocols with longer stage duration [14]. HRpeak was not significantly different between protocols [12,14,17], or was higher in protocols with longer stage duration [13,15,16]. Bentley et al. found that VO_2 and HR at VT1 were not significantly different between a 1-min and a 3-min stepwise protocol, whereas PO at VT1 was significantly higher in the 1-min stepwise protocol [12].

The effects of these different types of protocols and stage durations on VO_2, PO and HR, at both peak level and VTs have previously not been studied with synchronous arm exercise. Traditionally, protocols with longer stage duration are executed to determine submaximal responses and the position of thresholds. However in the last years, supported by technological innovations, ramp protocols became popular, also to detect VTs [19]. The advantage of ramp protocols is that the work changes over time are not affected by protocol steps, leading to linear physiological responses [19–21]. The consequence is, however, that the VO_2 response is specific to the non-steady-state character of the protocol. Typically, the VO_2 response shows a lag to the metabolic demand (i.e., mean response time) [20]. The measured VO_2 at any work rate will underestimate the steady-state VO_2 at that work rate [10,22]. Since VTs are often used to set up training schemes [6], it is of importance to know whether the position of VTs is affected by the used test protocol with corresponding test duration and step size. Therefore, the aim of this study was to examine the effects of stage duration with a ramp protocol, 1-min stepwise protocol and 3-min stepwise protocol on PO, VO_2 and HR at both peak level and at VT1 and VT2 during synchronous arm crank ergometry. We hypothesized that VO_2 and HR at VTs and peak level will not be affected by stage duration, whereas PO at VTs and peak level will be higher within short stage protocols, compared with the 3-min stepwise protocol.

Materials and methods

Participants

Nineteen able-bodied individuals were recruited to participate in the study: nine men / ten women, age (mean ± standard deviation) 30 ± 10 years with range: 21-58 years, body mass: 71.6 ± 9.9 kg, height: 1.78 ± 0.07 m. All participants were healthy and physically active. They participated recreationally in sports such as fitness, soccer and running with an average of four hours a week. They were non-specifically arm trained and had no experience with GXT on an arm crank ergometer. They had no restrictions or injuries of the upper extremities, and did not suffer from chronic diseases, such as heart or lung disease, diabetes or obesity. Before the start of the test, participants were medically screened using the Physical Activity Readiness Questionnaire (PAR-Q) [23]. Informed consent was obtained from all individual participants included in the study. The study was approved by the Ethics Committee of the Center for Human Movement Sciences, University Medical Center Groningen, University of Groningen, the Netherlands.

Test procedure

All participants performed three GXTs in synchronous cranking mode: one with a ramp protocol, one with a 1-min stepwise protocol, and one with a 3-min stepwise protocol. The order of the tests was counter-balanced and each test was separated by at least 3 days with a maximum of 7 days. The three tests were executed at the same time of the day within participants, but varied between participants. Participants were required to abstain from alcohol, caffeine and smoking 12 h before testing and from strenuous physical activity for 24 h.

All tests were conducted on an electrically braked arm crank ergometer (Lode Angio, Groningen, the Netherlands). The ergometer was wall-mounted using a height-adjustable bracket. Participants were seated so that the axis of rotation of the arm crank was at the same height as the axis of rotation of the shoulder joint, and positioned at a comfortable distance from the ergometer allowing for a slight bend (15-20°) of the participant's elbow at the furthest point of the range of movement [9,24]. Participants were required to sit back firmly in the chair to maintain a standardized position, and were instructed to keep their feet in front of them at shoulder width and flat on the floor throughout each test.

For the ramp protocol and the 1-min stepwise protocol the aim was to develop a protocol with a test duration between 8 and 12 min [25], and a longer test duration with at least six steps (18 min) for the 3-min stepwise protocol [13,18]. The starting workload and step size or ramp slope were based on pilot experiments and previous literature [9,26,27]. The pilot experiments were conducted in individuals comparable to the studied population and pilot results were not included in the present study. All tests started with a resting period of 2 min, followed by a warm-up of 5 min on 20 W. For male participants the test protocols were as follows: ramp: start at 0 W with increments 1 W/5 s (i.e., equivalent to 12 W / min); 1-min stepwise: start at 10 W with increments 12 W / min; 3-min stepwise: start at 10 W with increments 20 W / 3 min. For female participants the test protocols were as follows: ramp: start at 0 W with increments 1 W/7.5 s (i.e., equivalent to 8 W / min); 1-min stepwise: start at 10 W with increments 8 W / min; 3-min stepwise: start at 10 W with increments 14 W / 3 min. After all tests, a cool-down of 5 min was performed on 20 W. During the test, participants were instructed to maintain a crank rate of 60-80 revolutions per minute (RPM). Criteria to stop the test were volitional exhaustion or failure in keeping a constant cadence above 60 RPM. Verbal encouragements were given towards the end of the test. At the end of the test, ratings of perceived exertion (RPE) was recorded using the 10-point Borg scale [28].

Determination of peak physiological responses

Respiratory data, including VO_2, carbon dioxide production (VCO_2), minute ventilation (VE), and respiratory exchange ratio (RER), were collected and analysed per 10 s by mixing-chamber technique using the Cortex (Cortex, CORTEX Biophysik GmbH, Germany). The equipment was calibrated before each test. Criteria for a peak test were: RPE ≥ 8 at the end of the test and RERpeak ≥1.10 [19]. For all three protocols, VO_2peak, VCO_2peak, VE and RERpeak were defined as the highest 30-s average value of VO_2, VCO_2, VE and RER, respectively. HR was recorded continuously from rest through recovery using a wireless chest strap monitor (Polar T31, Finland). HRpeak was defined as the highest 10-s average value achieved. In the 1-min stepwise protocol POpeak was defined as the last completed PO step, plus ½ times the PO increment for each 30-s block in the non-completed PO step. In the 3-min stepwise protocol, POpeak was defined as the last completed PO step, plus 1/6 times the PO increment for each 30-s block in the non-completed PO step [29]. In the ramp protocol, POpeak was defined as the highest 10-s PO achieved at the end of the test. Additionally, total accumulated work done (TWD in kJ) was calculated as described in previous literature [9,17].

Determination of ventilatory thresholds

For the three protocols, all data were represented in plots as described by Wasserman et al. [30]. A combination of two plots was examined to determine VT1: 1) VCO_2 versus VO_2, 2) the ventilatory equivalents of oxygen (VE/VO_2) and carbon dioxide (VE/VCO_2) versus time. VT1 was defined as an increase in slope of more than 1 in the first plot (V-slope method) [6,31–33], and / or as the first sustained rise in VE/VO_2 without a concomitant increase in VE/VCO_2 in the second plot (ventilatory equivalents method) [31–35].

A combination of three plots was examined to determine VT2: 1) VE versus VCO_2, 2) VE/VO_2 and VE/VCO_2 versus time, 3) VCO_2 versus VO_2. VT2 was defined as the inflection in the VE versus VCO_2 slope in the first plot [6,35,36], and / or the first systematic increase in VE/VCO_2 (ventilatory equivalents method) in the second plot [6,35,36], and / or as a second increase in slope in the third plot with VCO_2 versus VO_2 [6,37].

Two trained researchers independently examined the plots visually to determine both VTs. Thereafter results were compared. The interrater reliability (intraclass correlation coefficient (ICC)) of the determined VTs was 0.93 (95% CI 0.84 – 0.97) for VT1 (based on N = 48) and 0.94 (95% CI 0.90 – 0.97) for VT2 (based on N = 48). On average, there was a 4.1% and 0.6% difference between raters for

VT1 and VT2, respectively. When the raters did not have the exact same point in time for a VT (N = 65), they examined the plots together to come to a mutually agreed VT. This value was then used for further analysis.

Thereafter, VO_2 and HR at the VTs were determined as the 10-s value on that point in time. PO at the VTs was determined as follows: the last completed PO step, plus ½ times the PO increment for each 30-s block in the non-completed PO step for the 1-min stepwise protocol; the last completed PO step, plus 1/6 times the PO increment for each 30-s block in the non-completed PO step for the 3-min stepwise protocol; and the PO achieved at that specific point in time for the ramp protocol. TWD at the VTs was calculated as the accumulated work (PO (W) x time (s)) at that point in time. The relative values of VO_2, HR, PO and TWD at both VTs were calculated as the absolute value at that VT divided by the peak value of that respective outcome measure (i.e., VO_2peak, HRpeak, POpeak and TWD at peak).

Statistical analyses

Data were analysed using IBM SPSS Statistics 24 (IBM SPSS Statistics 24, SPSS, Inc, Chicago, IL, USA). The data were tested for normality using the Kolmogorov-Smirnov test with Lilliefors Significance Correction, the Shapiro-Wilk test, and *z*-scores for skewness and kurtosis. The peak physiological responses and test duration, and the VO_2, HR, PO and TWD at VT1 and VT2, were compared among protocols. To test for systematic differences, a repeated measures analysis of variance (ANOVA) was performed. Mauchly's test was used to test the assumption of sphericity. A Bonferroni post-hoc test for multiple comparisons was used for pairwise comparisons. Due to the potential risk of bias, imputation of data was not considered. Cohen's *d* effect sizes were calculated and were evaluated according to Hopkins [38] as trivial (0–0.19), small (0.20–0.59), moderate (0.60–1.19), large (1.20–1.99), or very large (≥ 2.00). The ICC was used to measure relative agreement (2.1: two-way random, absolute agreement, single measures), and Bland-Altman plots with 95% limits of agreement (LoA) were used to measure absolute agreement (systematic error and random error) [39]. The following interpretation was used for the ICC: 0.00–0.25, little to no correlation; 0.26–0.49, low correlation; 0.50–0.69, moderate correlation; 0.70–0.89, high correlation; and 0.90–1.00, very high correlation [40]. Values were considered significant at $p < 0.05$, and data were reported as mean (± SD) unless otherwise stated.

Table 1. Descriptives and results of the repeated measures ANOVA for the ramp, 1-min and 3-min protocol.

	Ramp		1-min		3-min		ANOVA			Ramp vs 1-min		Ramp vs 3-min		1-min vs 3-min	
	N	M ± SD	N	M ± SD	N	M ± SD	N	F	p-value	p-value	Cohen's d	p-value	Cohen's d	p-value	Cohen's d
Peak values															
POpeak (W)	19	115 ± 37	19	108 ± 35	19	94 ± 31	19	71.68	< 0.01	< 0.01	0.19	< 0.01	0.62	< 0.01	0.42
VO_2peak (L/min)	19	1.95 ± 0.56	19	1.89 ± 0.56	19	1.99 ± 0.58	19	1.38	0.26	1.00	0.09	1.00	0.07	0.32	0.17
VCO_2peak (L/min)	19	2.67 ± 0.79	19	2.64 ± 0.81	19	2.59 ± 0.72	19	0.56	0.58	1.00	0.04	0.65	0.11	1.00	0.07
VE (L/min)	19	100 ± 33	19	101 ± 34	19	102 ± 32	19	0.17	0.84	1.00	0.03	1.00	0.06	1.00	0.03
HRpeak (bpm)	19	168 ± 17	19	170 ± 16	19	171 ± 15	19	1.06	0.36	0.88	0.14	0.51	0.16	1.00	0.02
RERpeak	19	1.42 ± 0.18	19	1.43 ± 0.19	19	1.36 ± 0.19	19	2.76	0.08	1.00	0.09	0.44	0.31	0.10	0.39
Test duration (min)	19	11.5 ± 2.1	19	10.8 ± 2.4	19	17.8 ± 3.8	19	260.48	< 0.01	< 0.01	0.31	< 0.01	2.06	< 0.01	2.21
Total work done (kJ)	19	41 ± 20	19	40 ± 19	19	59 ± 29	19	48.9	< 0.01	1.00	0.05	< 0.01	0.72	< 0.01	0.78
Ventilatory thresholds															
PO at VT1 (W)	18	51 ± 22	16	44 ± 20	15	36 ± 14	13	4.20	0.03	1.00	0.33	0.047	0.81	0.19	0.46
% of POpeak	18	43 ± 8	16	40 ± 10	15	39 ± 8	13	0.46	0.63	1.00	0.33	1.00	0.50	1.00	0.11
PO at VT2 (W)	18	80 ± 22	17	75 ± 23	17	61 ± 18	15	15.13	< 0.01	0.19	0.22	< 0.01	0.95	0.01	0.68
% of POpeak	18	70 ± 10	17	68 ± 9	17	63 ± 12	15	2.38	0.14	1.00	0.21	0.36	0.63	0.41	0.57
VO_2 at VT1 (L/min)	18	0.98± 0.36	16	1.00 ± 0.33	15	1.01 ± 0.25	13	0.48	0.63	1.00	0.08	0.74	0.26	1.00	0.14
% of VO_2peak	18	51 ± 13	16	52 ± 8	15	52 ± 9	13	0.39	0.68	1.00	0.20	1.00	0.27	1.00	0.07
VO_2 at VT2 (L/min)	18	1.45 ± 0.36	17	1.51 ± 0.47	17	1.38 ± 0.31	15	2.14	0.16	0.69	0.18	0.65	0.22	0.40	0.37
% of VO_2peak	18	75 ± 10	17	78 ± 11	17	70 ± 14	15	6.26	0.01	0.20	0.39	0.14	0.49	0.04	0.85
HR at VT1 (bpm)	18	110 ± 21	16	114 ± 22	15	115 ± 23	13	0.12	0.89	1.00	0.07	1.00	0.11	1.00	0.03
% of HRpeak	18	65 ± 8	16	67 ± 9	15	66 ± 9	13	0.04	0.97	1.00	0.07	1.00	0.07	1.00	0.00
HR at VT2 (bpm)	18	140 ± 18	17	143 ± 17	17	136 ± 24	15	1.73	0.20	0.45	0.34	1.00	0.11	0.43	0.40
% of HRpeak	18	82 ± 8	17	84 ± 6	17	79 ± 10	15	2.57	0.10	0.57	0.36	0.71	0.30	0.25	0.64
TWD at VT1 (kJ)	18	8 ± 6	16	8 ± 6	15	12 ± 8	13	5.35	0.01	1.00	0.00	0.06	0.56	0.05	0.56
% of TWD at peak	18	19 ± 6	16	20 ± 9	15	21 ± 6	13	0.75	0.48	1.00	0.13	0.62	0.33	1.00	0.13
TWD at VT2 (kJ)	18	20 ± 9	17	20 ± 10	17	27 ± 12	15	4.74	0.04	1.00	0.00	0.09	0.66	0.18	0.63
% of TWD at peak	18	50 ± 13	17	50 ± 12	17	46 ± 15	15	1.34	0.27	0.43	0.00	1.00	0.28	0.57	0.29

POpeak = peak power output, VO_2peak = peak oxygen uptake, VCO_2peak = peak carbon dioxide production, VE = minute ventilation, HRpeak = peak heart rate, RERpeak = peak respiratory exchange ratio. VT1 = first ventilatory threshold, VT2 = second ventilatory threshold. TWD = total work done. F = F (Fisher)-statistic of within-subject effects.

Results

All participants completed all tests successfully resulting in a total of 57 tests. In total, 101 out of 114 (89%) VTs could be determined. Of the 13 undetermined VTs, 2 were related to the ramp protocol, 5 to the 1-min stepwise protocol, and 6 to the 3-min stepwise protocol. Eight were VT1 and 5 were VT2 of these 13 undetermined VTs. Outcomes were normally distributed. Peak values and threshold characteristics are shown in table 1. RPE at peak was on average 10 ± 0 for the ramp and 3-min stepwise protocol, and 10 ± 1 for the 1-min stepwise protocol.

Systematic differences among test protocols

Results of the repeated measures ANOVA are shown in table 1. At peak level; VO_2, RER and HR were not significantly different among protocols. POpeak differed significantly among all three protocols, with the highest value for the ramp protocol (115 ± 37 W), followed by the 1-min stepwise (108 ± 35 W) and 3-min stepwise protocol (94 ± 31W). Test duration differed significantly among all three protocols, with the shortest test duration for the 1-min stepwise protocol (10.8 ± 2.4 min), followed by the ramp protocol (11.5 ± 2.1 min) and 3-min stepwise protocol (17.8 ± 3.8 min). TWD was significantly lower for the ramp (41 ± 20 kJ) compared with the 3-min stepwise protocol (59 ± 29 kJ), and for the 1-min stepwise (40 ± 19 kJ) compared with the 3-min stepwise protocol.

At both VTs, absolute values of VO_2 were not significantly different among protocols. The relative VO_2 as a percentage of VO_2peak at VT2 was significantly higher for the 1-min stepwise protocol compared with the 3-min stepwise protocol. Absolute and relative values of HR at VT1 and VT2 were not significantly different among protocols. At VT1, PO was significantly higher for the ramp protocol (51 ± 22 W) compared with the 3-min stepwise protocol (36 ± 14 W). At VT2, PO was significantly higher for the ramp (80 ± 22 W) compared with the 3-min stepwise protocol (61 ± 18 W), and for the 1-min stepwise (75 ± 23 W) compared with the 3-min stepwise protocol. The relative PO as a percentage of POpeak, at both VT1 and VT2, was not significantly different among protocols. Absolute and relative values of TWD at VT1 and VT2 were not significantly different among protocols.

Agreement among test protocols

The relative agreement varied (table 2). Twelve percent of correlations was very high (ICC ≥ 0.90), 24% of correlations was high (ICC ≥ 0.70), whereas 64% was moderate or less (ICC ≤ 0.69).

At peak level, the relative agreement was high to very high for VO_2, HR, PO and TWD. For POpeak and TWD, the lower boundaries of the confidence interval were, however, negative for two comparisons. Figure 1 shows the absolute agreement of POpeak among all protocols. The absolute agreement was low with large systematic error and large random error (i.e., wide 95% LoA).

At both VTs, the relative agreement was moderate to high for VO$_2$, low to high for HR, and low to very high for PO and TWD. The agreement of the relative values at both VTs was in general none to low for VO$_2$, HR, PO and TWD. Figure 2 shows the absolute agreement of HR at both VTs among all protocols. The absolute agreement was low with small-to-large systematic error and large random error (i.e., wide 95% LoA).

Table 2. Relative agreement at peak level and thresholds during arm crank testing for the ramp, 1-min and 3-min protocol.

		ICC (95% CI)		ICC (95% CI)		ICC (95% CI)
	N	Ramp vs 1-min	N	Ramp vs 3-min	N	1-min vs 3-min
Peak values						
POpeak (W)	19	0.97 (0.77 – 0.99)*	19	0.82 (-0.05 – 0.96)*	19	0.90 (-0.01 – 0.98)*
VO$_2$peak (L/min)	19	0.88 (0.71 – 0.95)*	19	0.93 (0.83 – 0.97)*	19	0.90 (0.76 – 0.96)*
VCO$_2$peak (L/min)	19	0.91 (0.79 – 0.97)*	19	0.94 (0.86 – 0.98)*	19	0.90 (0.76 – 0.96)*
VE (L/min)	19	0.91 (0.77 – 0.96)*	19	0.93 (0.84 – 0.97)*	19	0.92 (0.81 – 0.97)*
HRpeak (bpm)	19	0.85 (0.66 – 0.94)*	19	0.88 (0.72 – 0.95)*	19	0.87 (0.70 – 0.95)*
RERpeak	19	0.79 (0.53 – 0.91)*	19	0.60 (0.23 – 0.82)*	19	0.69 (0.34 – 0.87)*
Test duration (min)	19	0.92 (0.42 – 0.98)*	19	0.26 (-0.04 – 0.67)*	19	0.25 (-0.03 – 0.66)*
Total work done (kJ)	19	0.97 (0.92 – 0.99)*	19	0.73 (-0.07 – 0.93)*	19	0.71 (-0.08 – 0.92)*
Ventilatory thresholds						
PO at VT1 (W)	16	0.75 (0.43 – 0.91)*	15	0.56 (-0.02 – 0.84)*	13	0.68 (0.24 – 0.89)*
% of POpeak	16	0.30 (-0.20 – 0.68)	15	0.12 (-0.36 – 0.57)	13	0.36 (-0.25 - 0.76)
PO at VT2 (W)	16	0.93 (0.68 – 0.98)*	17	0.47 (-0.10 – 0.80)*	15	0.50 (-0.03 – 0.81)*
% of POpeak	16	0.79 (0.50 – 0.92)*	17	0.14 (-0.28 – 0.55)	15	0.02 (-0.42 – 0.49)
VO$_2$ at VT1 (L/min)	16	0.61 (0.17 – 0.85)*	15	0.78 (0.47 – 0.92)*	13	0.64 (0.16 – 0.87)*
% of VO$_2$peak	16	0.46 (-0.05 – 0.77)*	15	0.36 (-0.18 – 0.73)	13	0.13 (-0.49 – 0.63)
VO$_2$ at VT2 (L/min)	16	0.84 (0.61 – 0.94)*	17	0.74 (0.43 – 0.90)*	15	0.57 (0.14 – 0.83)*
% of VO$_2$peak	16	0.67 (0.27 – 0.87)*	17	0.57 (0.17 – 0.82)*	15	0.23 (-0.16 – 0.62)
HR at VT1 (bpm)	16	0.54 (0.06 – 0.81)*	15	0.86 (0.65 – 0.95)*	13	0.66 (0.17 – 0.88)*
% of HRpeak	16	0.35 (-0.18 – 0.72)	15	0.78 (0.50 – 0.93)*	13	0.49 (-0.09 – 0.82)*
HR at VT2 (bpm)	16	0.61 (0.21 – 0.84)*	17	0.64 (0.24 – 0.85)*	15	0.47 (0.00 – 0.78)*
% of HRpeak	16	0.47 (0.01 – 0.77)*	17	0.42 (-0.03 – 0.74)*	15	0.12 (-0.31 – 0.55)
TWD at VT1 (kJ)	16	0.67 (0.27 – 0.87)*	15	0.55 (0.02 – 0.83)*	13	0.61 (0.07 – 0.87)*
% of TWD at peak	16	0.27 (-0.26 – 0.67)	15	0.15 (-0.33 – 0.59)	13	0.30 (-0.30 – 0.72)
TWD at VT2 (kJ)	16	0.93 (0.80 – 0.97)*	17	0.49 (0.05 – 0.78)*	15	0.51 (0.06 – 0.80)*
% of TWD at peak	16	0.75 (0.43 – 0.90)*	17	0.22 (-0.30 – 0.62)	15	0.15 (-0.34 – 0.59)

POpeak = peak power output, VO$_2$peak = peak oxygen uptake, VCO$_2$peak = peak carbon dioxide production, VE = minute ventilation, RERpeak = peak respiratory exchange ratio, HRpeak = peak heart rate, TWD = total work done. VT1 = first ventilatory threshold, VT2 = second ventilatory threshold. ICC = intraclass correlation coefficient, CI = confidence interval. * ICC is significant at $p < 0.05$.

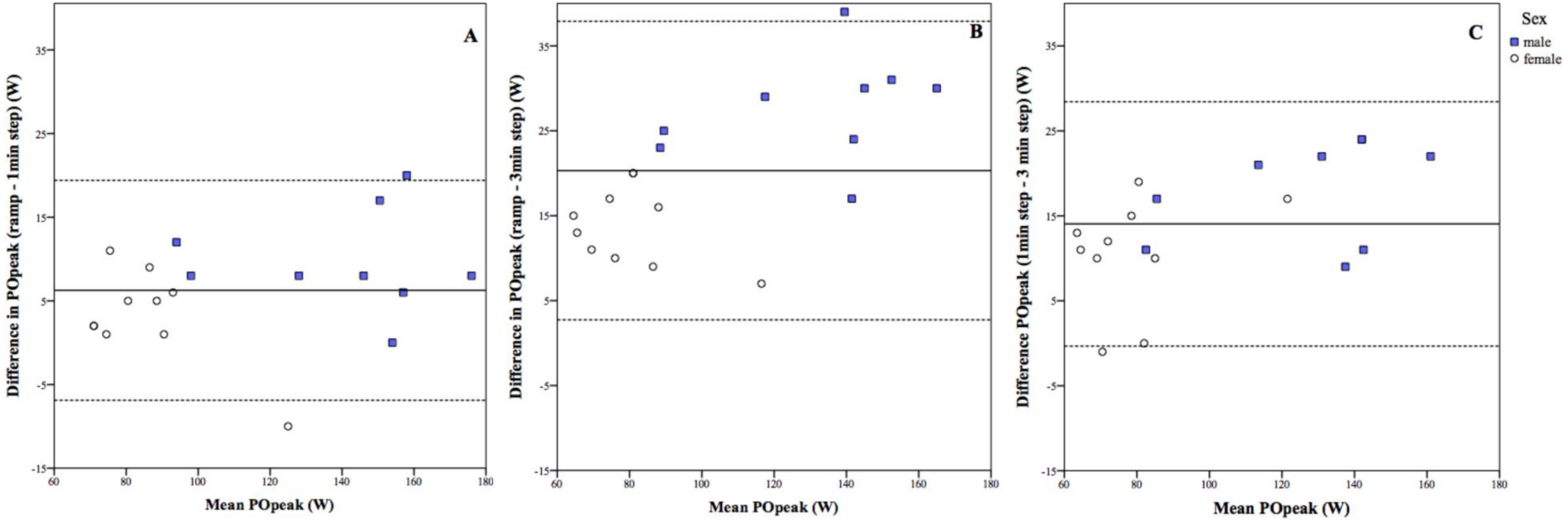


Figure 1. Bland-Altman plot representing absolute agreement. Solid line represents the mean (systematic error), dotted lines represent mean ± 2SD (95% LoA, random error). Each circle represents a participant (N=19). A. Absolute agreement of the peak power output (POpeak) between ramp and 1-min stepwise protocol. The intraclass correlation coefficient was very high (0.97), mean difference 6 W, 95% LoA -7 W to 19 W. B. Absolute agreement of the POpeak between ramp and 3-min stepwise protocol. The intraclass correlation coefficient was high (0.82), mean difference 20 W, 95% LoA 3 W to 38 W. C. Absolute agreement of the POpeak between 1-min and 3-min stepwise protocol. The intraclass correlation coefficient was very high (0.90), mean difference 14 W, 95% LoA 0 W to 28 W.

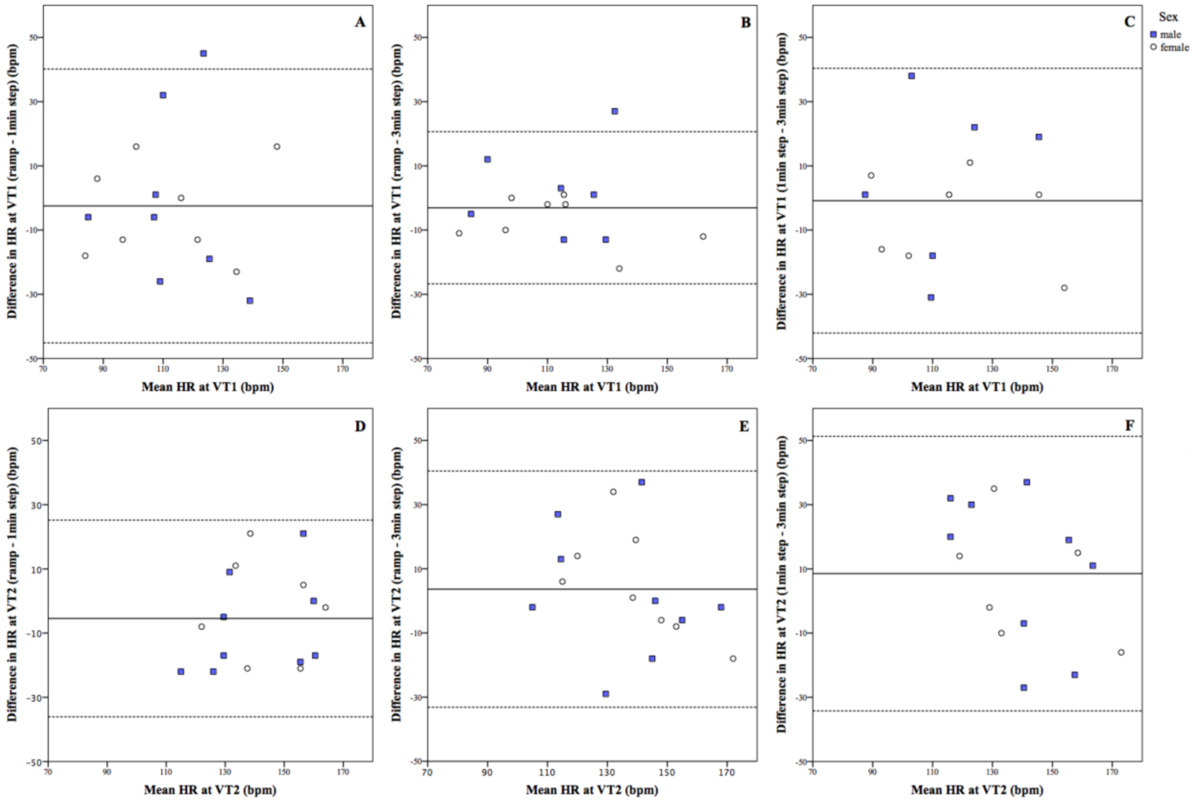


Figure 2. Bland-Altman plot representing absolute agreement. Solid line represents the mean (systematic error), dotted lines represent mean ± 2SD (95% LoA, random error). Each circle represents a participant. A. Absolute agreement of the heart rate (HR) between ramp and 1-min stepwise protocol (N=16). The intraclass correlation coefficient was moderate (0.54), mean difference 3 bpm, 95% LoA -45 bpm to 40 bpm. B. Absolute agreement of the HR between ramp and 3-min stepwise protocol (N=15). The intraclass correlation coefficient was high (0.86), mean difference 3 bpm, 95% LoA -27 bpm to 21 bpm. C. Absolute agreement of the HR between 1-min and 3-min stepwise protocol (N=13). The intraclass correlation coefficient was moderate (0.66), mean difference -1 bpm, 95% LoA -42 bpm to 40 bpm. D. Absolute agreement of the HR between ramp and 1-min stepwise protocol (N=16). The intraclass correlation coefficient was moderate (0.61), mean difference -5 bpm, 95% LoA -36 bpm to 25 bpm. E. Absolute agreement of the HR between ramp and 3-min stepwise protocol (N=17). The intraclass correlation coefficient was moderate (0.64), mean difference 4 bpm, 95% LoA -33 bpm to 40 bpm. F. Absolute agreement of the HR between 1-min and 3-min stepwise protocol (N=15). The intraclass correlation coefficient was low (0.47), mean difference 9 bpm, 95% LoA -34 bpm to 51 bpm.

Discussion

The aim of the present study was to examine the effects of stage duration with a ramp protocol, 1-min stepwise protocol and 3-min stepwise protocol on PO, VO_2 and HR at both peak physiological responses and VTs during synchronous arm crank ergometry. The results at peak level demonstrate that PO showed the highest value for the ramp protocol, followed by the 1-min stepwise and 3-min stepwise protocol. At VT1, PO was significantly higher for the ramp protocol compared with the 3-min stepwise protocol, but there was no significant difference between the ramp and 1-min stepwise and the 1-min stepwise and 3-min stepwise protocols. At VT2, PO was significantly higher for both short-stage protocols compared with the 3-min stepwise protocol. No systematic differences for HR and VO_2 were found among protocols, at both VTs and peak level. The relative agreement among protocols varied with low absolute agreement.

Systematic differences among protocols

The results of the present study are consistent with previous studies in able-bodied cycling [12,17,18,41,42] and arm crank ergometry [9] that demonstrated no significant differences in VO_2peak and HRpeak among protocols with varying stage duration. The differences in POpeak between the 3-min stepwise protocol and the protocols with short-stage duration (1-min stepwise and ramp) are in line with previous cycling literature [12–14,16,18]. Traditionally, short-stage protocols are executed to attain a valid VO_2peak, with the recommendation that test duration should not exceed 12 min [25]. The long-stage protocols, such as the 3-min stepwise protocol, are traditionally executed to attain valid lactate measurements during steady-state conditions to determine a threshold [43]. The accompanying recommendation is that increments should be small and step duration at least 3 min [13,44], resulting in a test duration longer than 12 min. The consequence is a different workload over time. Due to the steeper slope in the short-stage protocols, VO_2peak will be reached faster at a higher POpeak within a shorter test duration [10,18]. The lag in VO_2 response that is typically observed in protocols with short-stage duration results in an underestimation of the steady-state VO_2 at that work rate [10,22]. This also explains why studies that set protocols based on time (i.e., all protocols with expected test duration between 8 and 12 min, irrespective of stage duration) do not find a difference in POpeak [11,41,42]. This, however, does not explain why in the present study a higher POpeak was found in the ramp protocol compared with the 1-min stepwise protocol, as these protocols were set almost identically (only 2 W difference between protocols after 10-min testing and equal work increments between protocols). POpeak is highly dependent on test design and definition: next to stage duration, work increments and test duration; also starting workload, TWD and definition of POpeak are important aspects

[10,17,18,43]. An explanation for the higher POpeak achieved with the ramp protocol, might be the TWD. At a certain similar PO, the TWD was higher for the 1-min stepwise protocol compared with the ramp protocol. In other words, the TWD per minute was higher for the 1-min stepwise protocol compared with the ramp protocol. This higher TWD might lead to fatigue and thus a lower POpeak at the end of the test with shorter test duration [9,17,18]. The results of the ANOVA in the present study support this: TWD at peak level was not significantly different between the ramp and 1-min stepwise protocol, whereas the corresponding POpeak was significantly lower during the 1-min stepwise protocol. Due to the set-up and stepwise character of the 1-min stepwise protocol, participants seem to perform less than with the ramp protocol, whereas in fact, TWD at peak exercise is comparable.

In addition, Smith et al. argued that motivational factors might also play a role [9]. Participants might use external cues during a stepwise protocol to determine the point at which they finish the test, for example at the end of a distinct exercise stage, while increments in workload are less perceptible during a ramp protocol. Consequently, smaller increments in workload, for example 1 W every 5 s instead of distinct 12 W steps every minute, may have less psychological and physiological impact and, therefore, may postpone fatigue and allow participants to reach a higher POpeak [9].

Another important aspect is the definition of POpeak. In the present study POpeak of the ramp protocol was defined as the highest (10 s) PO value at the end of the test, whereas examples exist in which the final 30-s average PO value [10] or the mean minute (60 s) ramp power was calculated to be POpeak [9,45]. If the mean minute ramp power would have been calculated in the present study, POpeak would be 110 ± 36 W and not significantly different from POpeak of the 1-min stepwise protocol. In several studies using ramp protocols, calculation of POpeak is not clearly stated. This is unfortunate as the example stated above shows that this is a requisite to be able to compare literature.

This is the first study that investigated the effect of stage duration at VTs during synchronous arm ergometry. The results are comparable to the previous literature in able-bodied cycling: no differences in HR and VO_2 at VT1 and VT2 were found among protocols with different stage durations [12,15,42], whereas PO at VT1 is significantly lower in tests with longer stage duration [12,16]. Although in the present study there was no systematic difference in the relative PO (i.e., %POpeak) between protocols, we must emphasize that the relative agreement was mostly low or non-existent. Based on the findings in the present study, training zones based on PO at VTs will be at a higher intensity when a short-stage protocol is conducted.

Agreement among test protocols

In general, the results of the present study demonstrated that the level of relative agreement between the ramp, 1-min stepwise, and 3-min stepwise protocol for PO, HR, and VO_2 was not very promising since only 12% of ICC's were higher than 0.90. At peak level, the relative agreement between protocols was high to very high for peak values of VO_2, PO and HR. It must, however, be emphasized that the lower bound of the 95% CI for POpeak was negative in two out of three correlations for POpeak, with low absolute agreement. In addition, there might be a potential effect of heteroscedasticity. The effect is not really evident in figure 1 and, therefore, studies with more observations should be done to determine this more accurately. Maher et al. compared a 1-min stepwise with a 3-min stepwise protocol during arm crank exercise and reported a relative agreement of 0.96, 0.82 and 0.97 for VO_2peak, HRpeak and POpeak, respectively [11]. Smith et al. compared a ramp protocol with a 2-min stepwise protocol during arm crank exercise and reported a relative agreement of 0.67, 0.95 and 0.95 for VO_2peak, HRpeak and POpeak, respectively [9]. Nevertheless, Smith et al. concluded that the absolute agreement between protocols was low for all peak outcome measures and therefore unacceptable [9]. In the present study the 95% LoA were also wide, with a low absolute agreement at VT1, VT2 and peak level. In addition, considering biological variation of HR around 5 beats per minute with day-to-day testing [46], it must be concluded that the random variations in the present study are too large to be acceptable. Therefore, it is recommended that these different test protocols in synchronous arm crank exercise should never be used interchangeably within participants to asses cardiorespiratory fitness. In large studies focussing on physical capacity, the use of the same protocol between participants is advised. When this is not possible, multilevel statistical techniques are necessary to correct for possible differences.

Implications

The results of the present study show that there are systematic differences in PO between protocols at VTs and peak level. Moreover, the absolute agreement in HR at VTs was low due to large random error. Consequently, training zones based on HR or PO will be different among protocols and depending on the chosen protocol. Reviewing the short-stage protocols in the present study, most of the VTs could be determined. However, the non-steady state character of these protocols results in a certain anaerobic contribution to the PO [20]. Consequently, training zones for PO based on a ramp protocol or 1-min stepwise protocol will have a higher intensity than zones based on a 3-min stepwise protocol [43]. Future studies should investigate which protocol suits best to determine training zones for PO, e.g., whether the ramp protocol gives an overestimation with training zones that are at a too high intensity compared with other protocols in synchronous arm cranking, and whether this might result in overreaching. It is suggested that PO at VTs stemming from ramp and 1-min stepwise protocols could be used as objective means to monitor progress, adaptations and functional gains associated with training. However, to prevent overestimation, individual training prescription based on PO at VTs would be most secure based on 3-min stepwise protocols, until future studies are performed. An important side note is that protocols with long test duration might not be feasible for certain patient populations with a very low physical capacity or limited arm function. For example, in individuals with a tetraplegia, protocols with short test duration, such as the ramp or 1-min stepwise protocol, might be more appropriate. For these individuals, training intensity based on HR is often not applicable due to the altered sympathetic response to exercise [47]. It is, therefore, for this population even of more importance to know whether training zones for PO based on short-stage protocols will result in overreaching.

Limitations

The able-bodied participants in the present study were untrained in arm exercise, unlike wheelchair-bound individuals. We did, however, not find any effects of learning among test one, two, and three. An advantageous aspect of able-bodied participants is that the group is homogeneous and that all participants are physically able to complete all test conditions. The group was relatively small, but comparable to or larger than in previous studies [9,10,12–14,16–18,41]. Another general limitation of VT determination is that the position of the VT might be different between raters. In the present study the (relative) interrater reliability was very high, which is acceptable on group level. However, on an individual level in clinical practice, it is advised to evaluate the training zones during training, for example with a talk test [48]. Moreover, in the present study it was not investigated whether prescribing training intensity based on VT determination is favorable to prescription based on RPE or %POpeak in terms of improvements in cardiorespiratory fitness and in terms of over- or undertraining during upper-body exercise. These aspects need to be addressed in future research.

Future studies

An interesting aspect that was not investigated in the present study, is the test-retest reliability of a particular protocol (e.g., the 3-min stepwise protocol) within participants during synchronous arm ergometry. It might be interesting to investigate agreement at VT1, VT2 and peak level with repeated testing of the same exercise protocol in arm exercise, focussing on both trained and untrained individuals and subgroups, such as individuals with paraplegia or tetraplegia. In the light of the present study, it would be interesting to investigate the agreement of PO at both VTs within the 3-min stepwise protocol. It should, however, be considered that protocols with long test duration might be less feasible for individuals with a low physical capacity or limited arm function (e.g., individuals with tetraplegia). Especially for this population, agreement within short stage protocols is warranted.

Conclusion

This study showed that stage duration affects outcomes at both VTs and peak level during synchronous arm crank ergometry in able-bodied participants. No systematic differences for HR and VO_2 were found among protocols. However, PO differed significantly among all protocols at peak level, with the highest value for the ramp protocol, followed by the 1-min stepwise and 3-min stepwise protocol. At VT1, PO was significantly higher for the ramp protocol compared with the 3-min stepwise protocol. At VT2, PO was significantly higher for both short-stage protocols compared with the 3-min stepwise protocol. The relative agreement between protocols varied with low absolute agreement. Consequently, it is recommended that the ramp, 1-min stepwise, and 3-min stepwise arm crank ergometry protocol should never be used interchangeably within persons to assess cardiorespiratory fitness and/or monitor adaptations to training programmes. Furthermore, training prescription based on PO at VTs assessed in short-stage protocols might give an overestimation with training zones that could result in overreaching. Individual training prescription based on PO at VTs would be most secure based on 3-min stepwise protocols, however, protocols with long test duration might not be feasible for certain patient populations with a very low physical capacity. Future studies should pay attention to the effect of stage duration on both peak physiological responses and VTs during arm crank ergometry in subgroups with different abilities and to the consequences of these differences in training zones on training response and overreaching.

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

Acknowledgments

The authors would like to thank Marcel Post, Center of Excellence in Rehabilitation Medicine, De Hoogstraat Rehabilitation, Utrecht, the Netherlands for his support with the manuscript. This study was funded by HandicapNL, Stichting Mitialto, Stichting Beatrixoord Noord-Nederland, University Medical Center Groningen, Heliomare Rehabilitation Center and Stichting Handbike Events.

References

1. Abel T, Vanlandewijck Y, Verellen J. Handcycling. In: Goosey-Tolfrey V, editor. *Wheelchair Sport*. Human Kinetics, Champaign, IL; 2010:187-197.

2. Valent LJM, Dallmeijer AJ, Houdijk H, et al. Effects of hand cycle training on physical capacity in individuals with tetraplegia: a clinical trial. *Phys Ther*. 2009;89(10):1051-1060.

3. Hoekstra SP, Valent LJM, Gobets D, van der Woude LHV, de Groot S. Effects of four-month handbike training under free-living conditions on physical fitness and health in wheelchair users. *Disabil Rehabil*. 2017;39(16):1581-1588.

4. De Groot S, Postma K, van Vliet L, Timmermans R, Valent LJM. Mountain time trial in handcycling: Exercise intensity and predictors of race time in people with spinal cord injury. *Spinal Cord*. 2014;52(6):455-461.

5. Wolpern AE, Burgos DJ, Janot JM, Dalleck LC. Is a threshold-based model a superior method to the relative percent concept for establishing individual exercise intensity? A randomized controlled trial. *BMC Sports Sci Med Rehabil*. 2015;7(1):16.

6. Meyer T, Lucía A, Earnest CP, Kindermann W. A conceptual framework for performance diagnosis and training prescription from submaximal gas exchange parameters - Theory and application. *Int J Sport Med.* 2005;26(1):38-48.

7. Lucia A, Sánchez O, Carvajal A, Chicharro JL. Analysis of the aerobic-anaerobic transition in elite cyclists during incremental exercise with the use of electromyography. *Br J Sports Med*. 1999;33(3):178-185.

8. Seiler KS, Kjerland GØ. Quantifying training intensity distribution in elite endurance athletes: Is there evidence for an “optimal” distribution? *Scand J Med Sci Sport*. 2006;16(1):49-56.

9. Smith PM, Doherty M, Drake D, Price MJ. The influence of step and ramp type protocols on the attainment of peak physiological responses during arm crank ergometry. *Int J Sports Med*. 2004;25(8):616-621.

10. Smith PM, Amaral I, Doherty M, Price MJ, Jones AM. The influence of ramp rate on VO_2peak and “excess” VO_2 during arm crank ergometry. *Int J Sports Med*. 2006;27(8):610-616.

11. Maher JL, Cowan RE. Comparison of 1- versus 3-minute stage duration during arm ergometry in individuals with spinal cord injury. *Arch Phys Med Rehabil*. 2016;97(11):1895-1900.

12. Bentley D, McNaughton L. Comparison of W_{peak}, VO_{2peak} and the ventilation threshold from two different incremental exercise tests: relationship to endurance performance. *J Sci Med Sport*. 2003;6(4):422-435.

13. Bishop D, Jenkins DG, Mackinnon LT. The effect of stage duration on the calculation of peak VO_2 during cycle ergometry. *J Sci Med Sport*. 1998;1(3):171-178.

14. Gullestad L, Myers J, Bjornerheim R, et al. Gas exchange and neurohumoral response to exercise: influence of the exercise protocol. *Med Sci Sport Exerc*. 1997;29(4):496-502.

15. Larson RD, Cantrell GS, Ade CJ, et al. Physiologic responses to two distinct maximal cardiorespiratory exercise protocols. *Int J Sports Exerc Med*. 2015; 1:013.

16. Roffey DM, Byrne NM, Hills AP. Effect of stage duration on physiological variables commonly used to determine maximum aerobic performance during cycle ergometry. *J Sports Sci*. 2007;25(12):1325-1335.

17. Zuniga JM, Housh TJ, Camic CL, et al. Metabolic parameters for ramp versus step incremental cycle ergometer tests. *Appl Physiol Nutr Metab*. 2012;37(6):1110-1117.

18. Amann M, Subudhi A, Foster C. Influence of testing protocol on ventilatory thresholds and cycling performance. *Med Sci Sports Exerc*. 2004;36(4):613-622.

19. Mezzani A. Cardiopulmonary exercise testing : Basics of methodology and measurements. *Ann Am Thorac Soc*. 2017;14(Supplement 1):3-11.

20. Boone J, Bourgois J. The oxygen uptake response to incremental ramp exercise: Methodogical and physiological issues. *Sport Med*. 2012;42(6):511-526.

21. Myers J, Bellin D. Ramp exercise protocols for clinical and cardiopulmonary exercise testing. *Sports Med*. 2000;30(1):23-29.

22. Davis JA, Whipp BJ, Lamarra N, Huntsman DJ, Frank MH, Wasserman K. Effect of ramp slope on determination of aerobic parameters from the ramp exercise test. *Med Sci Sport Exerc*. 1982;14(5):339-343.

23. Chisholm D, Collis M, Kulak M, Davenport W, Gruber N. Physical activity readiness. *B C Med J*. 1975;17:375-378.

24. Mossberg K, Willman C, Topor MA, Crook H, Patak S. Comparison of asynchronous versus synchronous arm crank ergometry. *Spinal Cord*. 1999;37(8):569-574.

25. Buchfuhrer MJ, Hansen JE, Robinson TE, Sue DY, Wasserman K, Whipp BJ. Optimizing the exercise protocol for cardiopulmonary assessment. *J Appl Physiol Respir Environ Exerc Physiol*. 1983;55(5):1558-1564.

26. Widman LM, Abresch RT, Styne DM, McDonald CM. Aerobic fitness and upper extremity strength in patients aged 11 to 21 years with spinal cord dysfunction as compared to ideal weight and overweight controls. *J Spinal Cord Med*. 2007;30(sup1):S88-S96.

27. Hopman MTE, van Teeffelen WM, Brouwer J, Houtman S, Binkhorst RA. Physiological responses to asynchronous and synchronous arm-cranking exercise. *Eur J Appl Physiol Occup Physiol*. 1995;72(1):111-114.

28. Borg GA. Psychophysical bases of perceived exertion. *Med Sci Sport Exerc*. 1982;14(5):377-381.

29. Kuipers H, Verstappen FTJ, Keizer P, Geurten P, van Kranenburg G. Variability of aerobic performance in the laboratory and its physiologic correlates. *Int J Sports Med*. 1985;6(4):197-201.

30. Wasserman K, Hansen JE, Sue DY, Stringer WW, Sietsema, KE, Sun X-G, Whipp BJ. *Principles of Exercise Testing and Interpretation 5th Edition*. Philadelphia, USA. Lippincott Williams & Wilkins, 2012.

31. Gaskill SE, Ruby BC, Walker AJ, Sanchez OA, Serfass RC, Leon AS. Validity and reliability of combining three methods to determine ventilatory threshold. *Med Sci Sport Exerc*. 2001;33(11):1841-1848.

32. Leicht CA, Griggs KE, Lavin J, Tolfrey K, Goosey-Tolfrey VL. Blood lactate and ventilatory thresholds in wheelchair athletes with tetraplegia and paraplegia. *Eur J Appl Physiol*. 2014;114(8):1635-1643.

33. Beaver WL, Wasserman K, Whipp BJ. A new method for detecting anaerobic threshold by gas exchange. *J Appl Physiol*. 1986;60(6):2020-2027.

34. Caiozzo VJ, Davis JA, Ellis JF, et al. A comparison of gas exchange indices used to detect the anaerobic threshold. *J Appl Physiol*. 1982;53:1184-1189.

35. Binder RK, Wonisch M, Corra U, et al. Methodological approach to the first and second lactate threshold in incremental cardiopulmonary exercise testing. *Eur J Cardiovasc Prev Rehabil*. 2008;15(6):726-734.

36. Mezzani A, Hamm LF, Jones AM, et al. Aerobic exercise intensity assessment and prescription in cardiac rehabilitation. *J Cardiopulm Rehabil Prev*. 2012;32(6):327-350.

37. Aunola S, Rusko H. Reproducibility of aerobic and anaerobic thresholds in 20-50 year old men. *Eur J Appl Physiol Occup Physiol*. 1984;53(3):260-266.

38. Hopkins W. A scale of magnitudes for effect statistics. http://sportsci.org/resource/stats/effectmag.html. Accessed 25 Nov 2018.

39. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*. 1986;327:307-310.

40. Munro BH. *Statistical methods for health care research*. 7th edition. Philadelphia, USA. Lippincott Williams & Wilkins; 2004. p. 239 - 258.

41. Bogaard HJ, Woltjer HH, van Keimpema AR, Serra RA, Postmus PE, de Vries PM. Comparison of the respiratory and hemodynamic responses of healthy subjects to exercise in three different protocols. *Occup Med*. 1996;46(4):293-298.

42. Zhang Y, Johnson M, Chow N, Wasserman K. Effect of exercise testing protocol on parameters of aerobic function. *Med Sci Sport Exerc*. 1991;23(5):625-630.

43. Bentley DJ, Newell J, Bishop D. Incremental exercise test design and analysis: Implications for performance diagnostics in endurance athletes. *Sport Med*. 2007;37(7):575-586.

44. Weltman A, Snead D, Stein P, et al. Reliability and validity of a continuous incremental treadmill protocol for the determination of lactate threshold, fixed blood lactate concentrations, and VO_{2max}. *Int J Sports Med*. 1990;11(1):26-32.

45. Ingham SA, Pringle JS, Hardman SL, Fudge BW, Richmond VL. Comparison of step-wise and ramp-wise incremental rowing exercise tests and 2000-m rowing ergometer performance. *Int J Sports Physiol Perform*. 2013;8(2):123-129.

46. McArdle W, Katch F, Katch V. Individual differences and measurements of energy capacities. In: Balado D, editor. *Exercise Physiology: Nutrition, Energy, and Human Performance*. Philadelphia, USA. Lippincott Williams & Wilkins, 2010, p 245.

47. Valent LJM, Dallmeijer AJ, Houdijk H, et al. The individual relationship between heart rate and oxygen uptake in people with a tetraplegia during exercise. *Spinal Cord*. 2007;45(1):104-111.

48. Kouwijzer I, Cowan RE, Maher JL, et al. Interrater and intrarater reliability of ventilatory thresholds determined in individuals with spinal cord injury. *Spinal Cord*. 2019;57(8):669-678.

Chapter 5

Training for the HandbikeBattle: an explorative analysis of training load and handcycling physical capacity in previously untrained wheelchair users

Kouwijzer I, Valent LJM, van Bennekom CAM, HandbikeBattle Group, Post MWM, van der Woude LHV, de Groot S. Training for the HandbikeBattle: an explorative analysis of training load and handcycling physical capacity in previously untrained wheelchair users. *Under review*.

Abstract

Purpose: (1) to analyze training characteristics of previously untrained wheelchair users during handcycle training, and (2) to examine the associations between training load and change in physical capacity.

Methods: Former rehabilitation patients (N=60) with health conditions such as spinal cord injury or amputation were included. Participants trained for five months. A handcycling / arm crank graded exercise test was performed before and after the training period. Outcomes: peak power output per kg (POpeak/kg) and peak oxygen uptake per kg (VO_2peak/kg). Training load was defined as Training Impulse (TRIMP), which is rating of perceived exertion (sRPE) multiplied by duration of the session, in arbitrary units (AU). Training intensity distribution (TID) was also determined (time in zone 1, RPE ≤ 4; zone 2, RPE 5-6; zone 3, RPE ≥ 7).

Results: Multilevel regression analyses showed that $TRIMP_{sRPE}$ was not significantly associated with change in physical capacity. Time in zone 2 (RPE 5-6) was significantly associated with ΔVO_2peak, $\%\Delta VO_2$peak, ΔVO_2peak/kg and $\%\Delta VO_2$peak/kg.

Conclusion: This study shows no significant associations between change in physical capacity and (components of) training load. However, TID could be a key component in physical capacity improvement as training time in moderate intensity showed a significant association with change in physical capacity.

Keywords: sRPE, training intensity distribution, monotony, strain, upper-body exercise

Introduction

Physical capacity is generally reduced in manual wheelchair users [1]. A low physical capacity is associated with a high prevalence of cardiometabolic disease [2]. Therefore, exercise interventions to increase physical capacity in wheelchair users are important. An interesting goal to train for is the HandbikeBattle [3]. The HandbikeBattle is organized as an annual event in the mountains of Austria and is an uphill handcycling mountain race among teams of Dutch rehabilitation centers. The teams consist of former patients with, among others, a spinal cord injury (SCI) or amputation. The event was created to initiate an active lifestyle by means of free-living handcycle training.

Handcycling is a common exercise mode for manual wheelchair users during and after rehabilitation [4,5]. It is shown that handcycle training results in improvement in physical capacity during and after rehabilitation even in the most vulnerable patients with a tetraplegia [5]. Furthermore, handcycling has a higher efficiency than handrim wheelchair propulsion and leads to lower shoulder loads [6,7]. This is important as 30-73% of wheelchair users with a SCI experience musculoskeletal pain in the shoulder [8,9]. Handcycle training studies during or after rehabilitation are, unfortunately, scarce. In addition, studies related to upper body training often have a small heterogeneous sample size or do not take training load into consideration.

Elite athletes commonly monitor their training load, yet it is less common in rehabilitation interventions. Monitoring of training load is important to structure the training effort and intensity over time and to eventually optimize performance capacity. In turn, critical assessment of training load helps to prevent undertraining or overtraining [10]. Indices of training load relate to form, frequency, duration and intensity. Training load can be divided in external and internal training load [10]. In (hand)cycling, external training load is often represented by the training stress score (TSS) based on power output (PO (W)) [11]. This is an objective measure of external training load, but it is costly, and only applicable to handcycling and not to others forms of exercise (therapy) in rehabilitation. Internal training load measures are, among others, the training impulse (TRIMP) based on heart rate reserve (HRR) [12] or the session rating of perceived exertion (sRPE) [13]. TRIMP based on sRPE ($\text{TRIMP}_{\text{sRPE}}$) is calculated by multiplying the overall RPE of the session by the duration of the session in minutes [13]. $\text{TRIMP}_{\text{sRPE}}$ is an easy to use and cheap method to monitor internal training load, is applicable to different training modes, and gives an overall representation of the individual's perception of training, potentially taking into account physical, psychological and environmental factors [14]. These subjective factors are very important to the individual's training response, in addition to the imposed objective external training load [15]. An additional advantage of $\text{TRIMP}_{\text{sRPE}}$ as internal training load measure is that for individuals with tetraplegia training intensity based on heart rate (HR) is often not applicable due to the altered sympathetic response to exercise, which makes heart rate difficult to interpret [16]. It would, therefore, be an ideal method to monitor training sessions in rehabilitation. Previous studies showed large to nearly perfect correlations (0.5 – 0.97) among $\mathrm{TRIMP}_{\mathrm{sRPE}}$ and HR-based TRIMP methods in sprint kayak, wheelchair basketball, soccer, cycling and recreational handcycling [15,17–20]. Whereas very large to nearly perfect correlations (0.81 – 0.95) were found between $\mathrm{TRIMP}_{\mathrm{sRPE}}$ and TSS in cycling and recreational handcycling [19,20].

Although training load measures generally correlate well and training monitoring based on training load seems to be useful during the training process [14,21], dose-response relationships with improvements in physical capacity remain controversial [21]. Foster et al. found a correlation of 0.029 between increase in $\mathrm{TRIMP}_{\mathrm{sRPE}}$ and improvement in time trial performance [22]. In addition, $\mathrm{TRIMP}_{\mathrm{sRPE}}$ explained only 12% of the variance of change in $\mathrm{VO_2max}$ in rugby players [23], and small to moderate correlations were found between $\mathrm{TRIMP}_{\mathrm{sRPE}}$ and change in performance in hurling [24]. A recent study in elite cyclists found that different training load measures ($\mathrm{TRIMP}_{\mathrm{sRPE}}$, HR-based TRIMP and TSS) were only correlated to submaximal outcome measures (PO at 2mmol/L and 4 mmol/L blood lactate), and not to changes in POmax or $\mathrm{VO_2max}$ [25]. In recreational cyclists, no relationships were found among different HR-based TRIMP methods and change in POmax [26]. In addition to training load itself, it was proposed that training time in each intensity zone (training intensity distribution, TID) [26,27], lack of day-to-day variability in training load (monotony) and training strain could all play a role in adaptations to training [13,28,29].

Taken together, knowledge on training adaptations is rapidly increasing but far from complete or consistent. Especially in adaptive sports and upper-body training in wheelchair users during rehabilitation there is a lack of knowledge about suitable training regimes, loads and dose-response relationships. Previously it has been shown that training for the HandbikeBattle leads to improvements in physical capacity and health [4]. It is, however, unknown what training regimes led to these improvements. In an attempt to unravel more details on training regimes and dose-response relationships of handcycle training, the purpose of this explorative prospective cohort study was (1) to analyze training characteristics, and (2) to examine the associations between training load and the change in physical capacity.

Materials & methods

Participants

Inclusion criteria for the HandbikeBattle event were: being a former rehabilitation patient from one of the twelve rehabilitation centers; impairment of the lower extremities due to e.g., SCI, amputation, cerebral palsy or spina bifida; and commitment to the HandbikeBattle challenge. Exclusion criterion:

contra-indications to participate in the HandbikeBattle as diagnosed during the medical screening before the training period. In the present study, data were used from participants of the HandbikeBattle 2013 and 2015-2019 cohorts. In total 227 individuals were recruited to start monitoring their training sessions in this period. Twenty-six individuals dropped out during the training period for the HandbikeBattle due to motivational problems (N=4), medical reasons (N=16), family matters (N=1), not being able to combine training with activities of daily living (N=4), or financial reasons (N=1). No individuals dropped out due to overuse injuries. Twenty-one individuals did not complete the GXT before or after the training period. Another 120 individuals did not have complete training data. Training data were considered complete if more than 80% of training sessions had a filled out RPE. Hence, data from 60 participants were used in the present study, whereas data from 167 individuals could not be used. All participants provided written informed consent. The study was approved by the Local Ethics Committee of the Center for Human Movement Sciences, University Medical Center Groningen, the Netherlands (ECB/2012_12.04_I_rev/Ml).

Procedures

Design

The HandbikeBattle event is a serious challenge (20.2-km length and 863 m elevation gain) each year in June. At the start of the 5-month training period, most participants are relatively untrained handcyclists. Connected to, but not part of, the HandbikeBattle event is a prospective observational cohort study that was initiated to monitor effects of participation in the training period and the event. Measurements were performed at the start of the training period (January, T1), during the training period, and after the training period prior to the event (June, T2). At T1 a medical screening was performed by a rehabilitation physician or sports physician at the rehabilitation center. The screening comprised a medical anamnesis, physical examination and a handcycling / arm crank graded exercise test (GXT). At T2 the GXT was repeated with the same protocol and equipment. At T1 participants were asked to fill out a questionnaire about musculoskeletal shoulder pain. Guidance during the training period was provided by therapists from the respective rehabilitation centers, but otherwise the training period was free-living, i.e., no specific training program was provided by the researchers. After the GXT at T1, participants started to train indoors and outdoors. The main part of the training was done individually or together with HandbikeBattle participants from the same rehabilitation center. All participants were asked to monitor all their sporting activity with an online app (Strava) or a training diary on paper.

Physical capacity

Physical capacity was measured during an incremental handcycling / arm crank GXT to volitional exhaustion at T1 and T2, organized in and conducted by the staff of each of the participating rehabilitation centers. All tests were performed in synchronous mode of cranking. Dependent on the rehabilitation center, the GXTs were performed with the use of an arm ergometer (Lode Angio, Groningen, the Netherlands) or a recumbent sport handcycle attached to the Cyclus 2 ergometer (RBM elektronik-automation GmbH, Leipzig, Germany). Either a 1-min stepwise protocol, 3-min stepwise protocol or continuous ramp protocol was used, and was individualized for each participant. The set-up and protocol choice were consistent within participants over time. Criteria to stop the test were volitional exhaustion or failure in keeping a constant cadence above the preset value. PO (W) and oxygen uptake (VO_2 (L/min)) were measured during the test. For the 1-min stepwise protocol, POpeak was defined as the highest PO that was maintained for at least 30 s. For the 3-min stepwise protocol POpeak was determined as the highest PO maintained over 3 minutes, plus 1/6 x step size in Watts for every additional 30 s in the next step [30]. For the ramp protocol, the highest PO achieved during the test was considered POpeak. Peak oxygen uptake (VO_2peak) was defined as the highest 30-s average for VO_2. Outcome parameters in the analyses were the absolute and relative changes in POpeak/kg and VO_2peak/kg between T1 and T2 (ΔPOpeak/kg, %ΔPOpeak/kg, and ΔVO_2peak/kg, and %ΔVO_2peak/kg).

Training load calculation

Participants were asked to fill out after each training session: the type of training, duration of the training (minutes) and the sRPE score on a scale from 0 to 10 (Modified CR-10 scale) [13]. If the sRPE score was missing for a session, the average sRPE score of the same type of training was used for the analysis to calculate the $TRIMP_{sRPE}$ for that session [23]. $TRIMP_{sRPE}$ was calculated by multiplying the overall RPE of the session by the duration of the session in minutes [13]. Total $TRIMP_{sRPE}$ in arbitrary units (AU) was calculated as the sum of $TRIMP_{sRPE}$ of all training sessions during the training period for each participant. Average monotony per week (AU) was calculated per participant per week as the average daily $TRIMP_{sRPE}$ (AU) divided by the SD of the daily $TRIMP_{sRPE}$ of that week [29]. Total monotony (AU) was calculated for each participant as the sum of the weekly monotony for all weeks during the training period. Average strain per week (AU) was calculated per participant per week as the average $TRIMP_{sRPE}$ per week multiplied by average monotony per week [29]. Total strain (AU) was calculated for each participant as the sum of the weekly strain for all weeks during the training period. TID was calculated as the relative and absolute time and number of sessions spent in low intensity (zone 1, RPE ≤ 4), moderate intensity (zone 2, RPE 5-6) and high intensity (zone 3, RPE ≥ 7) [27].

Possible confounding variables

Possible confounding variables were musculoskeletal shoulder pain at T1, and handcycling classification. Age and sex were not considered as they seem to have less/no influence on training adaptations [31].

Musculoskeletal shoulder pain comprised two locations (shoulder (L/R)) with range 1=no pain, 6=very severe pain. Two groups were created: no-mild pain=0, moderate-severe pain=1. Having moderate-severe pain was defined as ≥4 (moderate pain) at one or both locations.

Handcycling classification was used as a proxy for severity of impairment and determined by an UCI certified Paracycling classifier, following the UCI Para-cycling Regulations: resulting in five classes, ranging from H1 (most impaired) to H5 (least impaired) [32]. H1 and H2 handcyclists have limitations in arm-hand function, whereas H3, H4 and H5 handcyclists have intact arm-hand function and limitations in trunk and/or lower extremities only. For the analyses in the present study, participants were divided in two large groups: (1) H1-H3 and (2) H4-H5.

Statistical analyses

The analyses were performed using SPSS (IBM SPSS Statistics for Windows, Version 24.0. Armonk, NY: IBM Corp.) and MLwiN Version 3.02 [33]. Descriptive statistics were calculated for outcome measures and determinants. Data were tested for normality with the Kolmogorov–Smirnov test with Lilliefors significance correction and the Shapiro–Wilk test, combined with z-scores for skewness and kurtosis. To ascertain possible response bias, characteristics of included participants in the present study (N=60) were compared with non-participants (N=167) using independent-samples t-tests, Mann-Whitney U tests and chi-squared tests. Changes in physical capacity were tested with paired-samples t-tests. Cohen's *d* effect sizes were calculated and were evaluated according to Hopkins as trivial (0–0.19), small (0.20–0.59), moderate (0.60–1.19), large (1.20–1.99), or very large (≥ 2.00) [34]. The Pearson product-moment correlation (r) was used to examine the associations among the training load determinants and changes in physical capacity, with a Spearman's rank correlation (ρ) in case of non-normality. The strength of the correlation coefficients was evaluated according to Hopkins as trivial (0-0.09), small (0.10-0.29), moderate (0.30-0.49), large (0.50-0.69), or very large (≥ 0.70) [34]. In addition, multilevel regression analyses were used to examine specific multivariate associations. Two-level models were created with participant as first level and rehabilitation center as second level to be able to make adjustments for the dependency of participants within centers. The first set of regression analyses comprised the association between change in physical capacity (ΔPOpeak/kg, %ΔPOpeak/kg, ΔVO_2peak/kg, and %ΔVO_2peak/kg) and Total $TRIMP_{sRPE}$ (basic models). Each regression analysis was corrected for baseline value of the outcome measure (POpeak/kg or VO_2peak/kg at T1) and duration of the training period (weeks). Thereafter, shoulder pain and handcycling classification were added as possible confounders (final models). A variable was included as confounder in the final model if its inclusion changed the beta of training load with more than 10% [35]. The second set of multilevel regression analyses comprised the association between change in physical capacity (ΔPOpeak/kg, %ΔPOpeak/kg, ΔVO_2peak/kg, and %ΔVO_2peak/kg), and separate determinants for frequency, duration and intensity: duration of the training period (weeks), number of training sessions per week (N), average training volume per training session (min), and average sRPE per training session. Additional explorative analyses were performed with TID, total monotony, total strain and high intensity training sessions only (RPE > 5) as determinants [22]; and with ΔPOpeak and ΔVO_2peak as outcome parameters. Significance was set at $p < 0.05$ for all statistical analyses.

Results

Participants had more often a high classification (H4-H5) than non-participants (table 1). Within the non-participants group, POpeak and POpeak/kg at T1 were lower for dropouts compared with individuals who completed the training period but had incomplete training data (table 1). All outcome measures were normally distributed. A total of 4617 training sessions were analyzed for this study. The most common training sessions comprised: handcycling (N=3269), strength and conditioning (N=895), swimming (N=60), wheelchair basketball (N=50), and wheelchair rugby (N=45). Handcycling was the main sport for all participants. Twenty-one participants had a filled out sRPE in all training sessions. Thirty-nine participants had missing sRPE with an average of 6.1% missing data (SD: 4.6, range: 1 – 17%). Participants trained for 21 ± 6 weeks with an average of 3.6 ± 1.4 training sessions per week (table 2). Mean weekly $TRIMP_{sRPE}$ was 1654 ± 579 AU (table 2). Physical capacity showed a significant increase between T1 (before training period) and T2 (after training period) (table 3). Figure 1 shows two typical examples of training characteristics of participants training for the event.

Correlations between training characteristics and outcome parameters were trivial to small (table 2). Total $TRIMP_{sRPE}$ was not significantly associated with change in physical capacity (table 4). After adding confounders to the models, associations remained non-significant (table 4). Separate determinants for frequency, duration and intensity showed no significant associations except for a negative association between duration of the training period and ΔVO_2peak/kg (table 5).

Table 1. Characteristics and outcomes at T1 for participants (N=60) and non-participants (N=167).

Characteristics	**Participants**			**Non-participants**								
	N			**N**	**Total**		**N**	**Incomplete data**		**N**	**Drop-outs**	
Sex (male/female) (%)	60	39/21	(65/35)	167	115/52	(69/31)	141	99/42	(70/30)	26	16/10	(62/38)
Age (years) (SD)	60	40	(12)	166	41	(14)	141	41	(14)	25	41	(13)
Impairment type	60			166			141			25		
Spinal cord injury (%)		26	(43)		95	(57)		80	(57)		15	(60)
Tetraplegia		5	(8)		16	(10)		14	(10)		2	(8)
Paraplegia		21	(35)		79	(47)		66	(47)		13	(52)
Amputation (%)		8	(13)		19	(11)		15	(11)		4	(16)
Multi trauma (%)		1	(2)		6	(4)		6	(4)		0	(0)
Spina bifida (%)		7	(12)		12	(7)		10	(7)		2	(8)
Other (%)		18	(30)		34	(21)		30	(21)		4	(16)
POpeak (W) (SD) T1	59	118	(39)	155	109	(41)	134	$112^{\dagger}$	(42)	21	$90^{\dagger}$	(24)
ΔPOpeak (W) (SD)	59	22	(18)	122	19	(17)	122	19	(17)	0	-	-
POpeak/kg (W/kg) (SD) T1	59	1.51	(0.51)	147	1.44	(0.52)	127	$1.48^{\dagger}$	(0.54)	20	$1.18^{\dagger}$	(0.30)
ΔPOpeak/kg (W/kg)	59	0.30	(0.24)	113	0.27	(0.25)	113	0.27	(0.25)	0	-	-
VO_2peak (L/min) (SD) T1	59	1.91	(0.57)	154	1.76	(0.54)	134	1.79	(0.55)	20	1.57	(0.45)
ΔVO_2peak (L/min) (SD)	59	0.30	(0.27)	118	0.24	(0.27)	118	0.24	(0.27)	0	-	-
VO_2peak/kg (ml/kg/min) (SD) T1	59	24.75	(7.88)	146	23.49	(6.82)	127	23.89	(6.95)	19	20.82	(5.35)
ΔVO_2peak/kg (ml/kg/min)	59	4.13	(3.37)	110	3.31	(3.82)	110	3.31	(3.82)	0	-	-
Shoulder pain (no-mild/moderate-severe) (%) T1	54	44/10	(81/19)	123	101/22	(82/18)	109	90/19	(83/17)	14	11/3	(79/21)
Handcycling classification (H1–H3/H4–H5) (%)	60	22/38*	(37/63)	153	85/68*	(56/44)	141	79/62	(56/44)	12	6/6	(50/50)

Data represent N (%) or mean (SD). POpeak: peak power output; VO_2peak: peak oxygen uptake. Shoulder pain: two categories: (1) no-mild pain and (2) moderate-severe pain. Handcycling classification: two categories: (1) H1–H3 and (2) H4–H5. * Significant difference with $p < 0.05$ between participants and non-participants. † Significant difference with $p < 0.05$ between non-participants with incomplete training data and non-participants who dropped-out.

Table 2. Overview of training characteristics (N=60) and correlations with outcome parameters.

			ΔPOpeak/kg	**%ΔPOpeak/kg**	**ΔVO$_2$peak/kg**	**%ΔVO$_2$peak/kg**
	Mean ± SD	Range (min – max)	*r* (*p*-value)	*r* (*p*-value)	*r* (*p*-value)	*r* (*p*-value)
Duration of training period (weeks)	21 ± 6	8 - 33	-0.01 (0.92)	0.10 (0.44)	-0.14 (0.31)	-0.04 (0.76)
Number of training sessions	77 ± 40	23 - 183	-0.00 (0.99)	0.14 (0.29)	0.01 (0.95)	0.08 (0.55)
Number of training sessions per week	3.6 ± 1.4	1 - 8	0.00 (0.99)	0.11 (0.40)	0.10 (0.45)	0.13 (0.34)
Total training volume (min)	6174 ± 2841	1635 - 13728	-0.03 (0.83)	0.06 (0.65)	-0.08 (0.57)	-0.01 (0.92)
Average training volume per week (min)	299 ± 102	112 - 572	-0.07 (0.62)	-0.06 (0.68)	-0.03 (0.82)	-0.03 (0.85)
Average training volume per training session (min)	86 ± 20	47 - 136	-0.07 (0.58)	-0.24 (0.07)	-0.16 (0.23)	-0.19 (0.16)
Total TRIMP$_{\text{sRPE}}$ (AU)	33892 ± 14746	9293 - 69440	-0.03 (0.84)	0.09 (0.50)	-0.09 (0.51)	-0.04 (0.77)
Average TRIMP$_{\text{sRPE}}$ per week (AU)	1654 ± 579	622 - 3350	-0.05 (0.72)	-0.01 (0.94)	-0.02 (0.90)	-0.03 (0.81)
Average TRIMP$_{\text{sRPE}}$ per training session (AU)	484 ± 154	199 - 919	-0.10 (0.44)	-0.19 (0.16)	-0.14 (0.28)	-0.18 (0.18)
Average sRPE per training session	5.4 ± 1.3	3 - 8	0.03 (0.83)	0.07 (0.60)	-0.02 (0.88)	-0.02 (0.86)
Total monotony (AU)	15.9 ± 7.3	4.5 – 36.0	-0.01 (0.95)	0.13 (0.34)	-0.06 (0.64)	0.01 (0.95)
Average monotony per week (AU)	0.8 ± 0.2	0.3 – 1.4	-0.03 (0.82)	0.06 (0.65)	-0.00 (0.98)	0.01 (0.93)
Total strain (AU)	29879 ± 19465	5615 - 99140	0.07 (0.58)*	0.10 (0.44)*	0.06 (0.68)*	0.07 (0.62)*
Average strain per week (AU)	1483 ± 835	429 - 4131	0.05 (0.72)*	0.06 (0.68)*	0.04 (0.78)*	0.01 (0.97)*
Training intensity distribution (RPE 1-4) (N sessions)	28.7 ± 30.5	0 - 130	0.03 (0.85)*	-0.01 (0.94)*	-0.03 (0.84)*	-0.05 (0.69)*
Training intensity distribution (RPE 5-6) (N sessions)	24.2 ± 21.6	1 - 117	-0.06 (0.63)*	0.04 (0.78)*	0.15 (0.27)*	0.19 (0.16)*
Training intensity distribution (RPE 7-10) (N sessions)	23.9 ± 22.1	0 - 105	0.09 (0.50)*	0.18 (0.17)*	0.00 (0.99)*	0.05 (0.70)*
Training intensity distribution (RPE 1-4) (% sessions)	35.3 ± 29.0	0 – 94	0.00 (0.99)	-0.07 (0.59)	-0.03 (0.81)	-0.04 (0.76)
Training intensity distribution (RPE 5-6) (% sessions)	30.9 ± 18.0	2 – 75	-0.05 (0.72)	-0.01 (0.96)	0.20 (0.13)	0.22 (0.10)
Training intensity distribution (RPE 7-10) (% sessions)	33.7 ± 28.2	0 – 98	0.03 (0.84)	0.07 (0.58)	-0.10 (0.46)	-0.10 (0.44)
Training intensity distribution (RPE 1-4) time (min)	2129 ± 2425	0 - 11998	0.04 (0.79)*	-0.04 (0.75)*	-0.03 (0.85)*	-0.07 (0.59)*
Training intensity distribution (RPE 5-6) time (min)	1878 ± 1437	30 – 7458	-0.11 (0.41)*	-0.04 (0.75)*	0.19 (0.14)*	0.22 (0.10)*
Training intensity distribution (RPE 7-10) time (min)	2155 ± 1826	0 – 7304	0.01 (0.94)*	0.09 (0.51)*	-0.10 (0.45)*	-0.06 (0.63)*
Training intensity distribution (RPE 1-4) %time	31.5 ± 27.8	0 - 96	0.01 (0.97)	-0.08 (0.56)	-0.07 (0.63)	-0.06 (0.63)
Training intensity distribution (RPE 5-6) %time	30.8 ± 18.1	1 – 80	-0.09 (0.52)	-0.06 (0.68)	0.24 (0.07)	0.26 (0.05)
Training intensity distribution (RPE 7-10) %time	37.6 ± 28.4	0 – 99	0.05 (0.73)	0.11 (0.43)	-0.10 (0.47)	-0.11 (0.40)

Data represent % or mean (SD). sRPE: session rating of perceived exertion; TRIMP: Training Impulse; POpeak: peak power output; VO$_2$peak: peak oxygen uptake.

* A Spearman's ρ instead of Pearson's *r*.

Table 3. Physical capacity before (T1) and after (T2) the training period.

	N	T1 (pre-training)	T2 (post-training)	Mean difference Δ (%)	*p*-value	Effect size	Qualitative outcome
POpeak (W)	59	118 ± 39	138 ± 45	22 ± 18 (20%)	< 0.001	0.52	Small effect
POpeak/kg (W/kg)	59	1.51 ± 0.51	1.80 ± 0.57	0.30 ± 0.24 (22%)	< 0.001	0.56	Small effect
VO$_2$peak (L/min)	59	1.91 ± 0.57	2.23 ± 0.66	0.30 ± 0.27 (17%)	< 0.001	0.48	Small effect
VO$_2$peak/kg (ml/min/kg)	59	24.76 ± 7.88	29.00 ± 8.03	4.12 ± 3.36 (18%)	< 0.001	0.52	Small effect

Data represent mean ± SD. POpeak: peak power output; VO$_2$peak: peak oxygen uptake. N = 60, however, 1 participant did not have POpeak and did have VO$_2$peak, whereas 1 other participant did not have VO$_2$peak and did have a POpeak.

Table 4. Basic and final models. Associations between absolute/relative change in physical capacity and total TRIMP$_{sRPE}$.

	ΔPOpeak/kg (x1000)			%ΔPOpeak/kg (x1000)			ΔVO$_2$peak/kg (x1000)			%ΔVO$_2$peak/kg (x1000)		
	Beta	SE	*p*-value	Beta	SE	*p*-value	Beta	SE	*p*-value	Beta	SE	*p*-value
Basic models												
Constant	322.795	174.474		41040.014	12216.585		8677.781	2373.900		43232.485	9857.727	
Total TRIMP$_{sRPE}$ (AU)	-0.000	0.003	0.85	0.021	0.180	0.91	0.000	0.035	1.00	0.003	0.146	0.98
Training period (weeks)	0.013	6.542	1.00	-46.652	458.077	0.92	-110.808	91.556	0.23	-366.253	380.192	0.34
Outcome at T1	-6.211	63.648	0.92	-12568.108	4456.565	0.005	-92.209	56.100	0.10	-710.645	232.958	0.002
Final models												
Constant	145.037	194.411		29628.436	13596.147		7442.817	2733.174		34979.079	11245.838	
Total TRIMP$_{sRPE}$ (AU)	0.001	0.003	0.73	0.118	0.195	0.55	0.027	0.039	0.49	0.122	0.161	0.90
Training period (weeks)	-2.321	7.206	0.75	-213.232	503.975	0.67	-163.174	102.917	0.11	-580.116	423.459	0.17
Outcome at T1	31.277	71.575	0.66	-10512.696	5005.644	0.04	-99.125	62.052	0.11	-658.526	255.318	0.01
Confounders												
Shoulder pain	142.326	89.764	0.47	11183.739	6277.665	0.07	966.735	1253.308	0.44	6850.936	5156.824	0.18
Classification	164.083	70.598	0.02	10943.692	4937.260	0.03	2134.731	980.212	0.03	9649.970	4033.152	0.02

Corrected for duration of training period (weeks) and value of the outcome parameter at T1. Shoulder pain: two categories: (1) no-mild pain and (2) moderate-severe pain (reference: no-mild). Handcycling classification: two categories: (1) H1–H3 and (2) H4–H5 (reference: H1-H3). A variable was included as confounder if the regression coefficient of total TRIMP$_{sRPE}$ changed more than 10%. Physical capacity outcome parameters are multiplied by 1000 to visualize the details in the beta.

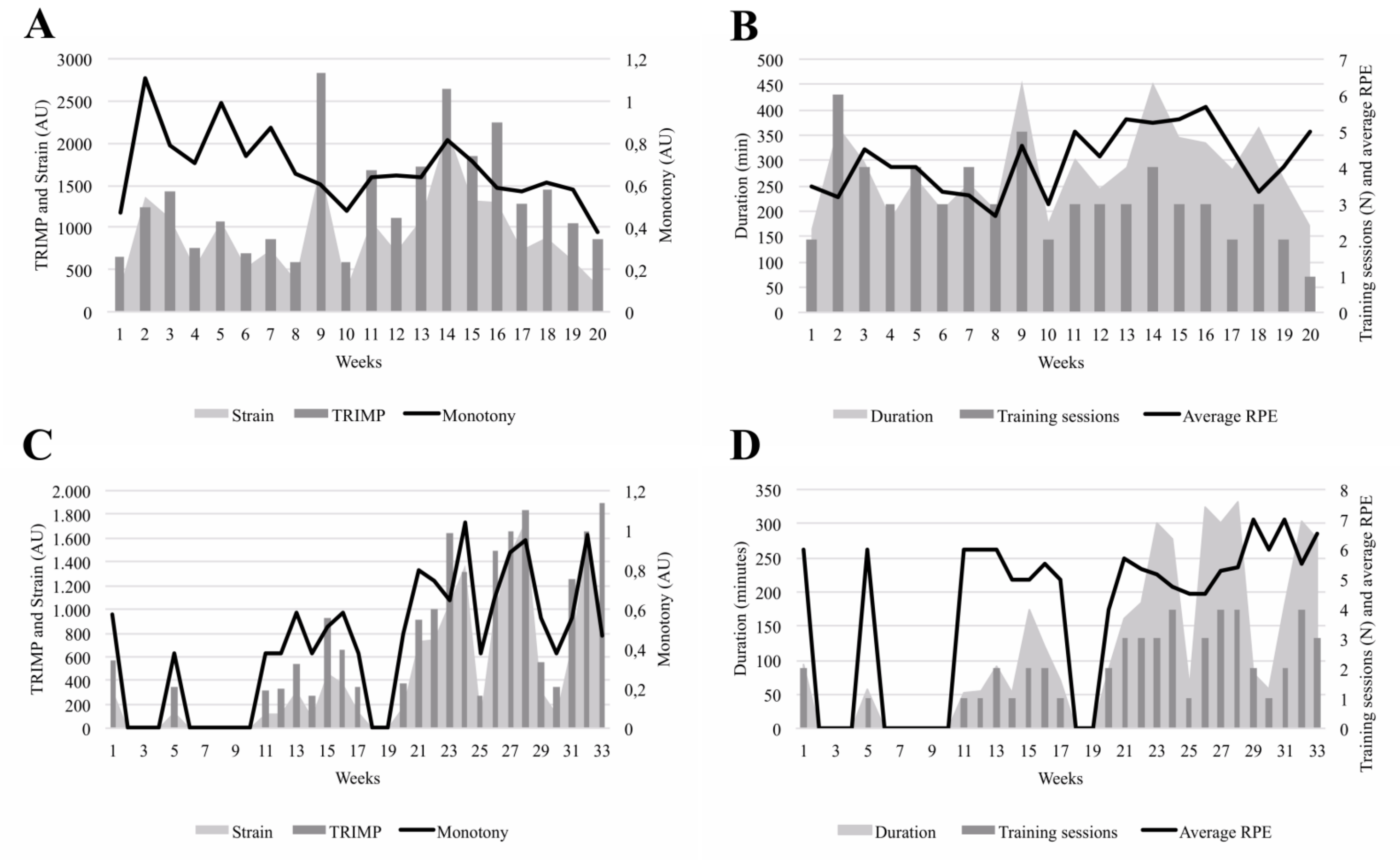


Figure 1 A and B. Typical example of a participant who showed a relatively consistent training period. H5 handcyclist with a paraplegia. At T1: VO_2peak 2.38 L/min, POpeak 115W. Relative change in VO_2peak/kg: 4%, relative change in POpeak/kg: 11%. Training period was 20 weeks, with 3 training sessions per week on average, and average TRIMP per week of 1330 AU. Training volume per training: 90 minutes, average RPE per session: 4.2. **Figure 1 C and D**. Typical example of a participant who showed a relatively long but inconsistent training period. H4 handcyclist with a paraplegia. At T1: VO_2peak 1.10 L/min, POpeak 78W. Relative change in VO_2peak/kg: 6%, relative change in POpeak/kg: 15%. Training period was 33 weeks, with 2 training sessions per week on average, and average TRIMP per week of 622 AU. Training volume per training: 71 minutes, average RPE per session: 5.5. At the start of the training period this participant had a lot of pain complaints related to the spinal cord injury, not related to training. In week 18 there was a surgery on the urinary tract.

Table 5. Associations between absolute/relative change in physical capacity and frequency (training sessions per week), duration (volume per training) and intensity (sRPE per training).

	ΔPOpeak/kg (x1000)			%ΔPOpeak/kg (x1000)			ΔVO_2peak/kg (x1000)			%ΔVO_2peak/kg (x1000)		
	Beta	SE	*p*-value	Beta	SE	*p*-value	Beta	SE	*p*-value	Beta	SE	*p*-value
Constant	440.869	354.203		44393.485	20744.201		10339.873	3888.866		56269.609	19515.935	
Training period (weeks)	-1.861	5.844	0.75	-111.711	406.651	0.78	-155.332	78.520	0.048*	-502.000	326.457	0.12
Training sessions per week (N)	-5.657	26.569	0.83	-301.639	1799.464	0.87	147.580	350.198	0.67	904.359	1479.801	0.54
Volume per training (min)	-1.387	1.974	0.48	-93.830	136.853	0.49	-27.957	27.021	0.30	-57.834	111.346	0.60
sRPE per training	3.100	25.660	0.90	4335.796	4928.621	0.38	481.676	981.019	0.62	-1416.477	1495.303	0.34
Outcome at T1	11.267	68.024	0.87	-11583.516	4682.740	0.01	-76.373	59.844	0.20	-742.301	253.707	0.003

Corrected for value of the outcome measure at T1. * Significant association with $p < 0.05$. Physical capacity outcome parameters are multiplied by 1000 to visualize the details in the beta.

Table 6. Associations between absolute/relative change in physical capacity and absolute training intensity distribution.

	ΔVO₂peak (x1000)			%ΔVO₂peak (x1000)			ΔVO₂peak/kg (x1000)			%ΔVO₂peak/kg (x1000)		
	Beta	SE	*p*-value	Beta	SE	*p*-value	Beta	SE	*p*-value	Beta	SE	*p*-value
Constant	462.526	206.001		45431.036	10275.538		9793.533	2393.556		48844.377	9848.711	
(RPE 1-4) (N sessions)	1.414	1.359	0.30	58.541	67.769	0.39	9.037	16.621	0.59	52.135	68.392	0.45
(RPE 5-6) (N sessions)	4.242	1.918	0.03*	231.431	95.658	0.02*	42.902	23.722	0.07	188.965	97.608	0.05
(RPE 7-10) (N sessions)	-0.042	1.788	0.98	-66.335	89.166	0.46	-2.563	22.261	0.91	-26.922	91.597	0.77
Training period (weeks)	-15.157	8.184	0.06	-882.131	408.223	0.03	-216.074	100.870	0.03	-836.225	415.046	0.04
Outcome at T1	1.499	64.946	0.98	-8581.190	3239.586	0.008	-100.012	55.863	0.07	-762.783	229.859	<0.001
Constant	435.294	207.101		45274.884	10158.960		10032.444	2402.470		50176.226	9925.690	
(RPE 1-4) time (min)	0.001	0.018	0.96	0.309	0.865	0.72	-0.060	0.212	0.78	0.181	0.874	0.84
(RPE 5-6) time (min)	0.067	0.030	0.03*	4.072	1.450	0.005*	0.738	0.365	0.04*	3.123	1.506	0.04*
(RPE 7-10) time (min)	-0.006	0.022	0.79	-0.726	1.063	0.49	-0.121	0.266	0.65	-0.591	1.098	0.59
Training period (weeks)	-13.154	8.086	0.10	-930.988	396.665	0.02	-208.867	99.388	0.04	-825.270	410.615	0.04
Outcome at T1	7.169	65.771	0.91	-8548.169	3226.254	0.008	-106.726	56.189	0.06	-809.064	232.141	<0.001

Corrected for duration of training period (weeks) and value of the outcome parameter at T1. * Significant association with $p < 0.05$. Physical capacity outcome parameters are multiplied by 1000 to visualize the details in the beta.

Additional explorative regression analyses with high intensity training sessions only (RPE > 5), total monotony or total strain as determinants; or ΔPOpeak and ΔVO_2peak as outcome parameters, showed no significant results. Multilevel multivariate regression analyses with TID showed a significant association between ΔVO_2peak as well as %ΔVO_2peak and absolute number of training sessions and time in moderate intensity (RPE 5-6). In addition, significant associations were found between ΔVO_2peak/kg as well as %ΔVO_2peak/kg and absolute time in moderate intensity (RPE 5-6) (table 6). None of the TID parameters were associated with change in POpeak or POpeak/kg, nor relative time or training sessions were associated with the change of any of the physical capacity outcome measures.

Discussion

Physical capacity improved with 17 – 22% during 21 ± 6 weeks of training. Correlations between training characteristics and outcome parameters were not significant and total $TRIMP_{sRPE}$ was not significantly associated with change in physical capacity. In addition, the separate components of frequency, duration and intensity were not unequivocally associated with change in physical capacity. Explorative analyses showed that absolute time and number of training sessions spent in moderate intensity (zone 2) were associated with an increase in physical capacity.

Physical capacity of the participants in the present study was comparable to other HandbikeBattle studies (table 1) [4,36] and other recreational handcyclists [7]. The changes in physical capacity were comparable to changes as described in a systematic review (10 – 30% for POpeak and VO_2peak) on upper body exercise in people with a SCI [37]. An 8-week training intervention for experienced handcyclists resulted in 20 – 26% improvement in VO_2peak/kg [38], whereas a 6-week home-based arm crank exercise intervention with four sessions per week at moderate intensity showed 19% improvement in VO_2peak/kg in untrained individuals with SCI [39].

Compared with previous studies on training load in able-bodied athletes, the duration of the training period was longer (21 ± 6 weeks in present study, versus 6 - 12 weeks in previous studies) [23–26,28,40]. The mean weekly $TRIMP_{sRPE}$ was lower than for able-bodied elite cyclists (1654 ± 579 AU in present study versus 4086 ±1460 AU) [25], but higher than or comparable to weekly loads in studies on rugby, hurling and soccer [23,24,28,40].

In the present study there were no significant dose-response relationships between $TRIMP_{sRPE}$ and changes in physical capacity. This is in agreement with several previous studies with able-bodied athletes. Previous studies in team sports have shown correlations of 0.22 – 0.70 between $TRIMP_{sRPE}$

and change in maximum velocity [40,41] and correlations of 0.20 – 0.24 between $TRIMP_{sRPE}$ and change in VO_2max [24,28]. In rugby, a curvilinear relationship between $TRIMP_{sRPE}$ and VO_2max was found with an explained variance of 12% [23]. One could argue that other training load parameters such as HR-based TRIMP or TSS might have a better association with changes in physical capacity in handcycling. Two recent studies in cycling showed, however, no conclusive results. In elite cyclists, there were no significant associations among different training load parameters ($TRIMP_{sRPE}$, different HR-based TRIMP methods and TSS) and change in POmax or VO_2max [25]. In recreational cyclists, there were no significant associations among different HR-based TRIMP methods and change in POmax [26]. Although the TSS is an objective parameter of external load, it is only applicable to (hand)cycle training sessions and not to other sporting activity, which is a disadvantage. In addition, HR-based TRIMP methods cannot be used for individuals with tetraplegia due to the altered sympathetic response to exercise [16]. In contrast, the $TRIMP_{sRPE}$ is a robust measure and can be used irrespective of mode or location [14]. An interesting focus for future handcycling research would be a combination of several (objective and subjective) internal and external training load methods.

Another interesting focus for future research would be the associations among training load and changes in submaximal responses in handcycling. In the study by Sanders et al. significant associations were found between $TRIMP_{sRPE}$ and change in PO at the first lactate threshold (LT1, $r = 0.54$), and change in PO at the second lactate threshold (LT2, $r = 0.60$) in elite cycling [25]. HR-based TRIMP methods and TSS were strongly associated with change in PO at the lactate thresholds as well ($r = 0.52 – 0.81$ and $r = 0.75 - 0.79$, respectively) [25].

In previous literature in elite athletes, a low day-to-day variability in training load, that is, a high weekly monotony, was associated with a decline in performance and risk of overtraining and illness, especially when combined with a high training load, resulting in a high strain [29]. Although the monotony threshold is different for each individual, in previous research a weekly monotony above 2.0 AU was mentioned to be associated with a decline in performance and risk of overtraining [28,29], whereas a weekly monotony around 1.0 AU indicates large day-to-day variability [42]. In the present study the weekly monotony was low for most participants, with only 3.6 ± 1.4 training sessions per week. Overtraining is, therefore, unlikely in the present study, whereas undertraining cannot be excluded. Especially considering the heterogeneity of the population, undertraining is likely in participants that were not able to maintain continuity in their training regime (Figure 1 C and D).

As training load consists of a combination of volume and intensity, different combinations may result in the same training load, but in a different response. In this view, TID and its effect on performance is widely studied. The threshold-training model, that is training in moderate intensity close to the second ventilatory threshold (VT2), has shown to be a guideline for training intensity in untrained participants [43,44]. In contrast, the polarized-training model is shown to be associated with improvements in performance in elite endurance athletes. In the polarized-training model, athletes train the majority of time (e.g., 75%) in the low intensity zone below the first ventilatory threshold (VT1) and the remaining time clearly above the VT2 in the high intensity zone [26,27]. The present study suggests that the threshold-training model is also applicable to relatively untrained wheelchair users during handcycle training, as only training at moderate intensity was associated with increase in physical capacity.

A limitation of the present study is that the training period was relatively long. In future studies it would be advised to perform additional measurements, such as a GXT or time trial, after every 4 – 6 weeks of training. In this way the associations between training load and the outcome parameters could become clearer and lack of consistency in training could be accounted for. In addition, this could aid in the adherence of training monitoring. It should be noted that given the large number of participants and logistics, this set up was not possible in the current study. Monitoring training load in a large group of non-elite participants was a challenge, which becomes clear from the 120 individuals with incomplete training data.

That said, the population of the present study was heterogeneous and several (unmeasurable) factors could play a role in the interaction between training load and training adaptations. Figure 1 C and D illustrate the inconsistency of training, due to all sorts of factors, not necessarily related to the training itself. Wheelchair users are a more vulnerable population than elite able-bodied athletes. However, even in a homogeneous group of able-bodied athletes, complex (temporal, fluctuating) inter-relationships exist among load, the ability to tolerate load (i.e., load capacity), performance and health [45]. Several components that influence these inter-relationships are for example fatigue, emotional disturbances, illness or training history [14,15]. An individualized approach is necessary, as the individual's psychophysiological response (internal training load) will determine training adaptation [46]. To get a grip on all these components, an integrated approach is proposed with monitoring of objective physiological measures, RPE, stress, coping, nutrition and sleep [14,45].

Conclusions

In conclusion, the present study shows that total $\text{TRIMP}_{\text{sRPE}}$ was not associated with changes in physical capacity during handcycle training. In addition, the separate components of frequency, duration and intensity were not unequivocally associated with change in physical capacity. However, the threshold-training model is suggested to be applicable to relatively untrained wheelchair users during handcycle training, as training at moderate intensity (zone 2) was significantly associated with increase in physical capacity. Future studies should focus on an individualized integrated approach to unravel more components associated with the training response in former rehabilitation patients.

Declaration of interest

The authors report no conflicts of interest.

Acknowledgements

The authors would like to thank Carl Foster, Department of Exercise and Sport Science, University of Wisconsin, La Crosse, United States; and Michel Brink, University of Groningen, University Medical Center Groningen, Center for Human Movement Sciences, Groningen, the Netherlands; for their advice on data analysis and interpretation. The help of students and coaches with data collection is highly appreciated: Sven Hoekstra, Julia Berentschot, Sanne van den Tol and Charlie Schillemans, University of Groningen, University Medical Center Groningen, Center for Human Movement Sciences, Groningen, the Netherlands; Marco Korf, Chantal Grove and Simone van der Riet, Department of Human Movement Sciences, Faculty of Behavioural and Movement Sciences, Vrije Universiteit Amsterdam, the Netherlands; Natasja ter Veer, Faculty Kinesiology and Rehabilitation Sciences, University of Leuven, Belgium; Bennie Bouwhuis, Roessingh Rehabilitation Center, Enschede, the Netherlands.

HandbikeBattle group name: Paul Grandjean Perrenod Comtesse, Adelante Zorggroep, Hoensbroek, the Netherlands. Eric Helmantel, University Medical Center Groningen, Center for Rehabilitation Beatrixoord, Groningen, the Netherlands. Mark van de Mijll Dekker, Heliomare Rehabilitation Center, Wijk aan Zee, the Netherlands. Maremka Zwinkels, Rehabilitation Center De Hoogstraat, Utrecht, the Netherlands. Misha Metsaars, Libra Rehabilitation and Audiology, Eindhoven, the Netherlands. Lise Wilders, Sint Maartenskliniek, Nijmegen, the Netherlands. Linda van Vliet, Amsterdam Rehabilitation Research Center | Reade, Amsterdam, the Netherlands. Wilbert Snoek, Rehabilitation center Revant, Breda, the Netherlands. Karin Postma, Rijndam Rehabilitation Center, Rotterdam, the Netherlands. Bram van Gemeren, Roessingh Rehabilitation Center, Enschede, the Netherlands. Selma Overbeek, Rehabilitation center Tolbrug, Den Bosch, the Netherlands. Alinda Gjaltema, Vogellanden, Zwolle, the Netherlands.

Funding: This study was funded by HandicapNL, Stichting Mitialto, Stichting Beatrixoord Noord-Nederland, University Medical Center Groningen, Heliomare Rehabilitation Center and Stichting Handbike Events.

References

1. Van den Berg-Emons RJ, Bussmann JBJ, Haisma JA, et al. A prospective study on physical activity levels after spinal cord injury during inpatient rehabilitation and the year after discharge. *Arch Phys Med Rehabil*. 2008;89(11):2094-2101.

2. Garshick E, Kelley A, Cohen SA, et al. A prospective assessment of mortality in chronic spinal cord injury. *Spinal Cord*. 2005;43(7):408-416.

3. De Groot S, Postma K, van Vliet L, Timmermans R, Valent LJM. Mountain time trial in handcycling: Exercise intensity and predictors of race time in people with spinal cord injury. *Spinal Cord*. 2014;52(6):455-461.

4. Hoekstra SP, Valent LJM, Gobets D, van der Woude LHV, de Groot S. Effects of four-month handbike training under free-living conditions on physical fitness and health in wheelchair users. *Disabil Rehabil*. 2017;39(16):1581-1588.

5. Valent LJM, Dallmeijer AJ, Houdijk H, et al. Effects of hand cycle training on physical capacity in individuals with tetraplegia: a clinical trial. *Phys Ther*. 2009;89(10):1051-1060.

6. Arnet U, Van Drongelen S, Scheel-Sailer A, van der Woude LHV, Veeger DHEJ. Shoulder load during synchronous handcycling and handrim wheelchair propulsion in persons with paraplegia. *J Rehabil Med*. 2012;44(3):222-228.

7. Dallmeijer AJ, Zentgraaff IDB, Zijp NI, van der Woude LHV. Submaximal physical strain and peak performance in handcycling versus handrim wheelchair propulsion. *Spinal Cord*. 2004;42(2):91-98.

8. Fullerton HD, Borckardt JJ, Alfano AP. Shoulder Pain: a comparison of wheelchair athletes and nonathletic wheelchair users. *Med Sci Sport Exerc*. 2003;35(12):1958-1961.

9. Dyson-Hudson TA, Kirshblum SC. Shoulder pain in chronic spinal cord injury, Part I: Epidemiology etiology, and pathomechanics. *J Spinal Cord Med*. 2004;27(1):4-17.

10. Halson SL. Monitoring training load to understand fatigue in athletes. *Sport Med*. 2014;44(Suppl 2):S139-S147.

11. Hunter A, Coggan A, McGregor S. Beyond average power. In: *Training and Racing with a Power Meter*. 3rd edition. Boulder, CO: Velopress; 2019. p. 105-138.

12. Banister E, Calvert T. Planning for future performance: implications for long term training. *Can J Appl Sport Sci*. 1980;5(3):170-176.

13. Foster C, Florhaug JA, Franklin J, et al. A new approach to monitoring exercise training. *J Strength Cond Res*. 2001;15(1):109-115.

14. Bourdon PC, Cardinale M, Murray A, et al. Monitoring athlete training loads: consensus statement. *Int J Sports Physiol Perform*. 2017;12(Suppl 2):161-170.

15. Impellizzeri FM, Rampinini E, Coutts AJ, Sassi A, Marcora SM. Use of RPE-based training load in soccer. *Med Sci Sport Exerc*. 2004;36(6):1042-1047.

16. Valent LJM, Dallmeijer AJ, Houdijk H, et al. The individual relationship between heart rate and oxygen uptake in people with a tetraplegia during exercise. *Spinal Cord*. 2007;45(1):104-111.

17. Borges TO, Bullock N, Duff C, Coutts A. Methods for quantifying training in sprint kayak. *J Strength Cond Res*. 2014;28(2):474-482.

18. Iturricastillo A, Yanci J, Granados C, Goosey-tolfrey V. Quantifying wheelchair basketball match load: a comparison of heart-rate and perceived-exertion methods. *Int J Sports Physiol Perform*. 2016;11(4):508-514.

19. Van Erp T, Foster C, de Koning JJ. Relationship between various training-load measures in elite cyclists during training , road races, and time trials. *Int J Sports Physiol Perform*. 2019;14(4):493-500.

20. De Groot S, Hoekstra SP, Grandjean Perrenod Comtesse P, Kouwijzer I, Valent LJM. Relationships between internal and external handcycle training load in people with spinal cord injury training for the HandbikeBattle. *J Rehabil Med*. 2018;50(3):261-268.

21. Borresen J, Lambert MI. The quantification of training load, the training response and the effect on performance. *Sport Med*. 2009;39(9):779-795.

22. Foster C, Daines E, Hector L, Snyder AC, Welsh R. Athletic performance in relation to training load. *Wis Med J*. 1996;95(6):370-374.

23. Taylor RJ, Sanders D, Myers T, Abt G, Taylor CA, Akubat I. The dose-response relationship between training load and aerobic fitness in academy rugby union players. *Int J Sports Physiol Perform*. 2018;13(2):163-169.

24. Malone S, Hughes B, Collins K, Akubat I. Methods of monitoring training load and their association with changes across fitness measures in hurling players. *J Strength Cond Res*. 2020;34(1):225-234.

25. Sanders D, Abt G, Hesselink MKC, Myers T, Akubat I. Methods of monitoring training load and their relationships to changes in fitness and performance in competitive road cyclists. *Int J Sports Physiol Perform*. 2017;12(5):668-675.

26. Vermeire KM, Vandewiele G, Caen K, Lievens M, Bourgois JG, Boone J. Training progression in recreational cyclists: no linear dose-response relationship with training load. *J Strength Cond Res*. 2019;Epub ahead:1-6.

27. Seiler KS, Kjerland GØ. Quantifying training intensity distribution in elite endurance athletes: Is there evidence for an “optimal” distribution? *Scand J Med Sci Sport*. 2006;16(1):49-56.

28. Clemente FM, Clark C, Castillo D, et al. Variations of training load, monotony, and strain and dose-response relationships with maximal aerobic speed, maximal oxygen uptake, and isokinetic strength in professional soccer players. *PLoS One*. 2019;14(12):e0225522.

29. Foster C. Monitoring training in athletes with reference to overtraining syndrome. *Med Sci Sport Exerc*. 1998;30(7):1164-1168.

30. Kuipers H, Verstappen FTJ, Keizer P, Geurten P, van Kranenburg G. Variability of aerobic performance in the laboratory and its physiologic correlates. *Int J Sports Med*. 1985;6(4):197-201.

31. Garber CE, Blissmer B, Deschenes MR, et al. Quantity and quality of exercise for developing and maintaining cardiorespiratory, musculoskeletal, and neuromotor fitness in apparently healthy adults: guidance for prescribing exercise. *Med Sci Sport Exerc*. 2011;43(7):1334-1359.

32. UCI Cycling Regulations, Part 16. Para-cycling; 2017. p1-75. Available from: UCI Para-cycling. https://www.uci.org/docs/default-source/rules-and-regulations/16-par-20200211-e.pdf. Accessed: 25 Nov 2019.

33. Charlton C, Rasbash J, Brown W, et al. MLwiN Version 3.02. Centre for Multilevel Modelling, Bristol, UK. University of Bristol; 2020.

34. Hopkins W. A scale of magnitudes for effect statistics. Available at: http://sportsci.org/resource/stats/effectmag.html. 2002. Accessed: November 25, 2019.

35. Twisk JWR. *Applied Longitudinal Data Analysis for Epidemiology. A Practical Guide.* 4th ed. Cambridge: Cambridge University Press; 2003.

36. Kouwijzer I, Valent LJM, Osterthun R, van der Woude LHV, De Groot S, HandbikeBattle Group. Peak power output in handcycling of individuals with a chronic spinal cord injury: predictive modeling,

validation and reference values. *Disabil Rehabil*. 2020;42(3):400-409.

37. Valent LJM, Dallmeijer A, Houdijk H, Talsma E, van der Woude L. The effects of upper body exercise on the physical capacity of people with a spinal cord injury: a systematic review. *Clin Rehabil*. 2007;21(4):315-330.

38. Nevin J, Smith P, Waldron M, et al. Efficacy of an 8-week concurrent strength and endurance training program on hand cycling performance. *J Strength Cond Res*. 2018;32(7):1861-1868.

39. Nightingale TE, Rouse PC, Walhin J-P, Thompson D, Bilzon JL. Home-based exercise enhances health-related quality of life in persons with spinal cord injury: a randomized controlled trial. *Arch Phys Med Rehabil*. 2018;99(10):1998-2006.

40. Fitzpatrick JF, Hicks KM, Hayes PR. Dose – response relationship between training load and changes in aerobic fitness in professional youth soccer players. *Int J Sports Physiol Perform*. 2018;13(Epub ahead of print):1-6.

41. Campos-Vazquez MA, Toscano-Bendala FJ, Mora-Ferera JC, Suarez-Arrones LJ. Relationship between internal load indicators and changes on intermittent performance after the preseason in professional soccer players. *J Strength Cond Res*. 2017;31(6):1477-1485.

42. Turner AN, Bishop C, Marshall G, Read P. How to monitor training load and mode using sRPE. *Prof Strength Cond*. 2015;(39):15-20.

43. Kindermann W, Simon G, Keul J. The significance of the aerobic-anaerobic transition for the determination of work load intensities during endurance training. *Eur J Appl Physiol*. 1979;42(1):25-34.

44. Gaskill SE, Walker AJ, Serfass RA, et al. Changes in ventilatory threshold with exercise training in a sedentary population: the heritage family study. *Int J Sports Med*. 2001;22(8):586-592.

45. Verhagen E, Gabbett T. Load, capacity and health: critical pieces of the holistic performance puzzle. *Br J Sports Med*. 2019;53(1):5-6.

46. Impellizzeri FM, Marcora SM, Coutts AJ. Internal and external training load: 15 years On. *Int J Sports Physiol Perform*. 2019;14(2):270-273.

Chapter 6

Changes in quality of life during training for the HandbikeBattle and associations with cardiorespiratory fitness

Kouwijzer I, de Groot S, van Leeuwen CMC, Valent LJM, van Koppenhagen CF, HandbikeBattle Group, van der Woude LHV, Post MWM. Changes in quality of life during training for the HandbikeBattle and associations with cardiorespiratory fitness. Archives of Physical Medicine and Rehabilitation 2020;101(6):1017-1024.

Abstract

Objective: To investigate 1) changes in life satisfaction and mental health during 5 months of training for the HandbikeBattle and 4-months follow-up; 2) associations between changes in handcycling cardiorespiratory fitness and changes in life satisfaction and mental health during the training period.

Design: This is a multicenter prospective cohort study with the following measurements: the start of the training (T1), after the 5-month training period, before the event (T2), after 4-months of follow-up (T3). At T1, T2, and T3, questionnaires were filled out. At T1 and T2 a graded exercise test was performed to measure cardiorespiratory fitness (peak oxygen uptake (VO_2peak) and peak power output (POpeak)).

Setting: Ten Dutch rehabilitation centers training for the HandbikeBattle event.

Participants: Patients with a rehabilitation history (N=136) and health conditions such as spinal cord injury, amputation, or multiple trauma history.

Interventions: Not applicable.

Main outcome measure: Life satisfaction as the sum score of 2 questions (range, 2-13), and the Mental Health subscale of the 36-item Short Form Health Survey (range, 0-100).

Results: Multilevel regression analyses showed that life satisfaction increased during the training period and did not significantly change during follow-up (mean ± SD, T1: 8.2 ± 2.2, T2: 8.6 ± 2.3, T3: 8.5 ± 2.4). Mental health showed no change over time (T1: 77.7 ± 14.5, T2: 77.8 ± 14.5, T3: 75.7 ± 16.5). An improvement in cardiorespiratory fitness was associated with an increase in life satisfaction (POpeak: β=0.014, p=0.046, VO_2peak: β=1.068, p=0.04). There were no associations between improvement in cardiorespiratory fitness and an increase in mental health (POpeak: p=0.66, VO_2peak: p=0.33).

Conclusions: This study shows a positive course of life satisfaction during training for the HandbikeBattle. An improvement in cardiorespiratory fitness was longitudinally associated with an increase in life satisfaction. Mental health showed no changes over time.

Keywords: Longitudinal study; Quality of life; Handcycling; Exercise; Physical capacity

Introduction

Handcycling is a popular mode of daily transportation and exercise in wheelchair users with, for example, a spinal cord injury (SCI), leg amputation or muscular disease [1]. Compared with wheelchair propulsion, handcycling has a higher efficiency [2] and lower mechanical strain [3], which possibly reduces the risks of upper-body overuse injuries. Moreover, with handcycling it is easier to cover large distances at higher speed compared with wheelchair propulsion. Therefore, handcycling has been introduced as an exercise mode to increase cardiorespiratory fitness during and after rehabilitation [4,5]. An increase in cardiorespiratory fitness is necessary for most wheelchair users, as their fitness is generally low [6–8]. Low cardiorespiratory fitness is associated with a high prevalence of cardiometabolic disease, which is the leading cause of mortality in this population [9,10].

In addition to the potential of reducing cardiometabolic risk factors, it has been shown that improvements in cardiorespiratory fitness are associated with increased quality of life (QoL) [11,12]. Moreover, it has been shown that, for example in people with SCI, life satisfaction (LS) is reduced and mental health problems are more common compared with the general population [13–15]. Therefore, interventions to increase QoL are important.

Participating in exercise and sports might be such an intervention [16]. Several studies showed positive associations between participation in sports and QoL in individuals with physical disabilities, such as SCI, amputation, cerebral palsy and neuromuscular disease;[12,17–24] they reported less stress, depressive symptoms, pain, and a higher satisfaction with physical functioning [12]. For example, a lifestyle intervention in adults with cerebral palsy resulted in an increase in health-related QoL (mental health domain) compared with the control group, with cardiorespiratory fitness (peak oxygen consumption (VO_2peak)) explaining 22.6% of this increase [24]. Most studies investigated cross-sectional associations, and longitudinal studies are, unfortunately, scarce. A previous longitudinal study showed that an increase in cardiorespiratory fitness was associated with an increase in LS in individuals with SCI [11]. Possible mechanisms to explain these longitudinal associations are intermediate effects such as a decrease in pain [12,21], and increase in self-efficacy [21], social integration [12,21] and functional independence [19,25].

Recognizing the importance of cardiorespiratory fitness and considering the potential positive effects of handcycling, since 2013, the HandbikeBattle is organized to promote handcycling among Dutch patients with a history of rehabilitation [26]. The HandbikeBattle is an annual uphill handcycling race in the mountains of Austria among teams of rehabilitation centers. The event was created to initiate an active lifestyle by means of handcycling with peers and to push the participant’s physical and mental boundaries. A previous publication showed positive effects of training for the

HandbikeBattle on cardiorespiratory fitness and health [5]. However, longitudinal changes in QoL (LS and mental health [13]) and the associations with cardiorespiratory fitness are unknown. Therefore, the purpose of this study was to examine: 1) changes in LS and mental health during 5 months of training prior to the HandbikeBattle and at 4-months follow-up, and 2) the associations among changes in handcycling cardiorespiratory fitness and changes in LS and mental health during the training period. It was hypothesized that LS and mental health would increase over time and that this increase would be associated with an increase in cardiorespiratory fitness.

Methods

The HandbikeBattle project

The HandbikeBattle event is a serious challenge (20.2-km length and 863 m elevation gain). At the start of the training period, most participants are relatively untrained handcyclists. Guidance during the training period is provided by therapists from the respective rehabilitation centers, but otherwise the training period is free-living, that is, no specific training program is provided by the researchers. Connected to, but not part of, the HandbikeBattle is an observational cohort study that was initiated to monitor effects of participation in the training period and the event.

Participants

Inclusion criteria for the HandbikeBattle event were: 1) being a former rehabilitation patient from 1 of the 10 rehabilitation centers; 2) impairment of the lower extremities (e.g., SCI, amputation, cerebral palsy or spina bifida); and 3) commitment to the HandbikeBattle challenge. The exclusion criterion included any contra-indications to participate in the HandbikeBattle as diagnosed during the medical screening. Additional inclusion criteria for the HandbikeBattle study were first time of participation in the HandbikeBattle event and sufficient knowledge of the Dutch language to understand the instructions. In the present study, data were used from participants of the HandbikeBattle 2013-2016 cohorts. In total, 187 individuals started training for the event in this period. Twenty-one individuals dropped out due to motivational problems (N=9), medical reasons (N=5), overuse injuries (N=2), or unknown reasons (N=5). Thirty further individuals did not fill out questionnaires at any time point (N=10), or only at 1 time point (N=20). Hence, data from 136 participants were used in the present study.

Procedure

The study has a prospective design. Measurements are performed at the start of the training period (January, T1); after the training period, prior to the event (June, T2); and at follow-up, 4 months after the event (October, T3) (figure 1). At T1, T2 and T3 participants were asked to fill out questionnaires on LS and mental health. At T1 and T2, participants were asked to fill out a questionnaire about musculoskeletal pain. The questionnaires were part of a set of questionnaires which took participants 30 – 45 minutes to complete in total. Participants were invited by e-mail with a link and could fill out the questionnaires in their own time at home. At T1, a medical screening was performed by a rehabilitation physician or sports physician at the rehabilitation center. The screening was comprised of a medical anamnesis, physical examination and a handcycling or arm crank graded exercise test (GXT). At T2, the GXT was repeated with the same protocol and equipment. All participants voluntarily signed an informed consent form. The study was approved by the Local Ethics Committee of the Center for Human Movement Sciences, University Medical Center Groningen, the Netherlands (ECB/2012_12.04_l_rev/Ml).

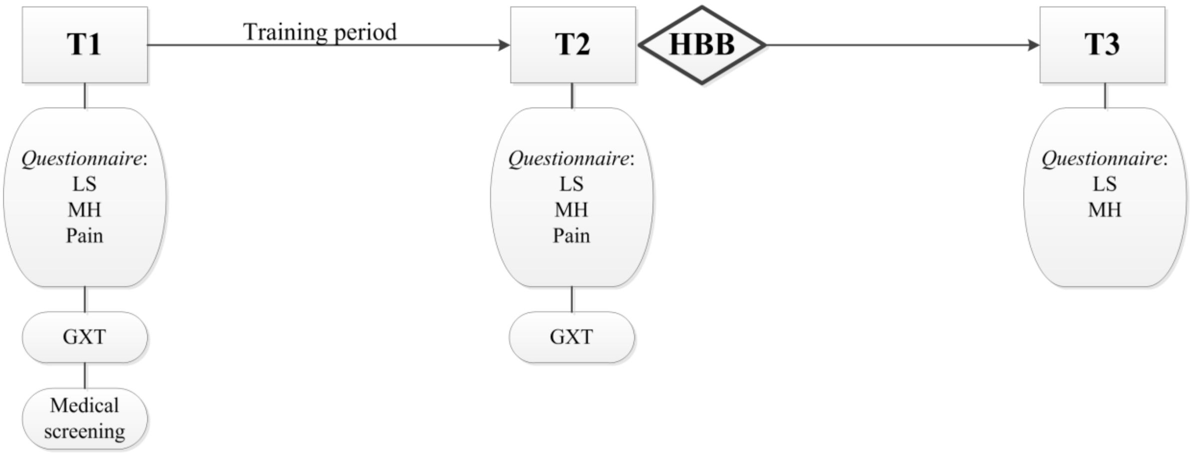


Figure 1. Study design. Measurements are performed at the start of the training period (January, T1); after the training period, prior to the event (June, T2); and follow-up, 4 months after the event (October, T3). Abbreviations: HBB, HandbikeBattle event; MH, mental health.

Measurement instruments

General participant information was collected at T1 and included age (y), sex (male or female), and impairment type (SCI, amputation, multiple traumas, spina bifida, or other).

Outcome measures

LS was assessed with the 2LS, consisting of 2 questions [27,28]. Question 1 stated satisfaction with life as a whole at the moment (LS Now, 1=very unsatisfying, 6=very satisfying). Question 2 stated satisfaction with life now compared with life before the onset of the condition (LS Comparison, range 1=much worse, 7=much better). The sum score (LS) ranges from 2-13. LS was only calculated for participants with an acquired impairment, as LS Comparison was not collected from individuals with a congenital impairment. The validity of the LS questions was supported [28].

Mental health was assessed with the Mental Health Inventory (MHI-5) of the 36-item Short Form Health Survey [29,30]. The MHI-5 consists of 5 items concerning nervousness, sadness, peacefulness, mood and happiness on a 6-point scale, a sum score was calculated ranging from 0 (lowest mental health) to 100 (highest mental health) [29]. A cutoff point of ≤72 refers to mental health problems, and a cutoff point of ≤60 refers to severe mental health problems [31,32]. Mental health showed a non-normal distribution (skewness -0.76, SE 0.13). Therefore, logarithmically transformed scores were used in the statistical analyses. The MHI-5 showed reliability and validity in persons with SCI [30].

Determinants

Cardiorespiratory fitness was measured during an incremental handcycling or arm crank GXT to volitional exhaustion at T1 and T2, organized in and conducted by the staff of each of the participating rehabilitation centers. Details on equipment and testing protocols have been described in a previous study.[33] Either a 1-minute step protocol or continuous ramp protocol was used, and was individualized for each participant. The set-up and protocol choice were dependent on the available equipment in the rehabilitation centers, but were consistent within participants over time. Power output (PO) in Watts and oxygen consumption in liters per minute were measured during the test. For the 1-minute protocol, POpeak was defined as the highest PO that was maintained for at least 30 seconds. For the ramp protocol, the highest PO achieved during the test was considered POpeak. Peak VO_2 was defined as the highest 30-second average for oxygen consumption.

Possible confounding variables

Possible confounding variables were age, sex, musculoskeletal pain, and handcycling classification. Musculoskeletal pain comprised 7 locations: the left and right hand and wrist, the left and right elbow, the left and right shoulder, and the neck. The pain was graded on a range from 1=no pain to 6=very severe pain. Having moderate-severe pain was defined as ≥4 (moderate pain) at ≥ 1 locations.

Handcycling classification was used as a proxy for severity of impairment and determined by an Union Cycliste Internationale certified paracycling classifier, following the Union Cycliste Internationale Para-cycling Regulations. The classification resulted in 5 different classes, ranging from H1 (most impaired) to H5 (least impaired) [34]. H1 and H2 handcyclists have limitations in arm-hand function, whereas H3, H4 and H5 handcyclists have intact arm-hand function and limitations in the trunk or lower extremities only. For the analyses in the present study, participants were divided in 2 large groups: (1) H1-H3 and (2) H4-H5.

Statistical analysis

The analyses were performed using SPSS (IBM SPSS Statistics 20, SPSS, Inc., Chicago, IL) and MLwiN Version 2.02 [35]. Descriptive statistics were calculated for outcome measures and determinants. Data were tested for normality with the Kolmogorov–Smirnov test with Lilliefors significance correction and the Shapiro–Wilk test, combined with z-scores for skewness and kurtosis. To ascertain possible response bias, characteristics of included participants in the present study (N=136) were compared with nonparticipants (N=51) using *t* tests, Mann-Whitney *U* tests and chi-squared tests.

Three-level multilevel models were created with observations within participants as first level, participant as second level, and rehabilitation center as third level to make adjustments for the dependency of the observations within participants, and participants within centers [36].

To answer the first research question on the longitudinal trajectory of LS and mental health over time, 2 models were created: one with LS and one with mental health as dependent variable. Time (T1, T2, T3) was included as a categorical variable with 2 dummies and T2 as reference category.

To answer the second research question on the longitudinal associations, 4 models were created, one for each of the combinations of POpeak or VO_2peak as independent variables and LS or mental health as dependent variables (T1, T2). To be able to distinguish between the between-subject component and the within-subject component, the 4 models were created as hybrid models [37]. The between-subject component reflects the cross-sectional association between 2 variables. The within-subject component reflects the longitudinal association; for example, an increase in cardiorespiratory fitness is associated with an increase in LS. In a standard longitudinal model, there is one regression coefficient that reflects both components of the association, whereas in the present study we are mostly interested in the within-subject component. A hybrid model accounts for this by providing separate regression coefficients with standard errors for each component [37]. Possible confounding variables including age, sex (reference: male), musculoskeletal pain (reference: no or mild) and handcycling classification (reference H1-H3) were added separately to each model first. A variable was included as confounder in the final model if its inclusion changed the regression coefficient with more than 10% [36].

Results

Participants were on average older, had a paraplegia more often, and had higher scores for mental health, than nonparticipants (table 1). Descriptive data of the outcome measures and determinants over time are depicted in table 2. Participants were classified with the following distribution: H1, N=7; H2, N=4; H3, N=49; H4, N=57; H5, N=19.

Table 1. Characteristics and outcomes at T1 for participants and nonparticipants.

Characteristics	***N***	**Participants**		***N***	**Nonparticipants**	
Sex (male/female) (%)	136	109/27	(80/20)	51	40/11	(78/22)
Mean age (y) ± SD	136	41	(13)	49	35*	(12)
Impairment type, n (%)	136			44		
Spinal cord injury		96	(71)		31	(71)
Tetraplegia		10	(8)		8 *	(18)
Paraplegia		86	(63)		23 *	(53)
Amputation		17	(13)		3	(7)
Multi trauma		2	(1)		2	(4)
Spina bifida		6	(4)		2	(4)
Other		15	(11)		6	(14)
Mean LS ± SD T1 (range, 2 – 13)	110	8.2	(2.2)	20	7.7	(2.4)
Mean mental Health ± SD T1 (range, 0 – 100)	122	77.7	(14.5)	20	68.2*	(18.1)
Mean POpeak (W) ± SD T1	107	126	(37)	32	113	(32)
Mean VO_2peak (L/min) ± SD T1	123	2.01	(0.57)	36	1.85	(0.40)
Musculoskeletal pain (no-mild/moderate-severe), n (%) T1	121	74/47	(61/39)	20	9/11	(45/55)
Handcycling classification (H1–H3/H4–H5) (%)	136	60/76	(44/56)	33	15/18	(46/54)

Data are represented as *N* (%) or mean (SD). Musculoskeletal pain has 2 categories: (1) no-mild pain and (2) moderate-severe pain. Handcycling classification has 2 categories: (1) H1–H3 and (2) H4–H5. * Significant difference with $p < 0.05$ between participants and nonparticipants.

Longitudinal trajectory of LS and mental health

LS showed a significant increase between T1 (start of training) and T2 (after training), and did not significantly change between T2 and T3 (follow-up) (table 3). When the model was recalculated with T1 as reference category, there was still a significant increase between T1 and T2, but no significant change between T1 and T3 (regression coefficient 0.333, SE 0.183, *p*=0.07).

Mental health showed no significant change over time (table 3). When the model was recalculated with T1 as reference category, there were no changes over time. However, a subgroup analysis with participants with mental health problems at T1 (score ≤72, N=38) showed a significant increase between T1 and T2, and no significant change between T2 and T3 (see table 3). In addition, there was a significant increase between T1 and T3 (regression coefficient 0.069, SE 0.028, *p*=0.01). The mean scores ± SD were as follows: T1: 60.0±10.1; T2: 67.1±14.7; T3: 63.4±15.0.

Table 2. Descriptive data and outcome measures of participants at all time points.

	N	**T1**		***N***	**T2**		***N***	**T3**	
LS	110	8.2	(2.2)	110	8.6	(2.3)	100	8.5	(2.4)
Mental health	122	77.7	(14.5)	124	77.8	(14.5)	109	75.7	(16.5)
POpeak (W)	107	126	(37)	105	141	(46)			
VO_2peak (L/min)	123	2.01	(0.57)	112	2.13	(0.63)			
Musculoskeletal pain (no-mild/moderate-severe) (%)	121	74/47	(61/39)	124	88/36	(71/29)			

Data are represented as *N* (%) or mean (SD). Musculoskeletal pain has 2 categories: (1) no-mild pain and (2) moderate-severe pain. T1 represents the start of the training period. T2 represents after the training period, prior to the HandbikeBattle event. T3 represents the follow-up measurement, 4 months after training.

Table 3. Longitudinal trajectory of LS and mental health.

	LS (N=124)			**Mental health (N=133)**			**Mental health subgroup (N=38)**		
	Regression coefficient	**SE**	***p*-value**	**Regression coefficient**	**SE**	***p*-value**	**Regression coefficient**	**SE**	***p*-value**
Constant (reference T2)	8.610	0.206		1.116	0.024		0.920	0.029	
Δ T2 – T1	-0.502	0.146	< 0.01*	-0.018	0.020	0.37	-0.099	0.028	< 0.01*
Δ T2 – T3	-0.169	0.183	0.36	-0.023	0.027	0.39	-0.032	0.032	0.32

T1 represents the start of the training period. T2 represents after the training period, prior to the HandbikeBattle event. T3 represents the follow-up measurement, 4 months after training. The transformed value of mental health was used for analysis. Mental health subgroup represents participants with score ≤ 72 at T1. The Δ T2 – T1 = a negative regression coefficient representing an improvement of the dependent variable over time. The Δ T2 – T3 = a negative regression coefficient representing a deterioration of the dependent variable over time. * Significance with $p < 0.05$.

Table 4. Basic models without confounders: longitudinal and cross-sectional associations between cardiorespiratory fitness and QoL (T1, T2).

	LS			**Mental health**			**Mental health subgroup**		
	Regression coefficient	**SE**	***p*-value**	**Regression coefficient**	**SE**	***p*-value**	**Regression coefficient**	**SE**	***p*-value**
POpeak (W)	(N=84)			(N=87)			(N=25)		
Constant	8.409	0.851		0.954	0.088		0.690	0.095	
Between-subject	0.000	0.006	1.00	0.001	0.001	0.10	0.001	0.001	0.06
Within-subject	0.017	0.007	0.02*	0.001	0.001	0.37	0.005	0.002	< 0.01*
VO$_2$peak (L/min)	(N=97)			(N=99)			(N=31)		
Constant	6.869	0.819		0.864	0.093		0.750	0.088	
Between-subject	0.741	0.385	0.05	0.103	0.044	0.02*	0.054	0.043	0.21
Within-subject	1.315	0.507	< 0.01*	0.091	0.061	0.14	0.240	0.113	0.03*

The transformed value of mental health was used for analysis. Mental health subgroup represents participants with score ≤ 72 at T1. * Significance with $p < 0.05$.

Table 5. Final models with confounders: longitudinal and cross-sectional associations between cardiorespiratory fitness and QoL (T1, T2).

	LS			**Mental health**			**Mental health subgroup**		
	Regression coefficient	**SE**	***p*-value**	**Regression coefficient**	**SE**	***p*-value**	**Regression coefficient**	**SE**	***p*-value**
POpeak (W)	(N=84)			(N=87)			(N=25)		
Constant	7.788	0.961		0.926	0.127		0.721	0.092	
Between-subject	0.004	0.007	0.57	0.001	0.001	0.32	0.001	0.001	0.03*
Within-subject	0.014	0.007	0.046*	0.000	0.001	0.66	0.004	0.002	0.01*
Confounders									
Age	-	-		0.003	0.002	0.14	-	-	
Sex	0.886	0.735	0.23	-0.043	0.069	0.54	-	-	
Musculoskeletal pain	-0.669	0.381	0.08	-0.098	0.040	0.01	-0.095	0.050	0.06
Classification	0.472	0.481	0.33	-	-		-	-	
VO_2peak (L/min)	(N=97)			(N=99)			(N=31)		
Constant	6.644	0.904		0.974	0.107		0.750	0.088	
Between-subject	0.904	0.406	0.03*	0.073	0.047	0.12	0.054	0.043	0.21
Within-subject	1.068	0.516	0.04*	0.060	0.062	0.33	0.240	0.113	0.03*
Confounders									
Age	-	-		-	-		-	-	
Sex	0.807	0.630	0.20	-0.071	0.068	0.29	-	-	
Musculoskeletal pain	-0.708	0.332	0.03	-0.101	0.038	< 0.01	-	-	
Classification	-	-		-	-		-	-	

The transformed value of mental health was used for analysis. Mental health subgroup represents participants with score ≤ 72 at T1. POpeak: peak power output; VO_2peak: peak oxygen uptake. Sex: reference male. Musculoskeletal pain: two categories: (1) no-mild pain and (2) moderate-severe pain (reference: no-mild). Handcycling classification: two categories: (1) H1–H3 and (2) H4–H5 (reference: H1-H3). A variable was included as confounder if the regression coefficient of POpeak or VO_2peak changed more than 10%. * Significance with $p < 0.05$.

Longitudinal associations between cardiorespiratory fitness and LS

Peak PO as well as VO_2peak showed a significant within-subject component association with LS (table 4). After adding confounders to the model with POpeak, the within-subject component association between POpeak and LS remained significant (table 5). After adding confounders to the model with VO_2peak, both the within-subject and between-subject component associations between VO_2peak and LS were significant (table 5).

Longitudinal associations between cardiorespiratory fitness and mental health

POpeak showed no significant association with mental health (see table 4). Peak VO_2 showed a significant between-subject, but no within-subject component association with mental health (table 4). After adding confounders to the models, there were no significant associations among cardiorespiratory fitness and mental health (see table 5). However, a subgroup analysis with participants with mental health problems at T1 (score ≤72, N=38) showed that POpeak had a significant between-subject and within-subject component association with mental health (see table 5). Peak VO_2 showed a only significant within-subject component association (see table 5).

Discussion

The present study showed that LS increased during 5 months of training for the HandbikeBattle and did not significantly change during follow-up. Life satisfaction at follow-up was, however, not significantly higher than LS at the start of the training period. Mental health showed no significant change over time for the total group. The results of this study further showed that, over time, an improvement in cardiorespiratory fitness was associated with an increase in LS. There were no associations between improvement in cardiorespiratory fitness and increase in mental health for the total group, however, those with low mental health showed a positive longitudinal association.

The results with respect to LS are in line with a previous longitudinal study in which a positive association was found between cardiorespiratory fitness and LS measured from the start of active rehabilitation until 5 years after discharge in individuals with an SCI [11]. Previously mentioned underlying mechanisms for the association between cardiorespiratory fitness and QoL include enhanced feelings of self-worth, self-efficacy and personal control, increased social interaction, reduced pain, and increased production of neurotransmitters regulating emotions.[20] Other intermediate effects might be the positive effect of an increase in cardiorespiratory fitness on functional independence and activities of daily living,[19,27,38,39] or body image and satisfaction [40,41].

In contrast to previous studies,[12,20,24] mental health did not significantly change over time. A possible explanation for this lack of change is that mental health scores were already high at baseline, the average of 77.7 is even above the mean score of 76.8 in the general Dutch population.[42] A similar mean score of 77.2 was found among Dutch individuals with an SCI 5 years after discharge from initial inpatient rehabilitation [30,43]. Although there was no ceiling effect in the present study, this high baseline score might be less susceptible to change. In addition, mental health was not significantly associated with cardiorespiratory fitness. It might be that mental health is less susceptible to changes in cardiorespiratory fitness and physical functioning than LS. In previous studies, LS was positively associated with functional independence,[27,39] whereas mental health was not [43,44]. Psychological factors such as self-efficacy, mastery, acceptance and purpose in life might be more important determinants to improve mental health [44–46].

Nevertheless, a subgroup analysis of participants with a mental health score ≤72 at T1 showed that there was an improvement over time and that this improvement was associated with an increase in cardiorespiratory fitness. Although it is a small group and regression to the mean could have played a role, these findings suggest that an increase in cardiorespiratory fitness might have a positive effect on mental health in individuals with lower mental health scores at baseline. There seems to be only a thin line, as nonparticipants also had a low mental health score. More intensive (mental health) guidance might be warranted for individuals with low mental health; this may prevent them from dropping out during the training period.

Implications and future studies

The present study is one of few studies that examined longitudinal associations between cardiorespiratory fitness and QoL. Unique aspects are that the training is free-living with peers and that participants are training towards a goal. In future studies it would be interesting to examine which training regimes led to these improvements in cardiorespiratory fitness and to provide long term follow-up results on the association between cardiorespiratory fitness and QoL in individuals participating in the HandbikeBattle. In addition, it would be interesting to unravel in more detail for which individuals in or after rehabilitation an increase in cardiorespiratory fitness would be most effective in increasing QoL and by which underlying mechanisms this could be achieved. Moreover, other determinants that could potentially affect QoL, such as peer support or psychological factors, should be investigated. The use of questionnaires or interviews would be valuable to gain knowledge about benefits (and adverse effects) of participating in and training for such an event.

Study limitations

There might be a selection bias, as individuals who dropped out before or during the training period were not included in the analyses. Therefore, individuals with a very low cardiorespiratory fitness might be less prominent in the present study. In addition, it might be possible that individuals with a very low mental health or LS, are not willing to participate in an event like the HandbikeBattle. It is, however, important to mention that the event was created to push physical and mental boundaries for individuals who might need this, and that participants were not professional athletes. Lastly, only the association between cardiorespiratory fitness and QoL was investigated. Other possible determinants of QoL should be addressed in future studies.

Conclusion

LS increased during the training period and did not significantly change during follow-up. Mental health showed no change over time for the total group of participants. The present study showed that cardiorespiratory fitness and LS are longitudinally associated, that is, an improvement in cardiorespiratory fitness was associated with an increase in LS.

Conflicts of interest

The authors report no conflict of interest.

Acknowledgments

This study was funded by HandicapNL, Stichting Mitialto, Stichting Beatrixoord Noord-Nederland, University Medical Center Groningen, Heliomare Rehabilitation Center and Stichting Handbike Events.

We thank Jos Twisk, PhD, Department of Epidemiology and Biostatistics, UMC Amsterdam, Amsterdam, the Netherlands for his support with the statistics, and Sjoerd Jansman, MSc, University of Groningen, University Medical Center Groningen, Center for Human Movement Sciences, Groningen, the Netherlands for his support with data processing.

HandbikeBattle group name: Paul Grandjean Perrenod Comtesse, Adelante Zorggroep, Hoensbroek, the Netherlands. Eric Helmantel, University Medical Center Groningen, Center for Rehabilitation Beatrixoord, Groningen, the Netherlands. Mark van de Mijll Dekker, Heliomare Rehabilitation Center, Wijk aan Zee, the Netherlands. Maremka Zwinkels, Rehabilitation Center De Hoogstraat, Utrecht, the Netherlands. Misha Metsaars, Libra Rehabilitation and Audiology, Eindhoven, the Netherlands. Lise Wilders, Sint Maartenskliniek, Nijmegen, the Netherlands. Linda van Vliet, Amsterdam Rehabilitation Research Center | Reade, Amsterdam, the Netherlands. Karin Postma, Rijndam Rehabilitation Center, Rotterdam, the Netherlands. Bram van Gemeren, Roessingh Rehabilitation Center, Enschede, the Netherlands. Alinda Gjaltema, Vogellanden, Zwolle, the Netherlands.

References

1. Abel T, Vanlandewijck Y, Verellen J. Handcycling. In: Goosey-Tolfrey V, editor. Wheelchair sport. Champaign, IL: Human kinetics; 2010. p. 187–197.

2. Dallmeijer AJ, Zentgraaff IDB, Zijp NI, van der Woude LHV. Submaximal physical strain and peak performance in handcycling versus handrim wheelchair propulsion. Spinal Cord 2004;42(2):91-98.

3. Arnet U, van Drongelen S, Scheel-Sailer A, van der Woude LHV, Veeger DHEJ. Shoulder load during synchronous handcycling and handrim wheelchair propulsion in persons with paraplegia. J Rehabil Med 2012;44(3):222-228.

4. Valent LJM, Dallmeijer AJ, Houdijk H, et al. Effects of hand cycle training on physical capacity in individuals with tetraplegia: A clinical trial. Phys Ther 2009;89(10):1051-1060.

5. Hoekstra SP, Valent LJM, Gobets D, van der Woude LHV, de Groot S. Effects of four-month handbike training under free-living conditions on physical fitness and health in wheelchair users. Disabil Rehabil 2017;39(16):1581-1588.

6. Van den Berg-Emons RJ, Bussmann JBJ, Haisma JA, et al. A prospective study on physical activity levels after spinal cord injury during inpatient rehabilitation and the year after discharge. Arch Phys Med Rehabil 2008;89(11):2094-2101.

7. Haisma JA, van der Woude LH V, 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(11):642-652.

8. Jacobs PL, Nash MS. Exercise recommendations for individuals with spinal cord injury. Sports Med 2004;34(11):727-751.

9. Garshick E, Kelley A, Cohen SA, et al. A prospective assessment of mortality in chronic spinal cord injury. Spinal Cord 2005;43(7):408-416.

10. Nightingale TE, Metcalfe RS, Vollaard NB, Bilzon JL. Exercise guidelines to promote cardiometabolic health in spinal cord injured humans: time to raise the intensity? Arch Phys Med Rehabil 2017;98(8):1693-1704.

11. Van Koppenhagen CF, Post MWM, de Groot S, 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(3):328-337.

12. Hicks AL, Martin KA, Ditor DS, et al. Long-term exercise training in persons with spinal cord injury: effects on strength, arm ergometry performance and psychological well-being. Spinal Cord 2003;41(1):34-43.

13. Post MWM, van Leeuwen CMC. Psychosocial issues in spinal cord injury: a review. Spinal Cord 2012;50(5):382-389.

14. Van Koppenhagen CF, Post MWM, van der Woude LHV, et al. Changes and determinants of life satisfaction after spinal cord injury: a cohort study in The Netherlands. Arch Phys Med Rehabil 2008;89(9):1733-40.

15. Dijkers MPJM. Quality of life of individuals with spinal cord injury: A review of conceptualization, measurement, and research findings. JRRD 2005;42(3):87-110.

16. Chekroud SR, Gueorguieva R, Zheutlin AB, et al. Association between physical exercise and mental health in

1.2 million individuals in the USA between 2011 and 2015: a cross-sectional study. The Lancet Psychiatry 2018;5(9):739-746.

17. Anneken V, Hanssen-Doose A, Hirschfeld S, Scheuer T, Thietje R. Influence of physical exercise on quality of life in individuals with spinal cord injury. Spinal Cord 2010;48(5):393-399.

18. Ditor DS, Latimer AE, Ginis KAM, Arbour KP, Mccartney N, Hicks AL. Maintenance of exercise participation in individuals with spinal cord injury: effects on quality of life, stress and pain. Spinal Cord 2003;41(8):446-450.

19. Sweet SN, Ginis KAM, Tomasone JR. Investigating intermediary variables in the physical activity and quality of life relationship in persons with spinal cord injury. Health Psychology 2013;32(8):877-885.

20. Ginis KAM, Jetha A, Mack DE, Hetz S. Physical activity and subjective well-being among people with spinal cord injury: a meta-analysis. Spinal Cord 2010;48(1):65-72.

21. Motl RW, McAuley E. Pathways between physical activity and quality of life in adults with multiple sclerosis. Health Psychology 2009;28(6):682-689.

22. Krops LA, Jaarsma EA, Dijkstra PU, Geertzen JHB, Dekker R. Health related quality of life in a Dutch rehabilitation population: reference values and the effect of physical activity. PLoS ONE 2017;12(1): e0169169.

23. Bragaru M, Dekker R, Geertzen JHB, Dijkstra PU. Amputees and Sports. A Systematic Review. Sports Med 2011;41(9):721-740.

24. Slaman J, van den Berg-Emons HJG, van Meeteren J, et al. A lifestyle intervention improves fatigue, mental health and social support among adolescents and young adults with cerebral palsy: focus on mediating effects. Clin Rehabil 2015;29(7): 717-727.

25. Haisma JA, Post MWM, van der Woude LHV, et al. Functional independence and health-related functional status following spinal cord injury: a prospective study of the association with physical capacity. J Rehabil Med 2008;40(10):812-818.

26. Groot de S, Postma K, van Vliet L, Timmermans R, Valent LJM. Mountain time trial in handcycling: Exercise intensity and predictors of race time in people with spinal cord injury. Spinal Cord 2014;52(6):455-461.

27. Van Koppenhagen CF, Post MWM, van der Woude LHV, et al. Recovery of life satisfaction in persons with spinal cord injury during inpatient rehabilitation. Am J Phys Med Rehabil 2009;88(11):887–895.

28. Post MWM, van Leeuwen CMC, van Koppenhagen CF, de Groot S. Validity of the Life Satisfaction Questions, the Life Satisfaction Questionnaire, and the Satisfaction With Life Scale in persons with spinal cord injury. Arch Phys Med Rehabil 2012;93(10):1832-7.

29. Ware JE, Snow KK, Kosinski M, Gandek B. Chapter 6: scoring the SF-36. In: SF-36 Health Survey: manual and interpretation guide. Boston: The Health Institute, New England Medical Center; 1993. p1-22.

30. van Leeuwen CMC, van der Woude LHV, Post MWM. Validity of the mental health subscale of the SF-36 in persons with spinal cord injury. Spinal Cord 2012; 50(9): 707–710;

31. Perenboom R, Oudshoorn K, van Herten L, Hoeymans N, Bijl R. Life expectancy in good mental health: establishing cut-offs for the MHI-5 and GHQ-12 [Dutch]. Leiden: TNO Publications; 2000.

32. Hoeymans N, Garssen AA, Westert GP, Verhaak PFM. Measuring mental health of the Dutch population: a

comparison of the GHQ-12 and the MHI-5. Health Qual Life Outcomes 2004, 2:23.

33. Kouwijzer I, Valent LJM, Osterthun R, van der Woude LHV, de Groot S, HandbikeBattle group. Peak power output in handcycling of individuals with a chronic spinal cord injury: predictive modeling, validation and reference values. Disabil Rehabil. 2020;42(3):400-409.

34. Union Cycliste Internationale. Cycling regulations, part 16 para-cycling. Available at: https://www.uci.org/docs/default-source/rules-and-regulations/16-par-20200211-e.pdf. Accessed November 25, 2018.

35. Rasbash J, Charlton C, Browne W, et al. MLwiN Version 2.02. Centre for Multilevel Modelling. Bristol (UK): University of Bristol; 2005.

36. Twisk JWR. Applied Longitudinal Data Analysis for Epidemiology. A Practical Guide. 4th ed. Cambridge: Cambridge University Press; 2003.

37. Twisk JWR, De Vente W. Hybrid models were found to be very elegant to disentangle longitudinal within- and between-subject relationships. J Clin Epidemiol 2019;107:66-70.

38. Van Leeuwen CMC, Post MWM, Hoekstra T, et al. Trajectories in the course of life satisfaction after spinal cord injury: identification and predictors. Arch Phys Med Rehabil 2011;92(2):207-213.

39. Van Leeuwen CMC, Post MWM, van Asbeck FWA, et al. Life satisfaction in people with spinal cord injury during the first five years after discharge from inpatient rehabilitation. Disabil Rehabil 2012;34(1): 76-83.

40. Bassett RL, Ginis KAM, Buchholz AC, SHAPE SCI Research Group. A pilot study examining correlates of body image among women living with SCI. Spinal Cord 2009; 47(6): 496–498.

41. Reboussin BA, Rejeski WJ, Martin KA, et al. Correlates of satisfaction with body function and body appearance in middle- and older aged adults: The activity counseling trial (ACT). Psychology & Health 2000;15(2): 239-254.

42. Aaronson NK, Muller M, Cohen PDA, et al. Translation, validation, and norming of the Dutch language version of the SF-36 Health Survey in community and chronic disease populations. J Clin Epidemiol 1998;51(11):1055–1068.

43. Van Leeuwen CMC, Hoekstra T, van Koppenhagen CF, de Groot S, Post MWW. Trajectories and predictors of the course of mental health after spinal cord injury. Arch Phys Med Rehabil 2012;93(12):2170-6.

44. Van Leeuwen CMC, Post MWM, Westers P, et al. Relationships between activities, participation, personal factors, mental health, and life satisfaction in persons with spinal cord injury. Arch Phys Med Rehabil 2012;93(1):82-9.

45. Van Leeuwen CMC, Kraaijeveld S, Lindeman E, Post MWM. Associations between psychological factors and quality of life ratings in persons with spinal cord injury: a systematic review. Spinal Cord 2012; 50(3), 174–187.

46. Van Leeuwen CMC, Edelaar-peeters Y, Peter C, Stiggelbout AM, Post MWM. Psychological factors and mental health in persons with spinal cord injury: an exploration of change or stability. J Rehabil Med 2015; 47(6): 531–537.

Chapter 7

The course of physical capacity in wheelchair users during training for the HandbikeBattle and at one-year follow-up

Kouwijzer I, Valent LJM, Post MWM, Wilders LM, Grootoonk A, HandbikeBattle Group, van der Woude LHV, de Groot S. The course of physical capacity in wheelchair users during training for the HandbikeBattle and at one-year follow-up. *Under review*.

Abstract

Objective: (1) to compare physical capacity at one year follow-up with physical capacity before and after the training period for the HandbikeBattle event; (2) to identify determinants of the course of physical capacity during follow-up.

Design: Prospective observational study. Former rehabilitation patients (N=33) with health conditions such as spinal cord injury or amputation were included. A handcycling / arm crank graded exercise test was performed before (January, T1) and after the training period (June, T2), and at one-year follow-up (June, T4). Outcomes: Peak power output (POpeak (W)) and peak oxygen uptake (VO_2peak (L/min)). Determinants: sex (M/F); age (years); classification; physical capacity, musculoskeletal pain, exercise stage of change, and exercise self-efficacy at T1; and HandbikeBattle participation at T4.

Results: multilevel regression analyses showed that POpeak and VO_2peak increased during the training period and did not significantly change during follow-up (T1: 112±37W, 1.70±0.48L/min; T2: 130±40W, 2.07±0.59L/min; T4: 126±42W, 2.00±0.57L/min). Participants who competed again in the HandbikeBattle showed slight improvement in physical capacity during follow-up, whereas participants who did not compete again showed a decrease.

Conclusion: Physical capacity showed an increase during the training period and remained stable after one-year follow-up. This study shows that training towards a goal is very important in exercise maintenance.

Key words: Cardiorespiratory Fitness, Longitudinal Studies, Rehabilitation, Exercise

Introduction

Wheelchair users in general have a low physical capacity compared to able-bodied individuals [1]. Apart from disability, this is due to the lower muscle mass in the upper body compared to legs, but also to a more sedentary / inactive lifestyle. In previous studies, improvements in physical capacity were associated with a lower risk for cardiovascular disease [2], a higher chance to return to work [3] and a higher life satisfaction [4]. Therefore, exercise interventions to increase upper-body physical capacity are important.

Several studies have shown positive effects of exercise on upper-body physical capacity in wheelchair users [5–8]. Exercise maintenance on the long term is, however, a challenge. In a previous follow-up study, which was undertaken three months after a controlled twice-weekly training study for nine months in individuals with spinal cord injury (SCI), exercise adherence dropped from 80.6% to 42.7% [6,9]. Possible explanations mentioned by the authors were: (1) the obligation that participants felt to come to the lab during the controlled lab-based study and the lack of this obligation during follow-up; (2) the presence of a goal, i.e., completing the nine-month study and the absence of a goal during follow-up; and (3) the degree of pain, which had an explained variance of 83% for exercise adherence during follow-up [9]. In a previous study on leisure time physical activity in individuals with SCI, it was shown that important factors for being stably active over time were: not having pressure ulcers, higher levels of exercise intentions, less severe SCI, age (being younger) and fewer years postinjury [10].

With respect to behavioral change and adopting or maintaining an active lifestyle, behavioral change models focus on several important constructs that are a prerequisite for engaging in exercise behavior. Examples of important constructs are the attitude towards exercise (exercise stage of change) and one’s confidence to regularly engage in physical activity and exercise (exercise self-efficacy) [11–15]. These constructs are thought to be both static and dynamic in nature and could, therefore, predict certain behavior, but could also be influenced and change over time.

Handcycling is a common mode of exercise for wheelchair users in the Netherlands. Today, handcycling is introduced already early in rehabilitation and is an easy mode to practice and cover larger distances at relatively high speeds. This can be explained by the higher efficiency and consequent higher power output (PO in W) in handcycling, while also accompanied by lower shoulder loads compared to handrim wheelchair propulsion [16,17]. Considering the beneficial effects of handcycling and the potential stimulating effect of training towards a goal, the HandbikeBattle was organized for the first time in 2013 [18]. In this Dutch annual event in the mountains of Austria, teams from twelve Dutch rehabilitation centers participate. Each team consists of former rehabilitation patients with a chronic disability such as a SCI, amputation or cerebral palsy. Prior to the event in June, participants train for a period of 4-5 months. At the start of the training period, most participants are relatively untrained handcyclists. Guidance during the training period is provided by therapists from the respective rehabilitation centers, but otherwise the training is self-organized and free-living for the full period, i.e., no specific training program is provided by the researchers. The aim of the training period and event is that participants learn to adopt an active lifestyle, experience positive effects in daily life, and continue to participate in sports on the long term. Previous studies have shown that training for the HandbikeBattle event results in improvement in physical capacity during the training period [4,5]. Long-term effects on physical capacity are, however, unknown. It is expected that participants who completed the HandbikeBattle are likely to maintain an active lifestyle because the training was not lab-based but self-organized in their own environment, they were physically active during the training period and possibly experienced positive effects of this lifestyle, and they have less barriers because they overcame certain barriers during the training period.

The purpose of the present study was, therefore, (1) to compare physical capacity one year after the HandbikeBattle event with physical capacity before and after the training period, (2) to identify determinants that influence the course of physical capacity during follow-up.

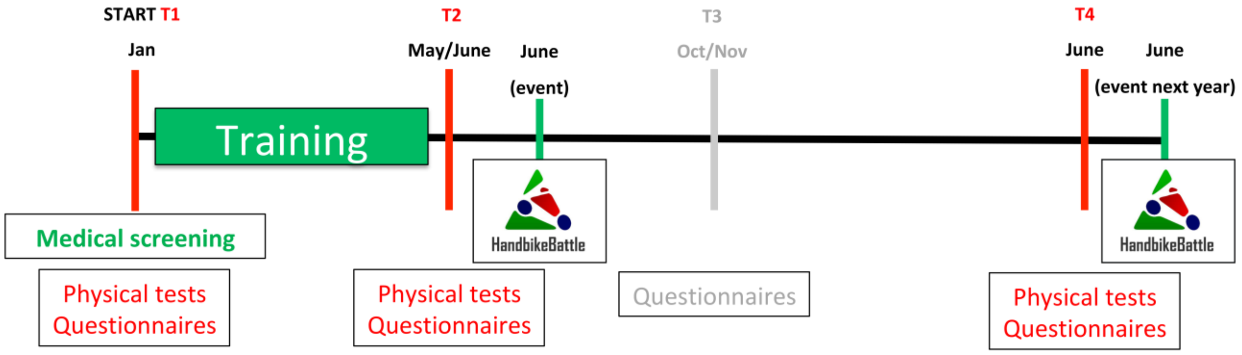
Methods

Participants

Inclusion criteria for the HandbikeBattle event were: being a former rehabilitation patient from one of the twelve participating rehabilitation centers; impairment of the lower extremities due to e.g., SCI, amputation, cerebral palsy or spina bifida; and commitment to participate in the HandbikeBattle event. Exclusion criterion: contra-indications to participate in the HandbikeBattle as diagnosed during the medical screening. In the present study, data were used from participants of the HandbikeBattle 2017 and 2018 cohorts (N=125). Four out of twelve rehabilitation centers were able (considering logistics, time constraints and financial situation) to conduct a follow-up graded exercise test (GXT) for the 2017 and 2018 cohorts one year after participation (in June 2018 and June 2019, respectively). As a result, 53 former HandbikeBattle participants were asked to perform a follow-up GXT one year after their participation in the HandbikeBattle event. All participants voluntarily signed an informed consent form. The study was approved by the Local Ethics Committee of the Center for Human Movement Sciences, University Medical Center Groningen, the Netherlands (ECB/2012_12.04_I_rev/Ml).

Procedure

The HandbikeBattle study has a prospective observational design. Measurements are performed at the start of the training period (January, T1); after the training period, prior to the event (June, T2); at follow-up, four months after the event (October/November, T3); and at follow-up, one year after the event (June, T4) (figure 1). At T1 a medical screening was performed by a rehabilitation physician or sports physician, which comprised a medical anamnesis, physical examination and a handcycling / arm crank GXT. At T2 and T4 the GXT was repeated with the same protocol and equipment. At all time points, participants were asked to fill out questionnaires about musculoskeletal pain, exercise stage of change and exercise self-efficacy.



HandbikeBattle **event**:	the mountain time trial in Austria
HandbikeBattle **project**:	medical screening, free-living training (with peers, guided by rehabilitation center), week in Austria among peers, participation in the HandbikeBattle event
HandbikeBattle **study**:	4 test occasions with physical tests and questionnaires. Aim: to study the effects of participating in the HandbikeBattle project for new participants

Figure 1. The design of the HandbikeBattle study. Time point T3 was not taken into account for the analyses in the present study.

Physical capacity

At T1, T2 and T4, physical capacity was measured during a synchronous incremental handcycling/arm crank GXT to volitional exhaustion. The GXTs were organized in and conducted by the staff of each of the participating rehabilitation centers. Dependent on the rehabilitation center, the GXTs were performed with the use of an arm ergometer (Lode Angio, Groningen, the Netherlands) or a recumbent sport handcycle attached to the Cyclus 2 ergometer (RBM elektronik-automation GmbH, Leipzig, Germany). Either a 1-min stepwise protocol or continuous ramp protocol was used and was individualized for each participant. The set-up and protocol choice were consistent within participants over time. PO (W) and oxygen uptake (VO_2 (L/min)) were measured during the test. Outcome parameters for physical capacity were POpeak and VO_2peak. For the 1-min stepwise protocol, POpeak was defined as the highest PO that was maintained for at least 30 s. For the ramp protocol, the highest PO achieved during the test was considered POpeak. Peak oxygen uptake (VO_2peak) was defined as the highest 30-s average for VO_2.

Determinants

Possible determinants that could explain differences among participants during follow-up were: sex (M/F), age (years), physical capacity at T1, handcycling classification, musculoskeletal pain at T1, exercise stage of change at T1, exercise self-efficacy at T1, and whether participants were going to participate again in the HandbikeBattle event at the time of their follow-up GXT (T4).

Handcycling classification was used as a proxy for severity of impairment and determined by an UCI certified Paracycling classifier, following the UCI Para-cycling Regulations. This resulted in five classes, ranging from H1 (most impaired) to H5 (least impaired) [19]. H1 and H2 handcyclists have limitations in arm-hand function, whereas H3, H4 and H5 handcyclists have intact arm-hand function and limitations in trunk and/or lower extremities only. For the analyses in the present study, participants were divided in two groups of equal size: (1) H1-H3 and (2) H4-H5.

Musculoskeletal pain comprised seven locations (hand/wrist (L/R), elbow (L/R), shoulder (L/R) and neck), with range 1=no pain, 6=very severe pain. Having moderate-severe pain was defined as ≥4 (moderate pain) at one or more locations. Two groups were created: (1) no-mild pain and (2) moderate-severe pain.

Exercise stage of change was measured with one question where participants had to select one of five statements reflecting their current exercise behavior. In these statements, the five stages of change were reflected: (1) precontemplation (no intention to become active), (2) contemplation (considering to become active), (3) preparation (irregularly active), (4) action (regularly active for less than six months) and (5) maintenance (regularly active for more than six months) [13]. For analyses, two groups were created: (1) 1 – 3 and (2) 4 - 5.

Exercise self-efficacy was measured with the Exercise Self-Efficacy Scale consisting of 10 items about self-confidence with respect to physical activity and exercise [20,21]. All items had a 4-point scale ranging from not at all true (1) to always true (4). A sum score of the 10 items was calculated ranging from 10 (lowest self-efficacy) to 40 (highest self-efficacy).

Statistical analyses

The analyses were performed using SPSS (IBM SPSS Statistics for Windows, Version 24.0. Armonk, NY: IBM Corp.) and MLwiN version 3.02 [22]. Descriptive statistics were calculated for outcome parameters and determinants. Outcome parameters were tested for normality with the Kolmogorov–Smirnov test with Lilliefors significance correction and the Shapiro–Wilk test, combined with z-scores for skewness and kurtosis. Individuals that performed the follow-up GXT (participants, N=33) were compared on baseline characteristics with individuals that did not perform the follow-up GXT (non-participants, N=20). Baseline characteristics were compared using independent-samples t-tests, Mann-Whitney U tests and chi-squared tests.

To study the longitudinal trajectory of physical capacity, three-level multilevel models were created with observations within participants as first level, participant as second level, and rehabilitation center as third level to make adjustments for the dependency of the observations within participants, and participants within centers [23]. Two models were created with either POpeak or VO_2peak as dependent variable. In each model, time (T1, T2, T4) was included as a categorical variable with two dummies and T2 as reference category.

To study determinants that influence the course of physical capacity during follow-up (T4), interaction terms with the time dummies were investigated in a series of separate models for each of the following determinants: sex (reference: male), age (years), physical capacity at T1, handcycling classification (reference: H1-H3), musculoskeletal pain at T1 (reference: no-mild pain), exercise stage of change at T1 (reference: 1-3), exercise self-efficacy at T1, and whether participants were going to participate again in the HandbikeBattle event at T4 (reference: no).

Results

Of the 53 participants who were asked to perform a follow-up GXT, 20 did not successfully perform the GXT, whereas 33 were successful. Reasons for not performing the GXT at T4 were: medical reasons (N=5; which were psychological problems (N=2), severe back pain, allergic reaction, and illness not specified), motivational problems (N=2), time constraints (N=1), family matters (N=1), loss of contact (N=4), unknown reasons (N=4), and one former participant passed away. Two more individuals were excluded as their follow-up GXT was performed with a different protocol than their previous GXTs. Hence, data from 33 individuals were used in the present study. There were no significant differences at baseline between participants and non-participants (table 1). Both outcome parameters were normally distributed. Exercise stage of change was considered not to be discriminative as 83% of participants scored in the highest category (table 1). Therefore, exercise stage of change was not taken into account as a determinant for change in physical capacity during follow-up. Of the 33 participants, eighteen participants competed again in the HandbikeBattle event at the time of T4, whereas 15 participants did not compete again.

Table 1. Characteristics and outcomes at T1 for participants and non-participants.

Characteristics	***N***	**Participants**		***N***	**Non-participants**	
Sex (male/female) (%)	33	22/11	(67/33)	20	16/4	(80/20)
Age (years) (SD)	33	40	(14)	20	41	(14)
Body mass (kg) (SD)	30	76	(22)	20	78	(22)
Impairment type	33			20		
Spinal cord injury (%)		17	(52)		12	(60)
Tetraplegia (%)		2	(6)		2	(10)
Paraplegia (%)		15	(46)		10	(50)
Amputation (%)		3	(9)		2	(10)
Cerebral palsy (%)		3	(9)		3	(15)
Stroke (%)		2	(6)		0	(0)
Multi trauma (%)		1	(3)		1	(5)
Spina bifida (%)		1	(3)		1	(5)
Other (%)		6	(18)		1	(5)
POpeak (W) (SD)	33	112	(37)	20	107	(41)
VO_2peak (L/min) (SD)	32	1.70	(0.48)	20	1.73	(0.56)
Handcycling classification (H1–H3/H4–H5) (%)	33	16/17	(48/52)	20	10/10	(50/50)
Musculoskeletal pain (no-mild/moderate-severe) (%)	26	15/11	(58/42)	17	9/8	(53/47)
Exercise stage of change (1-3/4-5) (%)	24	4/20	(17/83)	17	2/15	(12/88)
Exercise self-efficacy (SD)	24	35.8	(3.5)	17	35.1	(4.4)

Data represent *N* (%) or mean (SD). POpeak: peak power output; VO_2peak: peak oxygen uptake. Handcycling classification: two categories: (1) H1–H3 and (2) H4–H5. Musculoskeletal pain: two categories: (1) no-mild pain and (2) moderate-severe pain. Exercise stage of change: two categories: (1) 1-3 and (2) 4-5.

Longitudinal trajectory of physical capacity

Physical capacity over time is shown in table 2. At group level, POpeak and VO_2peak showed a significant increase between T1 (start training) and T2 (after training) and did not significantly change between T2 and T4 (one-year follow-up) (table 3). When the models were re-calculated with T1 as reference category, there was also a significant increase between T1 and T4 for both POpeak (beta 12.78, SE 2.99, $p<0.001$) and VO_2peak (beta 0.27, SE 0.06, $p<0.001$).

Table 2. Outcome parameters of participants at all time points.

	N	T1		*N*	T2		*N*	T4	
POpeak (W)	33	112	(37)	32	130	(40)	33	126	(42)
VO_2peak (L/min)	32	1.70	(0.48)	32	2.07	(0.59)	32	2.00	(0.57)

Data represent mean (SD). POpeak: peak power output; VO_2peak: peak oxygen uptake. T1 = start of the training period. T2 = after the training period, prior to the HandbikeBattle event. T4 = follow-up measurement, 1 year after the event.

Determinants of the course of physical capacity during follow-up

Sex, age, physical capacity at T1, handcycling classification, musculoskeletal pain at T1, and exercise self-efficacy at T1 showed no interaction effects with time (table 4). Participants who competed again in the HandbikeBattle event around T4 (N=18 competitors) showed a significantly different change in physical capacity between T2 and T4 than participants who did not compete again in the HandbikeBattle event (N=15 non-competitors) (figure 2). At T4, POpeak was 138 W for competitors versus 111 W for non-competitors, whereas VO_2peak was 2.18 L/min for competitors versus 1.80 L/min for non-competitors. Additional multilevel regression analyses for each subgroup showed that the increase in physical capacity between T2 and T4 for the competitors was not significant (POpeak: beta 4.39, SE 3.49, *p*=0.21; VO_2peak: beta 0.05, SE 0.07, *p*=0.50). However, the decrease in physical capacity between T2 and T4 for the non-competitors was significant (POpeak: beta -10.87, SE 4.20, *p*=0.01; VO_2peak: beta -0.17, SE 0.07, *p*=0.03). When the models for the non-competitors were re-calculated with T1 as reference category, there was no significant difference between T1 and T4 for POpeak (beta 4.33, SE 4.20, *p*=0.30). VO_2peak was, however, still significantly higher at T4 compared with T1 (beta 0.19, SE 0.08, *p*=0.01).

Discussion

Physical capacity showed a significant increase during the training period and at group level this remained stable at one-year follow-up. More detailed analyses showed that participants who competed again in the HandbikeBattle showed a slight (non-significant) improvement in physical capacity during follow-up, whereas participants who did not compete again in the HandbikeBattle showed a significant decrease.

Table 3. Longitudinal trajectory of physical capacity.

		Constant (reference: T2)	Δ T2 – T1		Δ T2 – T4	
	N	Regression coefficient (SE)	Regression coefficient (SE)	*p*-value	Regression coefficient (SE)	*p*-value
POpeak (W)	33	128.37 (6.91)	-15.21 (3.02)	<0.001	-2.43 (2.91)	0.40
VO$_2$peak (L/min)	32	2.05 (0.10)	-0.32 (0.06)	<0.001	-0.05 (0.06)	0.34

POpeak: peak power output; VO$_2$peak: peak oxygen uptake. T1 = start of the training period. T2 = after the training period, prior to the HandbikeBattle event. T4 = follow-up measurement, 1 year after the event. Δ T2 – T1 = a negative regression coefficient represents an improvement of the dependent variable over time. Δ T2 – T4 = a negative regression coefficient represents a deterioration of the dependent variable over time. SE = standard error.

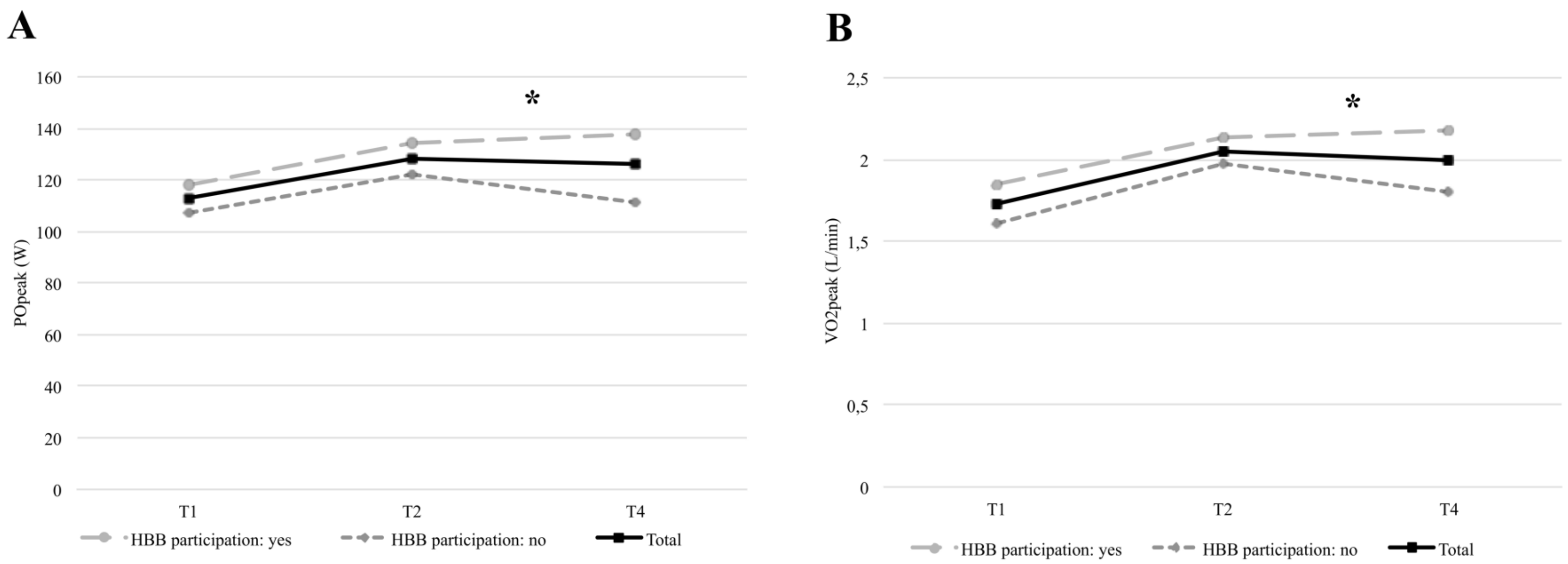


Figure 2. Multilevel regression analyses: longitudinal trajectory of physical capacity with interaction effects of HandbikeBattle (HBB) participation at the time of follow-up (T4). T1 = start of the training period. T2 = after the training period, prior to the HandbikeBattle event. T4 = follow-up measurement, 1 year after the event. A. Regression analysis for POpeak (W). B. Regression analysis for VO$_2$peak (L/min). * Significant difference in course of physical capacity with $p < 0.05$, between HBB participation yes vs. no.

Table 4. Longitudinal trajectory of physical capacity with interaction effects.

	Constant (reference:T2)	Δ T2 – T1	Δ T2 – T4	Determinant	(Δ T2 – T1) x determinant	(Δ T2 – T4) x determinant
POpeak (W)						
Sex	141.59 (7.64)	-17.93 (3.60)*	-4.00 (3.48)	-40.16 (13.30)*	8.62 (6.39)	5.20 (6.17)
Age	102.85 (21.49)	0.69 (9.86)	10.03 (9.13)	0.63 (0.51)	-0.38 (0.23)	-0.30 (0.21)
POpeak at T1	14.34 (8.51)	-14.34 (10.53)	1.49 (10.53)	1.01 (0.07)	-0.01 (0.09)	-0.04 (0.09)
Handcycling classification	113.08 (8.99)	-15.19 (4.20)*	-7.71 (4.10)	29.80 (12.49)*	-0.31 (5.85)	10.12 (5.65)
Musculoskeletal pain	139.80 (10.31)	-19.34 (3.65)*	-5.53 (3.56)	-8.98 (15.85)	6.40 (5.78)	1.90 (5.47)
Exercise self-efficacy	-16.00 (81.98)	-42.05 (31.68)	-11.54 (29.82)	4.28 (2.28)	0.72 (0.88)	0.19 (0.83)
HandbikeBattle participation at T4	121.85 (9.94)	-15.18 (4.16)*	-10.85 (4.06)*	12.15 (13.43)	-0.40 (5.62)	15.24 (5.43)*
VO$_2$peak (L/min)						
Sex	2.25 (0.11)	-0.30 (0.07)*	-0.08 (0.07)	-0.56 (0.19)*	-0.05 (0.12)	0.09 (0.12)
Age	1.83 (0.33)	-0.32 (0.19)	0.15 (0.19)	0.005 (0.008)	-0.000 (0.004)	-0.005 (0.004)
VO$_2$peak at T1	0.41 (0.17)	-0.41 (0.21)*	0.17 (0.21)	0.94 (0.10)*	0.06 (0.12)	-0.12 (0.12)
Handcycling classification	1.81 (0.13)	-0.23 (0.08)*	-0.05 (0.08)	0.47 (0.18)*	-0.18 (0.11)	-0.01 (0.11)
Musculoskeletal pain	2.10 (0.16)	-0.23 (0.08)*	-0.08 (0.08)	0.03 (0.24)	-0.18 (0.13)	0.01 (0.12)
Exercise self-efficacy	0.32 (1.27)	-0.68 (0.72)	-0.28 (0.68)	0.05 (0.04)	0.01 (0.02)	0.01 (0.02)
HandbikeBattle participation at T4	1.97 (0.14)	-0.36 (0.08)*	-0.17 (0.08)*	0.16 (0.19)	0.08 (0.11)	0.22 (0.11)*

Data represent multilevel regression model coefficients (SE). For both outcome parameters (POpeak and VO$_2$peak), seven separate models were created (one model for each determinant). Each model consisted of the time dummies, one determinant, and the interaction effect between time and determinant. POpeak: peak power output; VO$_2$peak: peak oxygen uptake. Sex: M/F, reference: male. Handcycling classification: two categories: (1) H1–H3 and (2) H4–H5, reference: H1-H3. Musculoskeletal pain: two categories: (1) no-mild pain and (2) moderate-severe pain, reference: no-mild pain. T1 = start of the training period. T2 = after the training period, prior to the HandbikeBattle event. T4 = follow-up measurement, 1 year after the event. Δ T2 – T1 = a negative regression coefficient represents an improvement of the dependent variable over time. Δ T2 – T4 = a negative regression coefficient represents a deterioration of the dependent variable over time. HandbikeBattle participation = whether participants were going to participate again in the HandbikeBattle event at the time of their follow-up GXT (0 = no, 1 = yes, reference: no). * Significance with $p < 0.05$.

Physical capacity of the participants at the start (T1: POpeak 112 ± 37 W; VO_2peak 1.70 ± 0.48 L/min) was slightly lower than in previous studies in the HandbikeBattle population (POpeak 119 – 126 W; VO_2peak 1.91 – 2.01 L/min) [4,5,24]. The increase in physical capacity (POpeak 16%; VO_2peak 22%) during the training period (T1-T2) is comparable with other HandbikeBattle studies and other intervention studies for wheelchair users with a SCI [5,25].

Long-term follow-up studies on physical capacity or physical activity among wheelchair users are scarce, which is unfortunate as long-term follow-up data are essential to gain knowledge on effects of exercise and training as well as on determinants of maintenance and relapse in physical activity behavior. In the present study, physical capacity remained stable after one-year follow-up for the total group. The only determinant that was associated with the course of physical capacity during follow-up, was participating in the HandbikeBattle event again at the time of follow-up. From these results it is suggested that having a goal to train for appears to be very important in exercise maintenance, which is in line with hypotheses in previous research [9,26,27]. The follow-up question would then be why certain participants choose to pursue this goal again, whereas others do not. Having a high physical capacity at the start, and therefore possibly having a more active lifestyle in general, was not associated with the course of physical capacity during follow-up. Again, it was also noted that this was not an extremely fit subgroup of the HandbikeBattle population. In addition, the change in physical capacity during the training period (T1-T2) did not have an interaction effect with participation in the HandbikeBattle during follow-up (table 4). This indicates that participants who showed the highest gains in physical capacity during the training period are not necessarily the participants competing again in the event next year.

Sex, age, handcycling classification, musculoskeletal pain, exercise stage of change and exercise self-efficacy were not associated with the course of physical capacity during follow-up. The mean age in the present study was 40 years with range 13 – 59 years, therefore all participants were in the age-category of potentially participating in school or work. The fact that participants with retirement age were not represented, could be an explanation for the finding that age was not associated with the course of physical activity. Compared with a previous study in individuals with SCI that concluded severity of the injury to be associated with leisure time physical activity, the participants in the present study were less severely injured [10]. In the present study only 9% of participants were classified as H1/H2 (comparable with tetraplegia), whereas in Sweet et al. 53% had a tetraplegia [10]. It is uncertain why musculoskeletal pain was not associated with long-term physical capacity. A possible explanation is that as a result of exercise, pain is fluctuating (decreasing) over time [6]. Therefore it could be that musculoskeletal pain at baseline is not a predictor of long-term exercise maintenance, but that longitudinal changes in pain are associated with changes in physical capacity over time. Another explanation is that individuals who have severe (exercise-limiting) pain are not participating in (training for) the HandbikeBattle and therefore the HandbikeBattle participants are a selection with relatively low pain scores.

Participants scored high on exercise stage of change. Eighty-three percent considered themselves as being regularly physically active at the start of the training period. Being regularly active was defined as performing activities like exercise and sports, but also cleaning and household activities for at least 30 minutes a day for at least five days a week. It could be that the participants were not necessarily involved in sports at the start of the study but were active in their household and daily commute to, for example, work or the supermarket.

Exercise self-efficacy was not associated with the course of physical capacity during follow-up. Participants had a mean score of 35.8 ± 3.5, which is fairly high but slightly lower than previous research in a Dutch population with sub-acute SCI in the Act-Active study (N=37, median: 37.0, IQR: 34.0-39.0) [28], and higher than another large study in a Dutch (inactive) population with long-standing SCI (ALLRISC, N=268, mean ± SD: 31.4 ± 7.8) [29]. In the last study, multivariate regression models showed a significant association between exercise self-efficacy and physical activity but with an explained variance of only 2% [29]. In a home-based exercise intervention study in individuals with SCI, exercise self-efficacy was not associated with physical activity, but a change in exercise self-efficacy was associated with a change in VO_2peak/kg over time [30].

Limitations

Due to missing data over time and a relatively small sample size, it was not possible to study the dynamic longitudinal character of exercise self-efficacy, exercise stage of change and musculoskeletal pain, and their associations with physical capacity over time. In the present study self-efficacy, exercise stage of change and musculoskeletal pain at baseline were not predictive of long-term physical capacity, but it would be interesting to investigate the course of these determinants over time and their association with long-term exercise maintenance.

Implications and future studies

Long-term follow-up studies on exercise maintenance in wheelchair users are scarce. The present study shows that physical capacity increases during the training period, and that this increase in physical capacity remains stable at one-year follow-up. The only determinant that was associated with the course of physical capacity during follow-up was whether participants were going to compete again in the event at the time of follow-up. These results showed that having a goal to train for is a very important determinant for exercise maintenance. The follow-up question would then be why certain participants choose to pursue this goal again, whereas others do not. In addition, goal setting in general is an important factor to focus on as pursuing other (even more challenging) goals could be equally or even more effective. Moreover, other (mediating) factors apart from the goal itself could be the competitive element or the social aspect of training with peers. Future studies should focus on which motivational factors and other determinants play a role in maintaining physical capacity on the long term in wheelchair users.

Conclusion

Physical capacity showed an increase during the training period and remained stable at one-year follow-up. Participants who competed again in the HandbikeBattle showed a slight (non-significant) improvement in physical capacity during follow-up, whereas participants who did not compete again in the HandbikeBattle showed a significant decrease. These results show that training towards a goal is very important in long-term exercise maintenance.

Author Disclosures

The authors report no conflict of interest.

Acknowledgements

HandbikeBattle group name: Paul Grandjean Perrenod Comtesse, Adelante Zorggroep, Hoensbroek, the Netherlands. Eric Helmantel, University Medical Center Groningen, Center for Rehabilitation Beatrixoord, Groningen, the Netherlands. Mark van de Mijll Dekker, Heliomare Rehabilitation Center, Wijk aan Zee, the Netherlands. Maremka Zwinkels, Rehabilitation Center De Hoogstraat, Utrecht, the Netherlands. Misha Metsaars, Libra Rehabilitation and Audiology, Eindhoven, the Netherlands. Ellen Moons-Langeweg, Sint Maartenskliniek, Nijmegen, the Netherlands. Linda van Vliet, Amsterdam Rehabilitation Research Center | Reade, Amsterdam, the Netherlands. Wilbert Snoek, Rehabilitation center Revant, Breda, the Netherlands. Karin Postma, Rijndam Rehabilitation Center, Rotterdam, the Netherlands. Bram van Gemeren, Roessingh Rehabilitation Center, Enschede, the Netherlands. Selma Overbeek, Rehabilitation center Tolbrug, Den Bosch, the Netherlands. Alinda Gjaltema, Vogellanden, Zwolle, the Netherlands.

This study was funded by HandicapNL, Stichting Mitialto, Stichting Beatrixoord Noord-Nederland, University Medical Center Groningen, Heliomare Rehabilitation Center and Stichting Handbike Events.

References

1. Van den Berg-Emons RJ, Bussmann JBJ, Haisma JA, et al. A prospective study on physical activity levels after spinal cord injury during inpatient rehabilitation and the year after discharge. *Arch Phys Med Rehabil*. 2008;89(11):2094-2101.

2. Nightingale TE, Walhin J, Thompson D, Bilzon JLJ. Impact of exercise on cardiometabolic component risk in spinal cord-injured humans. *Med Sci Sport Exerc*. 2017;47(12):2469-2477.

3. Van Velzen JM, de Groot S, Post MWM, Slootman JR, van Bennekom CAM, van der Woude LHV. Return to work after spinal cord injury. *Am J Phys Med Rehabil*. 2009;88(1):47-56.

4. Kouwijzer I, de Groot S, van Leeuwen CMC, et al. Changes in quality of life during training for the HandbikeBattle and associations with cardiorespiratory fitness. *Arch Phys Med Rehabil*. 2020;101(6):1017-1024.

5. Hoekstra SP, Valent LJM, Gobets D, van der Woude LHV, de Groot S. Effects of four-month handbike training under free-living conditions on physical fitness and health in wheelchair users. *Disabil Rehabil*. 2017;39(16):1581-1588.

6. Hicks AL, Martin KA, Ditor DS, et al. Long-term exercise training in persons with spinal cord injury: effects on strength, arm ergometry performance and psychological well-being. *Spinal Cord*. 2003;41(1):34-43.

7. Jacobs PL, Nash MS. Exercise recommendations for individuals with spinal cord injury. *Sport Med*. 2004;34(11):727-751.

8. Hicks AL, Martin Ginis KA, Pelletier CA, Ditor DS, Foulon B, Wolfe DL. The effects of exercise training on physical capacity, strength, body composition and functional performance among adults with spinal cord injury: a systematic review. *Spinal Cord*. 2011;49(11):1103-1127.

9. Ditor DS, Latimer AE, Martin Ginis KA, Arbour KP, McCartney N, Hicks AL. Maintenance of exercise participation in individuals with spinal cord injury: effects on quality of life, stress and pain. *Spinal Cord*. 2003;41(8):446-450.

10. Sweet SN, Martin Ginis KA, Latimer-Cheung AE. Examining physical activity trajectories for people with spinal cord injury. *Health Psychol*. 2012;31(6):728-732.

11. Bandura A. Self-efficacy: toward a unifying theory of behavioral change. *Psychol Rev*. 1977;84(2):191-215.

12. Prochaska JO, DiClemente CC. Stages and processes of self-change of smoking: toward an integrative model of change. *J Consult Clin Psychol*. 1983;51(3):390-395.

13. Kosma M, Ellis R, Cardinal BJ, Bauer JJ, McCubbin JA. The mediating role of intention and stages of change in physical activity among adults with physical disabilities: an integrative framework. *J Sport Exerc Psychol*. 2007;29(1):21-38.

14. Marcus BH, Simkin LR. The transtheoretical model: applications to exercise behavior. *Med Sci Sport Exerc*. 1994;26(11):1400-1404.

15. Ajzen I. The theory of planned behavior. *Organ Behav Hum Decis Process*. 1991;50(1):179-211.

16. Dallmeijer AJ, Zentgraaff IDB, Zijp NI, van der Woude LHV. Submaximal physical strain and peak performance in handcycling versus handrim wheelchair propulsion. *Spinal Cord*. 2004;42(2):91-98.

17. Arnet U, van Drongelen S, Scheel-Sailer A, van der Woude LHV, Veeger DHEJ. Shoulder load during synchronous handcycling and handrim wheelchair propulsion in persons with paraplegia. *J Rehabil Med*. 2012;44(3):222-228.

18. De Groot S, Postma K, van Vliet L, Timmermans R, Valent LJM. Mountain time trial in handcycling: exercise intensity and predictors of race time in people with spinal cord injury. *Spinal Cord*. 2014;52(6):455-461.

19. Union Cycliste Internationale Cycling Regulations, part 16 Para-cycling. Available at: https://www.uci.org/docs/default-source/rules-and-regulations/16-par-20200211-e.pdf. Accessed November 25, 2019.

20. Kroll T, Kehn M, Ho P, Groah S. The SCI exercise self-efficacy scale (ESES): development and psychometric properties. *Int J Behav Nutr Phys Act*. 2007;4(34):2-7.

21. Nooijen CFJ, Post MWM, Spijkerman DCM, Bergen MP, Stam HJ, van den Berg-Emons RJG. Exercise self-efficacy in persons with spinal cord injury: psychometric properties of the Dutch translation of the exercise self-efficacy scale. *J Rehabil Med*. 2013;45(4):347-350.

22. Charlton C, Rasbash J, Brown W, Healy M, Cameron B. MLwiN Version 3.02. Centre for Multilevel Modelling, University of Bristol, 2020.

23. Twisk JWR. Applied Longitudinal Data Analysis for Epidemiology. A Practical Guide. 4th ed. Cambridge (UK): Cambridge University Press; 2003.

24. Kouwijzer I, Valent LJM, Osterthun R, van der Woude LHV, de Groot S, HandbikeBattle Group. Peak power output in handcycling of individuals with a chronic spinal cord injury: predictive modeling, validation and reference values. *Disabil Rehabil*. 2020;42(3):400-409.

25. Valent LJM, Dallmeijer AJ, Houdijk H, Talsma E, van der Woude LHV. The effects of upper body exercise on the physical capacity of people with a spinal cord injury: a systematic review. *Clin Rehabil*. 2007;21(4):315-330.

26. Ajzen I, Kruglanski AW. Reasoned action in the service of goal pursuit. *Psychol Rev*. 2019;126(5):774-786.

27. Jaarsma EA, Dijkstra PU, Geertzen JHB, Dekker R. Barriers to and facilitators of sports participation for people with physical disabilities: A systematic review. *Scand J Med Sci Sport*. 2014;24(6):871-881.

28. Nooijen CFJ, Post MWM, Spooren AL, et al. Exercise self-efficacy and the relation with physical behavior and physical capacity in wheelchair-dependent persons with subacute spinal cord injury. *J Neuroeng Rehabil*. 2015;12:103:1-8.

29. Kooijmans H, Post MWM, Motazedi E, et al. Exercise self-efficacy is weakly related to engagement in physical activity in persons with long-standing spinal cord injury. *Disabil Rehabil*. 2019;Epub ahead of print:1-7.

30. Nightingale TE, Rouse PC, Walhin J-P, Thompson D, Bilzon JL. Home-based exercise enhances health-related quality of life in persons with spinal cord injury: a randomized controlled trial. *Arch Phys Med Rehabil*. 2018;99(10):1998-2006.

Chapter 8

General discussion

Aims and outline of this thesis

The main aims of the studies in this thesis were to investigate the effects of participation in the HandbikeBattle project and event on physical capacity and quality of life, and to answer questions from clinical rehabilitation practice regarding physical capacity testing and handcycle training. Taken together, this resulted in three themes with the following specific aims:

Physical capacity testing

1. To develop and validate predictive models for peak power output (POpeak (W and W/kg)) in a synchronous handcycling graded exercise test (GXT) for individuals with spinal cord injury (SCI) (**chapter 2**).
2. To define reference values for absolute and relative POpeak and peak oxygen uptake (VO_2peak) in handcycling based on lesion level and sex (**chapter 2**).
3. To examine whether it is possible to detect both ventilatory thresholds (VTs) in recreationally active individuals with tetraplegia or paraplegia (**chapter 3**).
4. To examine the interrater and intrarater reliability of VT determination (**chapter 3**).
5. To examine the effects of stage duration with a ramp protocol, 1-min stepwise protocol, and 3-min stepwise protocol on PO, VO_2, and heart rate (HR) at both peak level and at VT1 and VT2 during synchronous arm crank ergometry (**chapter 4**).

Handcycle training

6. To analyze training characteristics of the HandbikeBattle participants (**chapter 5**).
7. To examine the associations between training load and the change in physical capacity (**chapter 5**).

Effects of participation

8. To examine changes in life satisfaction and mental health during five months of training prior to the HandbikeBattle and at four months of follow-up (**chapter 6**).
9. To examine the associations among changes in handcycling physical capacity and changes in life satisfaction and mental health during the training period (**chapter 6**).
10. To compare physical capacity one year after the HandbikeBattle event with physical capacity before and after the training period (**chapter 7**).
11. To identify determinants that influence the course of physical capacity during follow-up (**chapter 7**).

Summary and interpretation of main findings

Physical capacity testing

In **chapter 2** the objective was to give practitioners guidance in selecting the correct individualized GXT protocol for handcycling and give insight in what is considered a “normal” handcycle-specific physical capacity in this diverse population. However, the developed predictive models had a relatively low explained variance (R^2), with a R^2 of 42% for the best model. Validation of the models showed a fair to good relative agreement but a low absolute agreement between the predicted and measured POpeak. Therefore, the models should be used with caution and only in addition to expert opinion of the practitioner when there is indecisiveness in what protocol to choose. Moreover, it should be noted that the models are only applicable to individuals with SCI, that the models are not based on individuals with a very low or very high physical capacity (i.e., elite athletes), and that women and individuals with a tetraplegia were underrepresented in the study.

The follow-up question would be why such a large proportion of unexplained variance remained. This could partly be due to the variation in measurement set-up and protocols among the eleven rehabilitation centers. Standardization of test setting and protocol to pursue homogeneity might further improve the predictive models. Second, the measurement of certain included determinants could be improved. For example, with the available 104 participants in the model development group, it was only possible to define lesion level based on two categories (>Th6 and ≤Th6), whereas including three or four categories in the model could have been more discriminative. In addition, an objective measure of daily physical activity or exercise might explain a larger part of the variance than self-reported handcycle training hours alone.

Other factors that might improve the model include isometric arm strength and anaerobic PO. Isometric arm strength and anaerobic POmean during a 30-second sprint were found to have an explained variance of 66% and 81% in persons with SCI conducting peak handrim wheelchair ergometry [1]. In a follow-up study by De Groot et al. a predictive model for POpeak was constructed in asynchronous arm ergometry (N=93) with the following determinants: anaerobic POmean, age, sex, body mass index (BMI), injury level and time since injury (TSI), which led to a higher explained variance (R^2 = 76%) than in **chapter 2** [2]. Validation of the model showed a relative agreement (ICC) of 0.89 [2]. It should be noted that although inclusion of the anaerobic POmean resulted in improved predictive models and a Wingate test is a relatively low time burden, it involves extra testing. Ideally, readily available or very easy to test determinants are used. In clinical practice, i.e., in most rehabilitation centers in the Netherlands, practitioners choose the protocol of the individual participant out of a set of three or four existing protocols. For example, ramp protocols with a slope of 1W/12s (50W after ten minutes), 1W/6s (100W after ten minutes) or 1W/4s (150W after ten minutes). It is, therefore, not necessary to predict the exact POpeak. The developed predictive models of **chapter 2** are less precise than was aimed for, but they could support in the decision of which of these protocols to choose. For clinical practice, it would especially be helpful when the prediction is based on a practical set of determinants. For example, researchers from the University of Miami found evidence that the ability to perform a floor to chair transfer would likely attain a POpeak of at least 0.8W/kg body weight [3,4]. A next step could be to further investigate these kind of associations between functional ability tests and physical capacity to eventually provide more personalized protocols and outcomes.

In **chapter 3** the feasibility of determination of VTs and their interrater and intrarater reliability were studied for a peak arm crank test in recreationally active individuals with SCI. The results showed that 90% of VTs could be determined, which is comparable to what is described for able-bodied individuals and athletes with SCI [5–8]. Although these results seem promising, we have to realize that compared with able-bodied leg exercise, relatively untrained individuals with an SCI have unique characteristics in terms of upper-body training and testing. This is for example reflected in the small range between resting values and peak values in individuals with SCI, and those with tetraplegia in particular. This small range could be a complicating factor in the determination of VTs [7,9]. For example, the mean VO_2peak in **chapter 3** was 1.50 ± 0.64 L/min for individuals with a paraplegia, and 0.76 ± 0.32 L/min for individuals with a tetraplegia. Especially the latter leaves little variation if one assumes a resting metabolic rate of 0.25-0.29 L/min [9]. In able-bodied men with the same age, VO_2peak would be 2.69 ± 0.36 L/min [10]. VO_2 at VT1 for individuals with paraplegia was 0.74 – 0.76 L/min, and for individuals with a tetraplegia 0.47 - 0.53 L/min. In able-bodied men with the same age, this would be 1.08 – 1.61 L/min [10,11]. Setting the resistance of the first step too high in the GXT might complicate determination of VT1, whereas peripheral fatigue may result in premature termination of the test and complicate determination of VT2. In a previous arm crank ergometry study with untrained individuals with SCI, VT1 could be detected in all individuals with a paraplegia, but only in 68% of individuals with tetraplegia [9]. Individuals without observable VT1 completed the GXT with lower POpeak and VO_2peak than individuals with an observable VT1 [9]. In **chapter 3**, the same patterns were observed: most of the VTs that could not be determined were VT2s and related to tests in individuals with tetraplegia. Although the overall success rate in **chapter 3** was still 90%, for 7 out of 11 tests of individuals with tetraplegia, one or both raters could not determine one or both VTs.

Multiple follow-up questions come to mind and are, so far, unanswered. It is, for example, unknown what the day-to-day variability of VO_2, HR and PO at VTs is in this population. It is necessary to gain knowledge about this aspect before change in VO_2, HR or PO at VTs can be used as outcome parameter for the dose-response relationship with training [9,12]. This could be done with, for example, repeated graded exercise testing with several days in between. Moreover, we now make the assumption that training intensity prescription based on VTs is the best method in this population of wheelchair users with an SCI. However, other options to set training intensity should be further considered in relation to VTs in more systematic research, and other options should be available if one or both VTs cannot be determined. One of these options is training based on %POpeak. This is feasible in handcycling as power output can be measured with a device in the cranks or hub. Next to %POpeak, and %Heart Rate Reserve (%HRR) in those with lower lesions (<Th1), setting intensity based on subjective measures such as RPE or the talk test are possibilities for this population and should be further investigated [13,14]. Also, the sensitivity to training load in this population is important to study. Since handcycling has a distinctly higher efficiency and POpeak compared to handrim wheelchair propulsion[15], the role of exercise mode on VTs is important to understand as well. A similar question holds for the choice of GXT protocol, as was studied in **chapter 4**.

The set-up of **chapter 4** closely resembles **chapter 3,** but the GXTs were conducted in healthy able-bodied individuals and the focus was on determining VTs while employing three different GXT protocols during arm cranking. In **chapter 4** PO was different at peak level and at VTs among the three protocols. No significant differences were found in HR and VO_2 at peak level and at VTs but absolute agreement in HR and VO_2 was low among protocols, so protocols should never be used interchangeably when evaluating change in physical capacity longitudinally within a participant. PO at peak level and VTs was higher for the short stage protocols than the 3-min stepwise protocol. Consequently, training prescription based on PO at VTs assessed in short stage protocols might give an overestimation with training zones that could potentially result in overreaching. In the future this study should be repeated with a patient population that is dependent upon arm exercise. In addition, it should address these differences in PO among protocols and the clinical implications of these differences for e.g. individuals with SCI. For example, a training study could be conducted where two matched groups of persons with paraplegia receive the exact same training regime except in one group intensity is set based on a 1-min stepwise protocol, whereas in the other group intensity is set based on a 3-min stepwise protocol for the GXT. Primary outcome parameters would be improvement in physical capacity, the occurrence of pain and overuse injuries, and indicators for overreaching. It is especially important for individuals with tetraplegia to investigate the clinical implications of the overestimation of PO at the VTs in short stage protocols for two reasons: 1) it is likely that the GXT will be conducted with a short stage protocol as a 3-min stepwise protocol with relatively long test duration is suggested to be less feasible for individuals with tetraplegia, because peripheral fatigue will limit them, 2) individuals with tetraplegia are more likely to train based on PO at the VTs as training based on HR at the VTs is often not applicable due to the altered sympathetic response to exercise [16].

Handcycle training

Participants of the HandbikeBattle trained free living. Therefore, monitoring training effort is crucial to understand the dose-response associations of handcycle training. In **chapter 5** it is shown that participants of the HandbikeBattle train on average 21 ± 6 weeks, with 3.6 ± 1.4 training sessions a week. The average duration of a training session was 86 ± 20 minutes. Training load based on session Rating of Perceived Exertion ($TRIMP_{sRPE}$) [17] was not significantly associated with change in physical capacity. In addition, individual characteristics of training dose, i.e., frequency, duration and intensity also showed no significant association. This is unfortunate, as $TRIMP_{sRPE}$ would be a practical tool to use in rehabilitation practice as it is easy to use, cheap, applicable to all types of training and gives an overall measure of the perceived effort by the participant. The results of **chapter 5** are, however, in line with previous research in able-bodied athletes that neither showed unequivocal associations between $TRIMP_{sRPE}$ and change in physical capacity [18–21]. In addition, other training load parameters such as different HR-based TRIMP methods and the training stress score (TSS) based on PO, showed no significant associations with change in POmax in able-bodied cyclists [22,23]. Compared with previous able-bodied literature, the participant group in **chapter 5** was heterogeneous with, among others, different disability types, the training period was relatively long and varied in length among participants, and there was a variety of sports activities according to their diaries (e.g., handcycling, wheelchair basketball etc.). In addition, several participants had to stop training for several days or weeks due to urinary tract surgery, spasticity treatment, infections and pressure ulcers. We should be aware that untrained wheelchair users are a more vulnerable population than elite able-bodied athletes and that this is an extra complicating factor in any study on exercise and training effects, indeed complicating the presence of uniform associations between training load and change in physical capacity. In general, it is unsatisfactory that dose-response relationships are not clear as athletes, coaches, researchers and also rehabilitation professionals want to know: at what intensity, frequency and duration should an athlete/patient train to achieve the best response? What is the optimal training regime? We should, however, be aware that training adaptation comprises of a very complex interplay among numerous (temporal) factors. Examples of factors are: fatigue, sleep, nutrition, mental state, stress (work-life balance), genetics and motivation. Additional factors that might be applicable to wheelchair users and potentially affecting training availability over time are, for example, recurrent infections, fluctuating pain (related or not related to training), bowel / bladder problems and pressure ulcers. In addition, numerous choices can be made in terms of training characteristics, for example, the number of training sessions in a week, how to work towards a goal and build training load during a certain period (i.e. periodization), the type of training (e.g. strength training, low-intensity cardio training, high intensity interval training (HIIT)), the quality of the training session (e.g., the weight and number of repetitions during strength training, the intensity and duration of intervals during HIIT).

In this respect, the $TRIMP_{sRPE}$ has the advantage compared with other training load parameters that it gives an overall representation of the individual's perception of training, potentially taking into account physical, psychological and environmental factors [24]. Disadvantages are that the $TRIMP_{sRPE}$ does not take into account the variability of intensity within a training session and the specific content of the training session. For example, a training session of one hour with a stable intensity might have a RPE 6, whereas HIIT training with alternated very high and very low intensity might also score an average RPE 6. In addition, a training session of 60 minutes with RPE 5 will give a $TRIMP_{sRPE}$ of 300 AU, whereas a training session of 100 minutes with RPE 3 will also give a $TRIMP_{sRPE}$ of 300 AU. In a recent editorial in British Journal of Sports Medicine, the authors advocated a continuous and prospective monitoring approach which consists of a combination of objective physiological measures (heart rate, power output), subjective measures (RPE), psychological measures (stress, coping) and lifestyle-related factors (sleep, nutrition) [25]. As a next step, an individualized approach based on these factors would be ideal for training prescription. However, it should be really easy to understand and fill out for participants or patients, otherwise it will be too time consuming and not feasible during or after rehabilitation and/or training practice. In short, for future HandbikeBattle studies it is recommended to use (as a minimum) a combination of subjective (RPE) and objective (heart rate, power output) measures and monitor pain after each training.

Effects of participation

In **chapter 6** the objective was to describe changes in quality of life (life satisfaction and mental health) over time and to study longitudinal associations between quality of life and physical capacity. Life satisfaction increased and this increase was associated with an increase in physical capacity, whereas mental health only increased in a subgroup with mental health problems (i.e., score ≤ 72) at baseline. The question remains which underlying mechanisms are responsible for this association between physical capacity and life satisfaction. A direct association might be possible, i.e., the feeling of increased capacity, strength etc. might give an increased life satisfaction. However, other mediating factors might also play a role. One hypothesis is that an increase in physical capacity has a positive effect on functional independence and activities of daily living, which might have a positive effect on life satisfaction [26–29]. A second hypothesis is that an increase in physical capacity has a positive effect on body image (i.e., satisfaction with physical functioning and appearance), which might have a positive effect on life satisfaction [30–32]. It should further be noted that an increase of physical capacity is probably not the only determinant for the increase in life satisfaction. It is hypothesized that, for example, peer support and increase in social network during the training period might be important determinants for life satisfaction as well [33]. In a next step, it would be interesting to look at effects of participation from a different perspective. Previous HandbikeBattle studies have shown an increase in physical capacity on a group level. It is, however, unknown what is considered a meaningful effect on an individual level. In other words, what is the smallest change in physical capacity that an individual defines as important (minimal clinical important difference (MCID) [34]) and how does that affect daily life? For example, if an individual always needs help with a certain transfer, and following a subsequent training period, physical capacity increased with 15W which allows him/her to perform the transfer independently, this change in physical capacity may indeed be meaningful.

On another note, it is mostly unknown why certain participants remain active in handcycling (or other sports) after the HandbikeBattle, whereas others quit sports or dramatically decrease their time spent on sport activity. One hypothesis on why certain participants show a decrease in sport activity and/or a physically active lifestyle after the event is that they are fully (perhaps too much) engaged in the project, and set everything aside to train for the event during the training period. This is feasible for a limited period of time, but after the event they return to their normal life as it was before the project. It is a future challenge, not only for the HandbikeBattle, but for all exercise interventions, to engage participants in such a way that training and physical activity remain integrated into their daily life on the long term. In **chapter 7** it was shown that physical capacity remained stable one year after participation in the HandbikeBattle event. This course of physical capacity was, however, mainly caused by the subgroup of participants that participated in the HandbikeBattle event again at the time of follow-up. The follow-up question would then be why certain participants choose to pursue this goal again, whereas others do not[35]. Goal setting in general should be investigated further, as pursuing a new goal (for example handcycling up the Mont Ventoux or Stelvio, or a less straining and more social touring-event) might be even more effective or motivating than participating in the HandbikeBattle again. From another perspective, it might be that the competitive element plays a role (i.e., improve finish time compared to last year (i.e., by training more or differently, losing weight etc.) or reach the finish before a team mate), or unique aspects of the HandbikeBattle event itself such as the social and nostalgic aspect (i.e., experience the surroundings of the Kaunertal again, meeting peers from last year etc.). In a recent HandbikeBattle study, a follow-up survey was sent to all participants of the 2013 – 2017 editions (N=203) [36]. The response rate was 47% (N=96). Forty-four (46%) participated in the event once, whereas 52 (54%) participated multiple times in the HandbikeBattle event. They were asked in which domains they had experienced benefits or losses of participation in the HandbikeBattle project (domains: fitness, health, handcycling, activities in daily life, personal development). Most responders reported they experienced benefits in fitness (90%), handcycling (87%), personal development (81%), activities in daily life (66%) and health (64%). Very few participants experienced loss in health (8%), fitness (5%), personal development (1%), handcycling

(1%), and activities in daily life (1%). Twenty percent of the respondents who participated in the HandbikeBattle only once experienced losses, which is in contrast to the only 2% of those who participated multiple times ($p \leq 0.01$). In addition, participants were asked about barriers of *current* sport participation. Sixty percent experienced no personal barriers and 64% experienced no environmental barriers. The personal barriers that were mentioned most frequently were lack of time (31%), less able to practice sport due to the disability (17%) and pain (15%). The most frequently mentioned environmental barriers were: transport to sport accommodation takes a lot of time (19%) and not having enough fellow athletes (16%). Those who participated less in sports indicated more personal ($p \leq 0.01$) and environmental barriers ($p = 0.02$), compared to those participating more in sports. The group that participated only once in the HandbikeBattle experienced more personal barriers than the group that participated multiple times ($p \leq 0.01$) [36]. This expresses the potential benefits of the HandbikeBattle for a given subsample of participants, however this may also be the consequence of the personal characteristics of those individuals. Future research is important here. The guided HandbikeBattle training period gives participants and therapists the opportunity to identify and take away barriers for long-term exercise adherence, but who at the start does or who does not pick up this challenge is still open for debate.

Practical implications

Physical capacity testing

Specificity of testing is very important and should be task-specific. As a consequence, wheelchair-specific physical capacity should be measured during a GXT in a wheelchair on a wheelchair ergometer or treadmill, whereas handcycling physical capacity should be measured during a synchronous handcycling GXT. In addition, individualization of the GXT protocol is very important to attain the true peak physical capacity. It is, however, hard to determine which individualized protocol should be used for a particular participant in advance. The predictive models developed in **chapter 2** are imperfect but can be supportive to the decision of which protocol to choose. A typical example is given below (textbox 1). After conducting the individualized GXT, an individualized training regime should be developed. In **chapter 3** it was shown that for the majority of individuals with SCI the VTs can be determined with a high to very high interrater and intrarater reliability. Therefore, defining training intensity zones based on VTs is feasible in the SCI population. However, the large random error within and between raters, and the number of VTs that could not be determined in individuals with tetraplegia show that training schemes based on VTs should be clinically evaluated at the individual level (textbox 2). If the training intensity based on VTs is either too high or too low, VT determination should be critically evaluated and possibly determined again in order to prevent over- or undertraining of that individual.

Typical example: prediction of POpeak (W) based on the theoretical model of chapter 2.

Male participant, 20 years of age.
Has a SCI level Th10, AIS A, TSI is 2 years.
His BMI is 20 kg/m^2. He participated in handcycle training for 5 hours a week during the past three months.
The sports physician in the rehabilitation center wants to perform a GXT but is unsure whether he should choose the 1W/6 s or 1W/4 s ramp protocol. What is the predicted POpeak?

POpeak (W) = 107.05 – (41.13 * sex) + (26.67 * lesion level) + (1.82 * handcycle training) + (0.52 * BMI) + (0.18 * TSI) + (10.92 * completeness) – (0.59 * age).

Sex	0 = male 1 = female
Lesion level	0 = above Th6 1 = equal to / below Th6
Handcycle training	Hours / week during the past three months
TSI	Years
Completeness	0 = motor complete (AIS A / B) 1 = motor incomplete (AIS C / D)

POpeak = 107.05 – (41.13*0) + (26.67*1) + (1.82*5) + (0.52*20) + (0.18*2) + (10.92*0) – (0.59*20) = 142 W.
If the rehabilitation center uses four different ramp protocols, e.g. with a slope of 1W/12 s, 1W/6 s, 1W/4 s or 1W/3 s, based on this prediction it is advised to choose the 1W/4 s protocol (equals 150W after 10 minutes).
If the participant would achieve this POpeak, according to the reference values he would be in the category “good” compared with other HandbikeBattle participants.

NB. In order to evaluate effects of training over time it is very important to choose the exact same protocol at T2. Do not use this prediction model again at T2.

Textbox 1.

“Based on the graded exercise test in my rehab center I have to train at these heart rate values. However, this cannot be correct. When I train, my heart rate gets easily higher than these values, I am not even sweating and I can easily talk. What are the correct intensity zones for me?” Female participant, 51 years old, SCI Th4.

Textbox 2.

For individuals with tetraplegia it is suggested to collect RPE during every step of the GXT to be able to train based on RPE if one or both VTs could not be determined. In addition, it would be interesting to study the associations between training intensity zones based on VTs and the talk test in individuals with SCI. During the talk test, a participant has to repeat a certain text at several time points during exercise. If the participant can speak comfortably the talk test is positive and the intensity is assumed to be below VT1. At higher intensities the participant will equivocate (at VT1), and eventually not speak comfortably (negative talk test, at VT2) [37–39]. Whether this also applies to individuals with SCI during upper-body exercise remains subject for future research. In a previous study with individuals with paraplegia it was shown that the point at which they no longer could talk comfortably was at an intensity of 75 ± 15 %VO_2 reserve which is considered vigorous intensity [14]. This intensity is considered sufficient to improve fitness and this is the first indication that the talk test could indeed be helpful in a population with SCI if the training intensity is unknown or if the intensity based on the VTs does not seem to be correct. The talk test is, like RPE, an easy to use measure without the need for expensive equipment. In addition, by applying the talk test, the participant becomes aware of the intensity and learns to feel how his/her body responses to exercise. **Chapter 3** and **chapter 4** both show that standardization within participants over time is important. If a second GXT is conducted after the training period to evaluate training effects, it is crucial to have the same rater(s) assessing the VTs (**chapter 3**) and to use the same GXT protocol (i.e., the same set-up, setting, starting load, step duration and step size) over time (**chapter 4**).

Handcycle training

Training for the HandbikeBattle is unique in a sense that participants train towards a goal that is not comparable to the training sessions itself, but far beyond. The HandbikeBattle track is long and the intensity is high on a large part of the track due to the inclination. In a previous HandbikeBattle study 17 participants were monitored with heart rate belts during the race [40]. Their average race time was 3:38 ± 1:19 (hh:mm). The intensity of the race was based on HRR. Five zones were defined according to the American College of Sports Medicine [41]: very light (<30% HRR), light (30-39% HRR), moderate (40-59% HRR), vigorous (60-89% HRR), near maximal (90-100% HRR). The participants exercised most of the time at a vigorous intensity (73% of time) (figure 1). It has to be noted that the time that participants stopped to rest and/or drink/eat were included in the results.

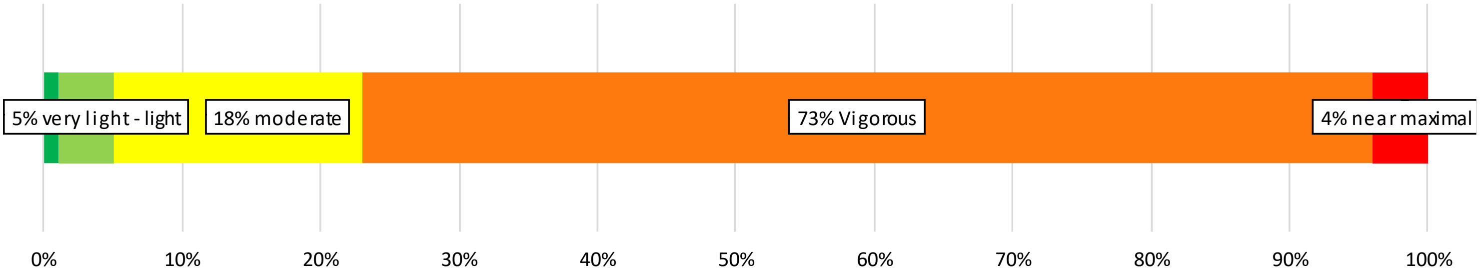


Figure 1. Exercise intensity distribution of the HandbikeBattle race based on %heart rate reserve (N=17) [40].

One could argue that in order to be prepared for such a physical and mental challenge, the training sessions should at least partly resemble this duration and intensity. Although the Netherlands is flat, we know from anecdotal evidence that a lot of participants tried to mimic the steep climb by training in the hills in the south of the Netherlands, the dunes of the coastal line, or by repeatedly climbing bridges and training with extra heavy weights in their handcycle. In the beginning, i.e., during the GXT at T1, several participants were relatively untrained and limited in their peak performance due to peripheral fatigue. Therefore, some of the rehabilitation centers advised to perform strength training twice a week in the beginning of the training period. So far, we have not studied the effects of strength training in the HandbikeBattle population, but we know from previous studies that strength training has a positive effect on body composition, VO_2peak, anaerobic and aerobic POpeak, muscle strength and shoulder pain in wheelchair users [42–44].

In **chapter 5** it is shown that the dose-response relationship between training load and change in physical capacity is not straightforward. Additional explorative analyses showed that training in the moderate intensity zone (zone 2, RPE 5-6) was positively associated with change in VO_2peak and VO_2peak/kg (figure 2). In future studies, we should take a closer look into these associations with, for example, HR monitoring during the training sessions in order to account for variability of the intensity within the training sessions and the quality of the training (e.g. a continuous training might have an overall RPE 5, whereas a HIIT training might also have an overall RPE 5). In addition, the combination of monitoring tools would be preferable as, for example, an increase in RPE in combination with a decrease in HR may be indicative of overreaching [45,46]. Another interesting focus could be the associations among training load and changes in submaximal responses. A previous study showed significant associations between training load and change in PO at the lactate thresholds (LTs) [22]. In this respect, the individualized TRIMP (iTRIMP) is found to be most consistently associated with changes in submaximal responses [18,22,47]. The iTRIMP is a HR-based training load parameter in which an (exponential) individual weighting factor is introduced based on the individual's blood lactate response to incremental exercise [47]. The iTRIMP has shown to be associated with changes in velocity at LT1 (r = 0.68 - 0.87) and LT2 (r = 0.74 - 0.78) in running and hurling [18,47] and with changes in PO at

LT1 (*r* = 0.81) and LT2 (*r* = 0.77) in cycling [22]. It should be evaluated whether this method is feasible in a rehabilitation setting.

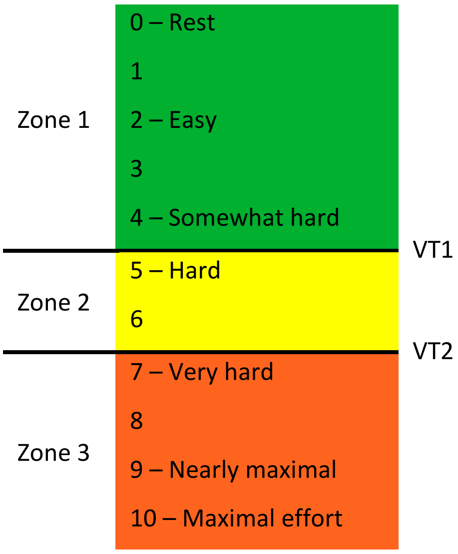


Figure 2. Training intensity distribution. Three intensity zones based on RPE 0-10 scale. VT = ventilatory threshold. Adapted from Seiler et al. [48]

Another important point that has not been touched upon in this thesis is the occurrence of (shoulder) pain and overuse injuries related to training. Although in previous literature [49] it is suggested that handcycling is less prone to lead to shoulder overload, it may play a role in training adherence and training activities, while participants are daily handrim wheelchair users in most cases. In previous years we made an attempt to monitor these aspects in relation to training, but it appeared very difficult to draw any conclusions. Ideally these aspects should be monitored very frequently (i.e., after each training session). In addition, it should be clear whether the pain is acute or chronic, and whether it is related to training (intensity, frequency, duration) or the ergonomic set-up of the handcycle, or the wheelchair, or daily life activities or a combination of these factors. The fact that these aspects are self-reported makes it more difficult to interpret, because of missing data and because participants do not always know what kind of overuse injury they have (in contrast to reports of the team physician / physiotherapist in a professional sports team). Easy to use apps or training logs might increase compliance with monitoring in the HandbikeBattle population.

Effects of participation

The effects of participation in the HandbikeBattle are positive on a group level. There is an increase in physical capacity [50], an increase in life satisfaction which is associated with physical capacity (**chapter**

6), and the increase in physical capacity is sustained on the long term (**chapter 7**). It would be helpful, for implementation purposes and generalization of these findings, to gain more knowledge about the included population. What are the (subconscious) mechanisms that play a role in the decision of the therapists in whether a participant might be suitable to start training for the event, whereas others are not? Very likely, several potential participants refused the invitation to join the HandbikeBattle project. What are the characteristics of these participants? In addition, within the included population, there is a drop-out rate of 17% before T2. Reasons for drop-out include medical reasons such as infections or pressure ulcers, but also motivational reasons or the inability to combine training with other daily activities such as work and family life. In **chapter 6** it was shown that in a subgroup with low mental health scores (i.e. ≤ 72), their mental health improved during the training period and that this improvement was associated with improvement in physical capacity. These findings suggest that an increase in cardiorespiratory fitness might have a positive effect on mental health in individuals with lower mental health scores at baseline. Moreover, this finding suggests that a low mental health is not necessarily a contra-indication to participate. In addition, nonparticipants also had a low mental health score, indicating the thin line between participants and drop-outs. More intensive (mental health) guidance might be warranted for individuals with low mental health; this may prevent them from dropping out during the training period. **Chapter 7** shows that remaining physically fit on the long term is a challenge. These long-term follow-up data are essential to gain knowledge on effects of exercise and training as well as on determinants of maintenance and relapse in physical activity behavior. How people feel during and after exercise may be critical in determining whether they continue [51]. Which factors determine whether people like the actual training and thus continue to do so? It could be the activity itself, the physiological phenomena that go with it, or for example, training with peers and an increased social network. Unravelling these factors would aid in the implementation of physical activity and sport in daily life on the long term.

Limitations and recommendations

The limitations of the study can be summarized in one word: heterogeneity. Participants were included from twelve rehabilitation centers. Therefore, GXTs were conducted by different physicians/test assistants, with different equipment and testing environment, different protocols and all participants had a unique training period. Individualization is key, but standardization of testing equipment and protocol (i.e., ramp versus 1-min stepwise) would probably have resulted in less unexplained variance in, for example, **chapter 2** and **chapter 5**. Moreover, the selection of participants by the rehabilitation centers is more or less a black box: why is a certain (former) patient found suitable for participating in such an event, and another patient not? In addition, patients with a very low physical capacity due to high co-morbidity will be less likely to be found suitable (or be able) to start training. Moreover, a substantial part of individuals might not be motivated to be involved in such a challenge. In other words: for which (former) patients will challenges like this work, and for which individuals will this not work? In the future, more information is needed about inclusion, for example: how many individuals were approached to participate but refused, and why? What were the reasons for drop-out? It would also be interesting to compare certain characteristics of the HandbikeBattle SCI population with the general Dutch SCI population.

In line with this, the current observational design of the study has its limitations. We have to acknowledge that certain questions cannot be answered with this design. From a certain research perspective, a more controlled set-up may have been favorable with fixed training regimes and a control group. However, we must be aware that the study is secondary to the HandbikeBattle event. The HandbikeBattle was perceived as an experiment of life with largely unknown outcomes, thus making our prospective observational design very befitting. The philosophy of the event is that participants learn to take initiative and train together with peers and at home with guidance from the rehabilitation centers. The purpose is that training is incorporated in daily life and that participants remain active on the long term. In the past years, the therapists from the rehabilitation centers have refined this process for their participants. By changing the design of the study to a lab-based controlled set-up, these effects will be lost. Given the multifactorial nature of the study and large quantity of data, for future HandbikeBattle research questions and studies, the possibility of data science technologies should be explored.

As an additional remark, due to the COVID-19 pandemic, the HandbikeBattle event in 2020 did not take place. The monitoring of training in the 2020 cohort was, therefore, incomplete and T2 did not take place. In addition, the T4 measurement of the 2019 cohort did not take place. For **chapter 7** this resulted in fewer participants than anticipated.

Besides these limitations with subsequent recommendations there are also some general recommendations for future HandbikeBattle research.

From physical capacity testing to training

In order to further optimize individualized training, the translation from GXT (protocol) to training should be investigated further. For example, it should be investigated whether the talk test could be a valuable and valid marker to set intensity in individuals with an SCI, and whether it could be used during training sessions in individuals with tetraplegia.

Handcycle training and overuse injuries

In-depth training analyses during the training period are necessary to gain more knowledge about the dose-response relationship between training load and the increase in physical capacity, and associations of this and other daily life activities in regular handrim wheelchair use with overuse injuries. Today there is no indication for increased musculoskeletal overuse risks in handcycling [49,52,53]; handcycling is even seen as a preventive mode of exercise by some [54]. Yet it cannot be neglected that individuals who are dependent on upper body loading in a wheelchair-dependent life might develop overuse injuries. This requires continued research. This should, however, be simple and realistic. For example, with an easy to use app or sensors that automatically store data. Ideally, a combination of subjective and objective internal and external training load measures are measured. A combination of RPE and HR would be valuable to observe decoupling patterns (i.e., high RPE with low HR). In addition, power meters during training are valuable for individuals with tetraplegia such that a combination of PO and RPE can be monitored. With respect to the development and prevention of overuse injuries, next to training load, optimal handcycle configuration is indispensable. Future studies should look further into the best individualized ergonomic set-up with the lowest risk of overuse injuries and best performance [55,56].

Gaining more insight into the physiological demands of an uphill race such as the HandbikeBattle would help in the decision of what would be the best training regime for these individuals. For example, a power balance model can be used to predict the power output needed to finish a race within certain time limits [57]. In previous years, power output (W) data during the race were collected in a subgroup of participants that, together with data from the GXT, can be used to develop such a predictive power model.

Long-term exercise adherence

Which positive effects of training or change in physical capacity are responsible for (long-term) exercise adherence? Examples of topics that should be further investigated are effects on body image, activities of daily living and social inclusion. A large part of the success of why participants commit to training and keep (or quit) training, might be due to (the lack of) social inclusion and peer support. Unfortunately, this was not investigated in the previous HandbikeBattle studies. From anecdotal evidence we know that former HandbikeBattle participants form “teams” themselves to compete in the HandbikeBattle event again next year (as individual participants, irrespective of the rehabilitation center). Or they start planning additional challenges themselves such as handcycling the Berlin Marathon or Mont Ventoux. In addition, in the last years handcycle clubs have grown in number, became affiliated to regular cycling clubs and are spread throughout the Netherlands (from 16 clubs in 2012 to 22 in 2020) (figure 3). Moreover, the Dutch handcycling competition (NHC) has grown from 7 annual competition events in 2012 to 16 in the 2020 season (March – September, that was planned before COVID-19) [58]. It all comes back to the question of which participants stay active in handcycling after the HandbikeBattle, and why. It would be interesting to further investigate whether training for and participating in the HandbikeBattle event meets the six experiential aspects of participation: 1) autonomy (feeling in control and capable of continuing training after the HandbikeBattle), 2) belongingness (being part of a group/community), 3) challenge, 4) engagement (feeling involved), 5) mastery (learning from others), 6) meaning (show to themselves / family / society that they are capable of doing more than they previously thought)[36,59].

Figure 3. Handcycle clubs in the Netherlands

The mission of the HandbikeBattle

The mission of the HandbikeBattle is threefold: 1) to encourage wheelchair users to initiate or keep training after the rehabilitation period, 2) to learn from others and gain confidence to achieve other goals in life, 3) to show that not only elite able-bodied athletes are capable of incredible performances, but recreationally active wheelchair users as well. Have we achieved these goals? In general, I would say: yes! Most participants really train for the event and part of the participant group remains active. Unfortunately, we only know from anecdotal evidence that participants learn much more than only handcycle training (textbox 3). For some participants it is a life-changing experience. On a group level we know that physical capacity and life satisfaction increased and participants experienced benefits in several domains such as personal development (81%) and activities in daily life (66%)[36,50,60]. In the future it would be really interesting to focus on whether participants reach their individual goals that they have set at the start of the training period (textbox 4).

"Crossing the finish line is a very emotional moment. Through the HandbikeBattle I relived my whole rehabilitation period. It took me weeks after the Battle before I had processed everything. I am another person now. My self-image is more positive and I gained confidence. I deal with setbacks different than before. And these are only the mental aspects. Because I started training for the HandbikeBattle I became aware of my (unhealthy) lifestyle. Now I eat healthier and I do a lot of exercise. I notice that my body has a higher capacity and that I recover faster from illnesses". Male participant, 39 years old, SCI Th9.

Textbox 3.

The HandbikeBattle event does not reach everybody. Individuals with a very low physical capacity or with cardiovascular contra-indications will not participate. In addition, individuals who really dislike exercise and are not motivated to train, will also not participate. The question remains how we could reach those individuals. Other life-style interventions might be better suited such as home-based training with an app [61] or a healthy lifestyle rehabilitation training program [62,63]. In addition, in the future we might think of an additional less "extreme" event next to the HandbikeBattle, and promote the membership of a handcycle club. Such that these individuals have a common goal to train for and might experience benefits of achieving this goal in several domains of life. Moreover, it is perhaps (physically and mentally) too soon for patients early in clinical rehabilitation to commit to a challenge like the HandbikeBattle or another strenuous training program, but it is very important that they know that these initiatives exist, for example by watching videos of the event and talking with peers with the same impairment who participated in these training initiatives such that they experience early in rehabilitation what is possible, and that they become motivated and have a (long-term) goal to work towards. It is the responsibility of the rehabilitation field to facilitate this process and to ensure that patients become acquainted with different sport activities and/or a physically active lifestyle early in rehabilitation.

What is your personal goal in the HandbikeBattle?

"I would like to increase fitness, have a lot of social contact with my team mates, become proud of myself, and conquer the mountain!" Male participant, 60 years old, amputation.

Textbox 4.

Reflections on the HandbikeBattle project and its role in the rehabilitation field

The HandbikeBattle is THE example of a large collaboration network in rehabilitation. The HandbikeBattle is an event where rehabilitation teams compete against each other, but everyone who is involved knows that it is about fraternization and mutual support (although we have to be realistic that everybody is grumpy that Rijndam Racers won again in 2019…). Team members become friends for life and therapists and participants keep cycling together after the event. Several participants from one team train together with other teams, and the GXTs of three teams are conducted in one rehabilitation center (without being bothered by rivalry or market forces). The strength of the event is the commitment of the rehabilitation centers and their individualized approach. This event bridges the gap in the rehabilitation field in the Netherlands at all levels. In addition to patients and therapists, policy makers, managers and directors are discussing and networking at the HandbikeBattle finish line in their lycra cycling clothes, and, let us be honest, wipe away a tear when a participant crosses the finish line. This is what the rehabilitation field in the Netherlands should look like. The HandbikeBattle is an event for everybody. We hope to see you all again at the finish line in June 2021.

References

1. Janssen TWJ, van Oers CA, Hollander AP, Veeger HEJ, van der Woude LHV. Isometric strength, sprint power, and aerobic power in individuals with a spinal cord injury. *Med Sci Sports Exerc*. 1993;25:863-870.

2. De Groot S, Kouwijzer I, Valent LJM, van der Woude LHV, Nash MS, Cowan RE. Good association between sprint power and aerobic peak power during asynchronuous arm-crank exercise in people with spinal cord injury. *Disabil Rehabil*. 2019;Epub ahead of print:1-8.

3. Maher JL, Cowan RE. Comparison of 1- versus 3-minute stage duration during arm ergometry in individuals with spinal cord injury. *Arch Phys Med Rehabil*. 2016;97(11):1895-1900.

4. Cowan RE, Guccione AA, Keyser RE, et al. Factors associated with transfer independence in men with paraplegia. Abstract. In: *6th INTERNATIONAL REHABMOVE STATE-OF-THE-ART CONGRESS*. 2018.

5. Gaskill SE, Ruby BC, Walker AVAJ, Sanchez OA, Serfass RC, Leon AS. Validity and reliability of combining three methods to determine ventilatory threshold. *Med Sci Sport Exerc*. 2001;33(11):1841-1848.

6. Shimizu M, Myers J, Buchanan N, et al. The ventilatory threshold: Method, protocol, and evaluator agreement. *Am Heart J*. 1991;122(2):509-516.

7. Leicht CA, Griggs KE, Lavin J, Tolfrey K, Goosey-Tolfrey VL. Blood lactate and ventilatory thresholds in wheelchair athletes with tetraplegia and paraplegia. *Eur J Appl Physiol*. 2014;114(8):1635-1643.

8. Coutts KD, McKenzie DC. Ventilatory thresholds during wheelchair exercise in individuals with spinal cord injuries. *Paraplegia*. 1995;33(7):419-422.

9. Au JS, Sithamparapillai A, Currie KD, Krassioukov AV, MacDonald MJ, Hicks AL. Assessing ventilatory threshold in individuals with motor-complete spinal cord injury. *Arch Phys Med Rehabil*. 2018;99(10):1991-1997.

10. Wasserman K, Hansen J, Sue D, Stringer W, Whipp B. Normal values. In: Wasserman K, Hansen J, Sue D, Stringer W, Whipp B, editorss. *Principles of Exercise Testing and Interpretation: Including Pathophysiology and Clinical Applications*. Lippincott Williams & Wilkins; 2005:160-182.

11. Vainshelboim B, Arena R, Kaminsky LA, Myers J. Reference standards for ventilatory threshold measured with cardiopulmonary exercise testing: The fitness registry and the importance of exercise: A national database. *Chest*. 2020;157(6):1531-1537.

12. Kouwijzer I, Cowan RE, Maher JL, et al. Interrater and intrarater reliability of ventilatory thresholds determined in individuals with spinal cord injury. *Spinal Cord*. 2019;57(8):669-678.

13. Van der Scheer JW, Hutchinson MJ, Paulson T, Martin Ginis KA, Goosey-Tolfrey VL. Reliability and validity of subjective measures of aerobic intensity in adults with spinal cord injury: A systematic review. *PM R*. 2018;10(2):194-207.

14. Cowan R, Ginnity K, Kressler J, Nash M. Assessment of the Talk Test and Rating of Perceived Exertion for exercise intensity prescription in persons with paraplegia. *Top Spinal Cord Inj Rehabil*. 2012;18(3):212-219.

15. Dallmeijer AJ, Zentgraaff IDB, Zijp NI, van der Woude LHV. Submaximal physical strain and peak performance in handcycling versus handrim wheelchair propulsion. *Spinal Cord*. 2004;42(2):91-98.

16. Valent LJM, Dallmeijer AJ, Houdijk H, et al. The individual relationship between heart rate and oxygen uptake in people with a tetraplegia during exercise. *Spinal Cord*. 2007;45(1):104-111.

17. Foster C, Florhaug JA, Franklin J, et al. A New approach to monitoring exercise training. *J Strength Cond Res*. 2001;15(1):109-115.

18. Malone S, Hughes B, Collins K, Akubat I. Methods of monitoring training load and their association with changes across fitness measures in hurling players. *J Strength Cond Res*. 2020;34(1):225-234.

19. Clemente FM, Clark C, Castillo D, et al. Variations of training load, monotony, and strain and dose-

response relationships with maximal aerobic speed, maximal oxygen uptake, and isokinetic strength in professional soccer players. *PLoS One*. 2019;14(12):e0225522.

20. Fitzpatrick JF, Hicks KM, Hayes PR. Dose – response relationship between training load and changes in aerobic fitness in professional youth soccer players. *Int J Sports Physiol Perform*. 2018;Epub ahead of print:1-6.

21. Campos-Vazquez MA, Toscano-Bendala FJ, Mora-Ferera JC, Suarez-Arrones LJ. Relationship between internal load indicators and changes on intermittent performance after the preseason in professional soccer players. *J Strength Cond Res*. 2017;31(6):1477-1485.

22. Sanders D, Abt G, Hesselink MKC, Myers T, Akubat I. Methods of monitoring training load and their relationships to changes in fitness and performance in competitive road cyclists. *Int J Sports Physiol Perform*. 2017;12(5):668-675.

23. Vermeire KM, Vandewiele G, Caen K, Lievens M, Bourgois JG, Boone J. Training progression in recreational cyclists: no linear dose-response relationship with training load. *J Strength Cond Res*. 2019;Epub ahead of print:1-6.

24. Bourdon PC, Cardinale M, Murray A, et al. Monitoring athlete training loads: consensus statement. *Int J Sports Physiol Perform*. 2017;12(Suppl 2):161-170.

25. Verhagen E, Gabbett T. Load, capacity and health: critical pieces of the holistic performance puzzle. *Br J Sports Med*. 2019;53(1):5-6.

26. Sweet SN, Ginis KAM, Tomasone JR. Investigating intermediary variables in the physical activity and quality of life relationship in persons with spinal cord injury. 2013;32(8):877-885.

27. Van Koppenhagen CF, Post MWM, van der Woude LHV, et al. Recovery of life satisfaction in persons with spinal cord injury during inpatient rehabilitation. *Am J Phys Med Rehabil*. 2009;88(11):887-895.

28. Van Leeuwen CMC, Post MWM, Hoekstra T, van der Woude LHV et al. Trajectories in the course of life satisfaction after spinal cord injury: Identification and predictors. *Arch Phys Med Rehabil.* 2011;92(2):207-213.

29. Van Leeuwen CMC, Post MWM, van Asbeck FWA, et al. Life satisfaction in people with spinal cord injury during the first five years after discharge from inpatient rehabilitation. *Disabil Rehabil*. 2012;34(1):76-83.

30. Bassett RL, Ginis KAM, Buchholz AC, SHAPE SCI Research Group. A pilot study examining correlates of body image among women living with SCI. *Spinal Cord*. 2009;47(6):496-498.

31. Reboussin BA, Rejeski WJ, Martin KA, et al. Correlates of satisfaction with body function and body appearance in middle- and older aged adults: The activity counseling trial (ACT). *Psychol Health*. 2000;15:239-254.

32. Astorino TA, Harness ET. Improved quality of life and body satisfaction in response to activity-based therapy in adults with spinal cord injury. *Neuroimmunol Neuroinflammation*. 2020;7:40-50.

33. Ginis KAM, Jetha A, Mack DE, Hetz S. Physical activity and subjective well-being among people with spinal cord injury: a meta-analysis. *Spinal Cord*. 2010;48(1):65-72.

34. De Vet H, Terwee C, Mokkink L, Knol D. Interpretability. In: *Measurement in Medicine*. Cambridge University Press; 2011:245-261.

35. Ajzen I, Kruglanski AW. Reasoned action in the service of goal pursuit. *Psychol Rev*. 2019;126(5):774-786.

36. De Groot S, Kouwijzer I, Valent LJM, et al. Sport participation after the HandbikeBattle: benefits, barriers, facilitators from the event—a follow-up survey. *Spinal Cord Ser Cases*. 2020;6(1):54.

37. Reed JL, Pipe AL. The talk test: A useful tool for prescribing and monitoring exercise intensity. *Curr Opin Cardiol*. 2014;29(5):475-480.

38. Persinger R, Foster C, Gibson M, Fater D, Porcari J. Consistency of the Talk Test for exercise prescription.

Med Sci Sports Exerc. 2004;36(9):1632-1636.

39. Rodriguez-Marroyo J, Villa J, Garcia-Lopez J, Foster C. Relationship between the talk test and ventilatory thresholds in well-trained cyclists. *J Strength Cond Res*. 2013;27(7):1942-1949.

40. De Groot S, Postma K, van Vliet L, Timmermans R, Valent LJM. Mountain time trial in handcycling: Exercise intensity and predictors of race time in people with spinal cord injury. *Spinal Cord*. 2014;52(6):455-461.

41. Garber CE, Blissmer B, Deschenes MR, et al. Quantity and quality of exercise for developing and maintaining cardiorespiratory, musculoskeletal, and neuromotor fitness in apparently healthy adults: guidance for prescribing exercise. *Med Sci Sport Exerc*. 2011;43(7):1334-1359.

42. Nevin J, Smith P, Waldron M, et al. Efficacy of an 8-week concurrent strength and endurance training program on hand cycling performance. *J Strength Cond Res*. 2018;32(7):1861-1868.

43. Nash MS, van de Ven I, van Elk N, Johnson BM. Effects of circuit resistance training on fitness attributes and upper-extremity pain in middle-aged men with paraplegia. *Arch Phys Med Rehabil*. 2007;88(1):70-75.

44. Jacobs PL. Effects of resistance and endurance training in persons with paraplegia. *Med Sci Sports Exerc*. 2009;41(5):992-997.

45. Fusco A, Knutson C, King C, et al. Session RPE during prolonged exercise training. *Int J Sports Physiol Perform*. 2020;15(2):292-294.

46. Rietjens GJWM, Kuipers H, Adam JJ, et al. Physiological, biochemical and psychological markers of strenuous training-induced fatigue. *Int J Sports Med*. 2005;26(1):16-26.

47. Manzi V, Iellamo F, Impellizzeri F, D'Ottavio S, Castagna C. Relation between individualized training impulses and performance in distance runners. *Med Sci Sport Exerc*. 2009;41(11):2090-2096.

48. Seiler KS, Kjerland GØ. Quantifying training intensity distribution in elite endurance athletes: Is there evidence for an "optimal" distribution? *Scand J Med Sci Sport*. 2006;16(1):49-56.

49. Arnet U, van Drongelen S, Scheel-Sailer A, van der Woude LHV, Veeger DHEJ. Shoulder load during synchronous handcycling and handrim wheelchair propulsion in persons with paraplegia. *J Rehabil Med*. 2012;44(3):222-228.

50. Hoekstra SP, Valent LJM, Gobets D, van der Woude LHV, de Groot S. Effects of four-month handbike training under free-living conditions on physical fitness and health in wheelchair users. *Disabil Rehabil*. 2017;39(16):1581-1588.

51. Ekkekakis P, Parfitt G, Petruzzello SJ. The pleasure and displeasure people feel when they exercise at different intensities: Decennial update and progress towards a tripartite rationale for exercise intensity prescription. *Sport Med*. 2011;41(8):641-671.

52. Valent LJM, Dallmeijer AJ, Houdijk H, et al. Effects of hand cycle training on physical capacity in individuals with tetraplegia: A clinical trial. *Phys Ther*. 2009;89(10):1051-1060.

53. Valent LJM, Dallmeijer AJ, Houdijk H, Slootman JR, Janssen TWJ, van der Woude LHV. Effects of hand cycle training on wheelchair capacity during clinical rehabilitation in persons with a spinal cord injury. *Disabil Rehabil.* 2010;32(26):2191-2200.

54. Van der Woude LHV, Dallmeijer AJ, Janssen TWJ, Veeger HEJ. Alternative modes of manual wheelchair ambulation: an overview. *Am J Phys Med Rehabil*. 2001;80(10):765-777.

55. Stone B, Mason BS, Bundon A, Goosey-Tolfrey VL. Elite handcycling: a qualitative analysis of recumbent handbike configuration for optimal sports performance. *Ergonomics*. 2019;62(3):449-458.

56. Stone B, Mason BS, Warner MB, Goosey-Tolfrey VL. Horizontal crank position affects economy and upper limb kinematics of recumbent handcyclists. *Med Sci Sports Exerc*. 2019;51(11):2265-2273.

57. Groen WG, van der Woude LHV, de Koning JJ. A power balance model for handcycling. *Disabil Rehabil*.

2010;32(26):2165-2171.

58. Handbiken.nl. www.handbiken.nl. Accessed August 14, 2020.

59. Shirazipour CH, Evans MB, Leo J, Lithopoulos A, Martin Ginis KA, Latimer-Cheung AE. Program conditions that foster quality physical activity participation experiences for people with a physical disability: a systematic review. *Disabil Rehabil*. 2020;42(2):147-155.

60. Kouwijzer I, de Groot S, van Leeuwen CMC, et al. Changes in quality of life during training for the HandbikeBattle and associations with cardiorespiratory fitness. *Arch Phys Med Rehabil*. 2020;101(6):1017-1024.

61. WHEELS app. www.allesoversport.nl/artikel/leefstijlapp-voor-rolstoelgebruikers-met-een-dwarslaesie-of-beenamputatie. Accessed August 14, 2020.

62. Behandelmodule gezonde leefstijl. Rijndam Revalidatie. www.rijndam.nl/speciale-programmas/behandelmodule-gezonde-leefstijl. Accessed August 14, 2020.

63. Blokland IJ, van Bennekom CAM, Appel R, Groot FP, Houdijk H. Fysiek Profiel - Fysieke testen en training binnen de revalidatie. *Ned Tijdschr voor Revalidatiegeneeskd*. 2018:149-152.