Doctoral theses at NTNU, 2018:342

Julia Kathrin Baumgart

Determination of peak oxygen uptake, the anaerobic threshold and efficiency in Paralympic sitting sports

- With emphasis on the influence of test mode and protocol

NTNU
Kunnskap for en bedre verden

Doctoral thesis

Julia Kathrin Baumgart

Doctoral theses at NTNU, 2018:342

NTNU
Norwegian University of Science and Technology
Thesis for the Degree of
Philosophiae Doctor
Faculty of Medicine and Health Sciences
Department of Neuromedicine and Movement
Science

NTNU

ISBN 978-82-326-3472-9 (printed ver.)
ISBN 978-82-326-3473-6 (electronic ver.)
ISSN 1503-8181

NTNU
Kunnskap for en bedre verden

Julia Kathrin Baumgart

Determination of peak oxygen uptake, the anaerobic threshold and efficiency in Paralympic sitting sports

- With emphasis on the influence of test mode and protocol

Thesis for the Degree of Philosophiae Doctor

Trondheim, October 2018

Norwegian University of Science and Technology
Faculty of Medicine and Health Sciences
Department of neuromedicine and Movement Science

NTNU
Norwegian University of
Science and Technology

NTNU
Norwegian University of Science and Technology

Thesis for the Degree of Philosophiae Doctor

Faculty of Medicine and Health Sciences
Department of Neuromedicine and Movement Science

ISBN 978-82-326-3472-9 (printed ver.)
ISBN 978-82-326-3473-6 (electronic ver.)
ISSN 1503-8181

Doctoral theses at NTNU, 2018:342

Printed by NTNU Grafisk senter

EN.DUR.ANCE (noun) –

The power to withstand pain or hardships;

The ability or strength to continue despite fatigue, stress or adverse conditions.

LIST OF PAPERS

Paper I – Baumgart JK, Brurok B, Sandbakk Ø. Peak Oxygen Uptake in Paralympic Sitting Sports: A Systematic Literature Review, Meta- and Pooled-Data Analysis. PLoS One. 2018;13(2):e0192903. doi:10.1371/journal.pone.0192903

Paper II – Baumgart JK, Gürtler L, Ettema G, Sandbakk Ø. Comparison of peak oxygen uptake and exercise efficiency between upper-body poling and arm crank ergometry in trained paraplegic and able-bodied participants. In revision. Eur J Appl Physiol[a]

Paper III – Baumgart JK, Skovereng K, Sandbakk Ø. Comparison of Peak Oxygen Uptake and Test-Retest Reliability of Physiological Parameters between Closed-End and Incremental Upper-Body Poling Tests. Front Physiol. 2017;8(857). doi:1.3389/fphys.2017.00857 (including corrigendum)

Paper IV – Baumgart JK, Moes M, Skovereng K, Ettema G, Sandbakk Ø. Examination of gas exchange and blood lactate thresholds in Paralympic athletes during upper-body poling. In revision. PLoS One[b]

---

[a] This article is now published: https://www.ncbi.nlm.nih.gov/pubmed/29936549

[b] This article has just been accepted for publication: link is not yet available

SUMMARY

Endurance exercise performance is mainly determined by peak oxygen uptake (VO_{2peak}), percentage of VO_{2peak} used at the anaerobic threshold and efficiency. Furthermore, a high VO_{2peak} increases tolerance for higher training volumes and reduces recovery times between training sessions. However, there is lack of knowledge on how to assess endurance exercise performance in sport-specific modes in athletes with disabilities competing in different Paralympic sitting sports disciplines. For sitting Para cross-country skiers, Para biathletes and Para ice hockey players, the upper-body poling mode (UBP) may be the most sport-specific.

Therefore, the main aim of the studies conducted in the course of this PhD was to determine VO_{2peak}, the anaerobic threshold and efficiency in Paralympic sitting sports, with emphasis on the influence of test mode and protocol on these factors.

Initially, a systematic literature review (Paper I) was conducted, including a meta-analysis to investigate differences in VO_{2peak} between Paralympic sitting sports and a pooled-data analysis to investigate the effect of age, sex, body-mass, disability and test mode on VO_{2peak}. Thereafter, three experimental studies (Paper II – IV) were conducted using the UBP mode. In the three experimental studies, VO_{2peak} and efficiency were compared between the UBP and the arm crank ergometry (ACE) mode (Paper II), the test-retest reliability of physiological parameters was investigated in the UBP mode (Paper III), and aerobic and anaerobic thresholds were investigated in the UBP mode (Paper IV).

The main findings were that differences in VO_{2peak} values between Paralympic sitting sports were fairly well reflected by the sport-specific demands and therefore, highest in sports with continuously high physical efforts such as Para cross-country sit skiing. In wheelchair athletes, being a man, not being tetraplegic or having an amputation compared to being paraplegic and testing in the wheelchair ergometry or wheelchair treadmill mode compared to the ACE mode was favourable for high VO_{2peak} (Paper I). VO_{2peak} did not differ between restricted UBP and restricted ACE, whereas VO_{2peak} was lower in the paraplegic participants compared to the able-bodied cross-country skiers. Furthermore, exercise efficiency was lower in restricted UBP compared to restricted ACE (Paper II). In the UBP mode, the relative test-retest reliability of VO_{2peak} during a 1-min, a 3-min and an incremental upper-body poling test was high in able-bodied cross-country skiers. However, VO_{2peak} was significantly higher during

a 3-min closed-end and an open-end incremental test compared to a 1-min closed-end test (Paper III). When testing Para ice hockey players in the UBP mode, the physiological outcome parameters identified at the ventilatory threshold were significantly higher than the ones identified with fixed methods at the first lactate threshold, even though both are used to determine the aerobic threshold. In comparison, the outcome parameters at the respiratory compensation threshold and the second lactate threshold, which both determine the anaerobic threshold, and the ones identified at the ventilatory threshold were closely located. Furthermore, continuous linear and curvilinear (i.e. no-breakpoint) models fitted the gas exchange data of most Para ice hockey players during UBP better than breakpoint models (Paper IV).

Overall, the findings of the studies conducted during this PhD show that participating in endurance sports with continuously high movement demands, being a man, not being tetraplegic or having an amputation, as well as testing in a wheelchair treadmill or wheelchair ergometer mode are favourable for high absolute and body-mass VO_{2peak} values. Movement differences between UBP and ACE do not seem to have an impact on VO_{2peak} when the upper-body is restricted, but the discontinuous power production in UBP leads to lower efficiency. Compared to able-bodied participants, spinal-cord injury related limitations negatively impact on VO_{2peak} but not on efficiency in paraplegic participants. In UBP, both the 3-min and an incremental test are reliable VO_{2peak} tests. Furthermore, the ventilatory threshold and first lactate threshold cannot be used interchangeably to identify the aerobic threshold. The close location of the ventilatory threshold, the respiratory compensation threshold and the second lactate threshold does not allow us to distinguish the aerobic and anaerobic threshold, indicating the presence of only one threshold in athletes with a disability exercising in UBP. In addition, it is questionable if clear breakpoints exist in the gas exchange data of athletes with a disability during UBP.

SAMMENDRAG

En utholdenhetsprestasjon avhenger i hovedsak av et høyt peak oksygen opptak (VO_{2peak}), prosentandelen av VO_{2peak} ved anaerob terskel og arbeidsøkonomi. I tillegg vil en høy VO_{2peak} øke toleransen for store treningsmengder samt redusere restitusjonstid i mellom treningsøkter. Derimot er kunnskapen om hvordan man tester idretts-spesifikk utholdenhet hos utøvere med ulike typer funksjonsnedsettelser og idretter begrenset. For sittende langrennsløpere, sittende skiskyttere og Para hockey spillere er overkroppsstaking trolig den mest idretts-spesifikke test-modaliteten.

Hovedmålsettingene med studiene som har blitt gjennomført i denne doktorgraden har derfor vært å undersøke VO_{2peak}, prosentandelen av VO_{2peak} ved anaerob terskel, og arbeidsøkonomi i Paralympiske sittende idretter, med fokus på å forstå test-modalitetens og testprotokollens innvirkning på disse faktorene.

En systematisk litteraturgjennomgang og Meta-analyse for å undersøke om det er forskjeller i VO_{2peak} mellom ulike sittende Paralympiske idretter ble gjennomført. I tillegg, ble en Meta-regresjon benyttet for å se på effekt av alder, kjønn, kroppsvekt, funksjonsnedsettelse og test-modalitet på VO_{2peak} (Paper I). Deretter ble tre eksperimentelle studier med bruk av overkroppsstaking som test-modalitet gjennomført: 1) VO_{2peak} og arbeidsøkonomi under overkroppsstaking og håndsykling ble sammenlignet, 2) reliabilitet av fysiologiske variabler under overkroppsstaking ble undersøkt, og 3) ulike mål for aerob og anaerob terskel under overkroppsstaking ble sammenlignet (Paper II – IV).

Hovedfunnene var at VO_{2peak} i de forskjellige sittende Paralympiske idrettene gjenspeiler de idretts-spesifikke kravene og er derfor høyest i idretter med høye bevegelseskrav, som i langrennspigging. I idretter der utøvere konkurrerer i rullestol, er det å være mann, ikke tetraplegiker eller å ha en amputasjon istedenfor paraplegi, og å teste på en rullestol-ergometer eller på en rullestol-tredemølle sammenlignet med å teste på en håndsykkel, gunstig for en høy VO_{2peak} (Paper I). VO_{2peak} var ikke forskjellig mellom overkroppsstaking og håndsykling, mens VO_{2peak} var lavere hos paraplegikere sammenlignet med funksjonsfriske langrennsløpere. Videre var arbeidsøkonomien dårligere under overkroppsstaking sammenlignet med håndsykling (Paper II). Under overkroppsstaking hos funksjonsfriske langrenns løpere er test-retest reliabiliteten for VO_{2peak} ved en 1-min, 3-min og

en inkrementell test høy. Derimot var VO_{2peak} under den 3-min og inkrementelle testen signifikant høyere sammenlignet med 1-min testen (Paper III). Når Para hockey spillere ble testet under overkroppsstaking, var prosentdelen av VO_{2peak} betydelig høyere på den ventilatoriske terskelen sammenlignet med den første laktat-terskelen. Videre var prosentandel av VO_{2peak} på respiratorisk kompensasjons terskel og på andre laktatterskel relativt lik den på ventilatorisk terskel (Paper IV).

Kort oppsummert, viser hovedfunnene av studiene gjennomført i denne doktorgraden at det er gunstig for en høy VO_{2peak} å konkurrere i idretter med høye kontinuerlige bevegelseskrav, samt å være mann, ikke å være tetraplegiker eller å ha en amputasjon, og å teste på rullestol-ergometer eller rullestol-tredemølle. Ulikheter i bevegelsen mellom overkroppsstaking og håndsykling resulterte ikke i forskjeller i VO_{2peak}, mens fysiologiske begrensninger knyttet til en paraplegi fører til en lavere VO_{2peak} sammenlignet med verdiene vi finner hos funksjonsfriske langrennsutøverne. I tillegg var arbeidsøkomien dårligere i overkroppsstaking enn håndsykling, noe som ikke var forskjellig mellom paraplegikere og funksjonsfriske langrennsutøvere. Under overkroppsstaking er VO_{2peak} reliabel både i en 3-min og en inkrementell test. I tillegg kan ventilatorisk terskel og første laktatterskel ikke bli brukt om hverandre for å identifisere aerob terskel. Ventilatorisk terskel, respiratorisk kompensasjons terskel og andre laktatterskel er veldig lik, noe som kan indikere at én og ikke to terskler eksisterer i utøvere med en funksjonsnedsettelse når de blir testet i en overkroppsmodalitet.

ABBREVIATIONS

AB	Able-bodied participants
ACE	Arm crank ergometry
AMP	Participants with an amputation
AT	Aerobic threshold
ANT	Anaerobic threshold
LT1	First lactate threshold
LT2	Second lactate threshold
MR	Metabolic rate
PARA	Participants with paraplegia
Para	Paralympic
PO	Power output
RCT	Respiratory compensation threshold
RER	Respiratory exchange ratio
RPE	Rating of perceived exertion
SCI	Spinal cord injury
TETRA	Participants with tetraplegia
UBP	Upper-body poling
VE	Minute ventilation
VO_2	Oxygen uptake
VO_{2max}	Maximal oxygen uptake
VO_{2peak}	Peak oxygen uptake
% of VO_{2peak}	Percentage of peak oxygen uptake
VCO_2	Carbon dioxide production
VCO_{2peak}	Peak carbon dioxide production
VT	Ventilatory threshold
WERG	Wheelchair ergometry
WTR	Wheelchair treadmill

TABLE OF CONTENTS

APPENDICES

Appendix 1 – Supplementary data 1: VO_{2peak} in different upper-body exercise modes

Appendix 2 – Supplementary data 2: Influence of workload increases and increment duration on VO_{2peak}

BACKGROUND

With its origin in the Stoke Mandeville Games in 1948, adaptive sports initially constituted an effective means of augmenting rehabilitation outcomes for people with disabilities[1]. However, being successful at the Paralympic Games of today requires top-level performance. The Paralympic games are now the world's second largest sporting event, with an increasing number of sports disciplines competed in, participating nations and athletes. 4342 athletes with 10 different eligible physical impairments from 159 nations participated in 23 summer disciplines in Rio 2016 and 569 athletes from 49 countries in 6 winter disciplines in Pyoengchang 2018[2]. Of these, 16 of the summer sports and 5 of the winter sports disciplines have at least one sitting class. Depending on the eligibility criteria of each sitting sports discipline, athletes with impaired muscle power, impaired passive range of movement, limb deficiency, leg length difference, hypertonia, ataxia and athetosis are allowed to compete[3]. Along with advances in equipment and technology, improvement in physical performance has led to increasingly smaller margins between winning a medal or not[4]. Depending on the sports discipline, physical performance is determined by a different relative contribution of strength, speed, flexibility, technique and endurance.

The endurance demands of the Paralympic sitting sport disciplines vary within a spectrum from typical endurance sports requiring high aerobic energy delivery over sustained periods to those performed with relatively low levels of displacement and corresponding low aerobic demands[5, 6]. However, endurance is also important for athletes in sports with low levels of displacement to be able to tolerate higher training volumes and reduce recovery time between training sessions[7]. Endurance is most commonly tested in the arm crank ergometry mode (ACE) with a variety of test protocols[8-13]. However, knowledge is lacking on how to test endurance exercise performance in some of the sport-specific modes in athletes with different disabilities competing in different sitting sports disciplines.

Determinants of endurance exercise performance

In endurance sports, three factors mainly determine endurance exercise performance: 1) maximal oxygen uptake (VO_{2max}), 2) percentage of VO_{2max} used at the anaerobic threshold (ANT) and 3) efficiency[14-17]. VO_{2max} and the ANT are the main determinants of the oxygen uptake (VO_2) that can be sustained over a given period of time, also described as 'performance VO_2' by Joyner et al.[16, 17] (Figure 1). The performance VO_2 and efficiency then determine the speed or power that can be sustained over time.

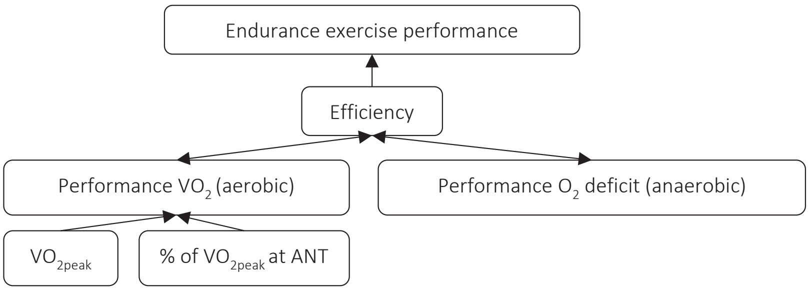


Figure 1. Main physiological factors related to endurance exercise performance (modified from Joyner et al.[17]).
Peak oxygen uptake (VO_{2peak}), oxygen uptake (VO_2), oxygen (O_2), anaerobic threshold (ANT)

VO_{2max} is an indicator of humans' maximal ability to deliver and utilize energy aerobically during dynamic exercise involving large muscle groups[18, 19]. VO_{2max} is determined by the variables that define oxygen delivery in the Fick equation (VO_2 = cardiac output (central factor) × arterial-venous O_2 content difference (peripheral factor))[20]. Cardiac output is equal to the stroke volume times the heart rate (HR)[21]. Oxygen extraction is measured by the difference in oxygen saturation of blood going to the tissues in the arteries and blood returning to the heart in the veins[21]. However, when testing individuals with different disabilities in an upper-body mode, VO_{2max} is rarely reached due to testing with a limited amount of active muscle mass and due to disability-specific limitations[22-26]. Therefore, peak oxygen uptake (VO_{2peak}), which denotes the highest oxygen uptake during exercise to voluntary exhaustion, is used instead[27].

Two physiological thresholds are commonly described in the literature (Figure 2): the aerobic threshold (AT) and the anaerobic threshold (ANT). The AT separates low- from moderate-intensity exercise and is determined by the ventilatory threshold (VT) or the first lactate threshold (LT1)[28, 29]. The VT is determined by a breakpoint in the VO_2-VCO_2 (V-slope method)[29] or VE/VO_2-time relationship (ventilatory equivalent method)[30], and the LT1 by a fixed rise in BLa above resting levels[31, 32] or a breakpoint in the log-log transformed VO_2-BLa data[33]. The ANT separates moderate- from high-intensity exercise and is determined by the respiratory compensation threshold (RCT)[29] or the second lactate threshold (LT2)[28]. The ANT marks the point beyond which any attempt of the body to maintain aerobic equilibrium fails, and fatigue (often indicated by an exponential rise in blood lactate (BLa) concentration) starts to increase[34]. The RCT is determined by a breakpoint in the VE/VCO_2-time (ventilatory equivalent method)[30] or VCO_2-VE relationship[29], and the LT2 by the D_{max} method[35] or a fixed BLa value of 4mmol·L^{-1} [36]. In contrast, it has been argued that the changes in gas exchange with increasing work rate are continuous transitions where fatigue gradually accumulates rather than clear breakpoints[37].

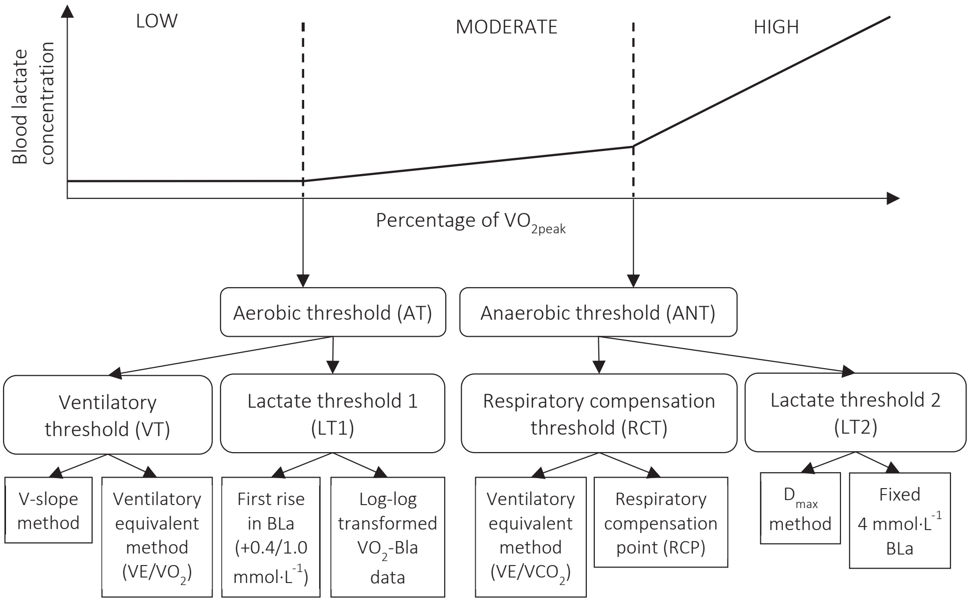


Figure 2. Aerobic and anaerobic threshold determined by gas exchange thresholds (ventilatory threshold and respiratory compensation threshold, respectively) and blood lactate thresholds (lactate threshold 1 and 2, respectively) (modified from Binder et al.[28]).

Peak oxygen uptake (VO_{2peak}), minute ventilation (VE), oxygen uptake (VO_2), blood lactate (BLa)

In exercise physiology, the overall definition of efficiency is the ratio of work produced to the energy expended per time unit[38]. Various definitions of efficiency – such as gross efficiency, net efficiency and delta efficiency – have been used in studies investigating energy expenditure in different forms of upper-body[9, 39, 40], lower-body[41-45] and whole-body exercise[46]. Gross efficiency is the ratio of power output (PO) produced to the total metabolic rate (MR)[47]. Gross efficiency uses the entire human body as energy converting system without any form of baseline subtraction[38, 47]. Net efficiency is the ratio of PO produced to the MR subtracted by the resting MR[47]. Delta efficiency is the ratio of delta PO produced to delta MR[47]. In brief, the challenges with net efficiency and delta efficiency, which are outlined more in detail by Ettema and Lorås (2009)[48], concern the assumption that the processes related to the resting metabolism are independent of the processes associated with producing work. In comparison, gross-efficiency is a theoretically sound concept. However, it is affected by the diminishing effect of the metabolism rate related to zero PO production with increasing PO[47]. Therefore, next to gross efficiency, the entire PO-MR relationship is used as a measure of exercise efficiency in this thesis. In this relationship, both the offset (i.e., the interpolated MR at zero PO) and the incline (i.e., the increase in MR to a given increase in PO) are interpreted as a whole, which integrates the understanding gained from net efficiency, delta efficiency and gross efficiency.

Upper-body-related physiological limitations

When testing individuals with different disabilities in an upper-body mode, differences in physiological responses compared to the lower-body mode might be attributed to testing in the upper-body mode itself as well as to disability-related physiological limitations. Even in able-bodied participants (AB), VO_{2peak} during upper-body exercise was 20-30 % lower compared to lower-body exercise[49, 50]. This is related to a limited amount of active muscle mass during upper-body exercise[26], possible restrictions to local muscle blood flow[26, 50, 51] and high levels of local muscle fatigue[52], which lead to the cardiorespiratory system not being fully taxed. In line with this, several studies show that at a given submaximal PO, HR and BLa are higher in upper-body compared to lower-body exercise[53, 54]. Furthermore, at a given submaximal PO, VO_2 (and with this MR) was shown to be higher in upper-body compared to lower-body or whole-body exercise in some studies[53-58], whereas in other studies that used a more similar testing position, VO_2 was not different[46, 50, 59]. Furthermore, at a given VO_2, higher HR and BLa were found during upper-body exercise compared to lower-body exercise[53, 54]. In line with the generally higher physiological responses to upper-body exercise compared to lower-body exercise at a given intensity, the percentage of peak oxygen uptake (% of VO_{2peak}) at the aerobic and anaerobic threshold was shown to occur at lower exercise intensities during upper-body exercise[60, 61], even though some studies also found a similar % of VO_{2peak}[62, 63].

Disability-related physiological limitations

Irrespective of the upper-body mode used during exercise testing, VO_{2peak} values were found to be even lower in paraplegic participants (PARA) compared to AB[11, 64, 65]. Depending on the level of the spinal cord injury (SCI), the lower VO_{2peak} is likely related to a limited active muscle mass during exercise testing[23], lack of blood redistribution below the level of injury including the splanchnic vascular bed in individuals with a SCI above Th6[24], and autonomic dysfunction including disruption of sympathetic stimulation to the myocardium in individuals with SCI above Th1[66, 67]. Whereas the influence of a SCI on VO_{2peak} is relatively well investigated, knowledge on the influence of other common disabilities, such as an amputation, poliomyelitis or spina bifida is scarce. One study found lower peak PO during an incremental test to exhaustion in athletes with low and high SCI (below and above Th6, respectively) compared to athletes with unilateral lower-limb amputations[68], which might arguably also reflect a difference in VO_{2peak}. Additionally, another study found similar absolute VO_{2peak} in PARA athletes compared to athletes with double, above knee amputations[69]. Furthermore, at a fixed submaximal PO or the same % of VO_{2peak}, a higher HR was found in PARA compared to AB, which compensates for the reduced stroke volume and leads to similar cardiac output[25, 70]. Furthermore, disability-related limitations do not seem to influence the AT, as it was shown to occur at a higher % of VO_{2peak} in PARA compared to AB[71, 72]. Additionally, efficiency was also higher in PARA compared to AB[73-75].

Upper-body test mode and protocol

ACE is the test mode most commonly used to assess endurance in persons who are primarily able to use their upper-body during exercise[8-13], such as Paralympic athletes. However, sport-specificity of the test mode is of importance in reflecting VO_{2peak}, percentage of VO_{2peak} used at the anaerobic threshold and efficiency in the respective sport[76]. In wheelchair athletes, the wheelchair ergometry (WERG) or the wheelchair treadmill (WTR) mode may provide a more sport-specific alternative compared to ACE mode. In Paralympic (Para) ice hockey players, sitting Para cross-country skiers and sitting Para biathletes, the upper-body poling (UBP) movement may be more sport-specific than the ACE mode. During WERG, WTR and UBP, the synchronous movement of the hands allows more movement of the trunk compared to ACE, where the asynchronous movement of the hands limits dynamic trunk involvement. A higher VO_{2peak} might, therefore, be expected in the WERG, WTR and UBP compared to the ACE mode due to more active muscle mass. Some studies are in line with this and show higher values in the WERG or WTR mode compared to the ACE mode[77, 78], whereas others show no differences[9, 78-81]. Furthermore, whether VO_{2peak} differs between the UBP and ACE mode has not yet been investigated.

In contrast to the lack of agreement on the presence or absence of differences in VO_{2peak} between modes, studies consistently show that the WERG or WTR is less efficient than the ACE mode in wheelchair athletes, physically less active PARA as well as AB[39, 40, 81]. This is likely caused by higher coordinative demands of the discontinuous movement and by production of power during a shorter portion of each cycle in the WERG and WTR modes[40]. In line with this, exercise efficiency is expected to be lower in the UBP compared to the ACE mode, even though the size of the difference in exercise efficiency remains to be investigated.

Various test protocols have been used to determine VO_{2peak} in upper-body modes. The most common test protocol comprises different incremental workload increases until voluntary exhaustion[82-87]. Smith et al.[86] found no differences in VO_{2peak} between 6 watt/min and 12 watt/min increases during an incremental test in the ACE mode. However, in an UBP mode, the influence of different increment durations and workload increases on VO_{2peak} has not yet been investigated. Furthermore, a 3-min self-paced all-out test has been used to assess VO_{2peak} in ACE[88] and UBP[46, 89]. No differences in VO_{2peak} between the 3-min and the incremental test were found in cyclists[90]. In addition, the ability to increase VO_2 more rapidly than during the 3-min and the incremental test might be important in sports where high POs are produced over short time. An example of such a sport is Para cross-country sit skiing, where hard work is performed in the steep uphill terrain, with subsequent recovery in the downhills. Therefore, it might be of interest to investigate VO_{2peak} in a test of shorter duration. However, in upper-body exercise modes, including UBP, it has not yet been investigated whether VO_{2peak} differs between closed-ended tests of different duration and an incremental protocol. In addition, the test-retest reliability of such test protocols needs to be established before meaningful differences between athletes and repeated tests can be interpreted.

Various methods have been employed to identify the VT and LT1, which are used to determine the AT, and the RCT and the LT2, which are used to determine the ANT[29, 31, 33, 34]. The assumption that the VT corresponds with the LT1, and the RCT with the LT2, are based on the initial studies by Beaver et al.[29, 33] and Wassermann et al.[34, 91, 92]. In line with this, the VT and the LT1 and the RCT and LT2 were shown to have high correlations in able-bodied participants during running and cycling[93, 94], whereas another study found low correlations during cycling[95]. In an upper-body mode, studies that correlate gas exchange and BLa thresholds are missing, but Ribeiro et al.[96] found no significant differences between the VT and the LT1 in able-bodied swimmers, whereas Leicht et al.[97] found significant differences between the VT and LT1, but no significant differences between the RCT and LT2 in tetraplegic participants (TETRA) and PARA during exercise on the wheelchair treadmill. However, whether these gas exchange and BLa threshold methods can be used interchangeably to determine the AT and ANT in athletes with different disabilities in an upper-body exercise mode remains to be investigated. In addition, it is not clear whether breakpoints or continuous transitions occur in the gas exchange data in upper-body exercise. Therefore, it should be further looked into if a breakpoint model compared to continuous linear or curvilinear (i.e. no-breakpoint) models better fit the gas exchange data of athletes with different disabilities in an upper-body exercise mode.

Aims

The overall aim of this thesis was to determine VO_{2peak}, the anaerobic threshold and efficiency in Paralympic sitting sports, with an emphasis on the influence of test mode and test protocol on these factors. This was done by means of four main studies (a systematic literature review and three experimental studies).

The aim of **Paper I** was to investigate between- and within-sports differences in VO_{2peak}. Furthermore, the influence of the test mode and other determinants (sex, age, body-mass, type of disability) on VO_{2peak} in Paralympic sitting sports was investigated.

> Hypotheses: VO_{2peak} was expected to be highest in athletes who compete in sports with continuously high movement demands and lowest in athletes with a tetraplegia. Furthermore, in wheelchair athletes, VO_{2peak} was expected to be higher in the sport-specific modes, i.e. in the wheelchair ergometry and wheelchair treadmill compared to the arm crank ergometry mode.

The aim of **Paper II** was to compare the VO_{2peak} and efficiency between upper-body poling and arm crank ergometry in trained able-bodied and paraplegic participants.

> Hypotheses: No differences in VO_{2peak} were expected between modes, but higher VO_{2peak} was expected in the able-bodied participants. Furthermore, higher efficiency was expected in the arm crank ergometry mode.

The aim of **Paper III** was to investigate the reliability and peak physiological responses of a 1-min and a 3-min closed-end as well as an incremental upper-body poling test.

> Hypotheses: High test-retest reliability was expected in the 1-min, 3-min and incremental test. Furthermore, no differences in VO_{2peak} were expected between the 3-min and the incremental test.

The aim of **Paper IV** was to compare outcome parameters identified with different gas exchange and blood lactate thresholds at the aerobic and the anaerobic threshold in endurance trained Para ice hockey players with different disabilities during upper-body poling.

Furthermore, it was investigated whether breakpoint or continuous no-breakpoint models fit the gas exchange data better.

> Hypotheses: It was hypothesized that outcome parameters identified with gas exchange and blood lactate threshold methods would differ at the aerobic threshold but not at the anaerobic threshold. Furthermore, continuous no-breakpoint models were expected to fit the gas exchange data better than breakpoint models.

METHODS

Overall design

To get an overview over VO_{2peak} values in all Paralympic sitting sports, a systematic literature review and meta-analysis was performed (Paper I). Furthermore, the influence of other factors, such as sex, age, body-mass, disability, test-mode on VO_{2peak} in wheelchair athletes was investigated with regression analyses. Sport-specificity is important for obtaining VO_{2peak}, % of VO_{2peak} used at the anaerobic threshold and efficiency that are reflective of the endurance capacity of the respective sport. In Para ice hockey players, sitting Para cross-country skiers and Para biathletes, the upper-body poling (UBP) movement may be the most sport-specific. Therefore, VO_{2peak} and exercise efficiency were compared between UBP and ACE, which is the most commonly used mode during upper-body exercise testing (Paper II). Furthermore, the test-retest reliability of VO_{2peak} was investigated during two closed-end and an incremental UBP test (Paper III). Performance at the anaerobic threshold is important both during competitions and in training. Therefore, it was investigated whether gas exchange and BLa threshold methods can be used interchangeably to identify the AT and the ANT in Para ice hockey players during upper-body poling. In addition, it was investigated whether breakpoint compared to continuous no-breakpoint models fit the gas exchange data better (Paper IV). Next to VO_{2peak}, the three experimental papers (Paper II to IV) included submaximal data, which allowed us to investigate efficiency across these papers (Figure 3).

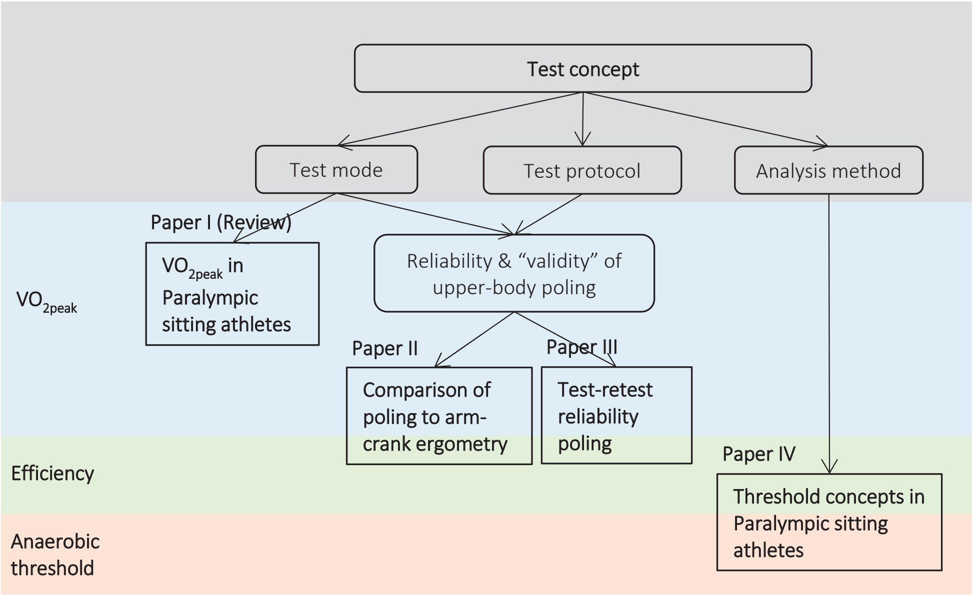


Figure 3. Overview over the four studies conducted in the course of this PhD.

Participants

In the meta-analyses, the data of 771 athletes was analysed in Paper I. In the experimental studies (Paper II to IV), 22 upper-body trained participants with a disability and 35 upper-body trained AB partook (Table 1). The seven participants with a disability in Paper II were PARA, upper-body trained and participated in various sports. The 15 participants in Paper IV were Para ice hockey players of the Norwegian national team and had all participated at national or international competitions or both. The Para ice hockey players used UBP both during training as well as regular testing. The AB participants were recreationally upper-body trained cross-country skiers (N = 34) and a rower used to cross-country skiing (N = 1), and as such all used to UBP. There was an overlap of one male participant with a disability for Paper II and IV, and of one male able-bodied participant for Paper II and III. The PARA of Paper II were recruited through collaboration with the St. Olavs University Hospital in Trondheim. The Para ice hockey players of Paper IV and the AB of Paper II and III were recruited through collaboration with the Olympic centre of Mid-Norway. The studies were approved by the Regional Ethics Committee for Medical and Health Research in Mid-Norway or the Norwegian Data Protection Authority and conducted in accordance with the Declaration of Helsinki of 1964 and its later amendments. All participants signed an informed consent form prior to testing, and were made aware that they could withdraw from the study at any point without providing an explanation.

Table 1. Anthropometrics of the participants of the four papers included in this thesis. Age and body mass are presented as mean ± SE for Paper I, and as mean ± SD for the three experimental studies. Training hours per week are presented as mean ± SD.

	Sex	Disability/ Able-bodied	Number of participants	Age (years)	Body mass (kg)	Training status (hours/week)
Paper I*	W	Disability	90	30.6 ± 0.8	65.7 ± 1.2	-
	M	Disability	669			
Paper II	W	Paraplegia	1	40	65	6
	W	Able-bodied	2	20.5 ± 0.7	64.3 ± 11.3	13.5 ± 3.5
	M	Paraplegia	6	31.8 ± 11.2	75.9 ± 12.9	5.6 ± 2.1
	M	Able-bodied	9	22.4 ± 2.6	78.1 ± 6.2	11.5 ± 3.2
Paper III	M	Able-bodied	24	28.3 ± 9.3	77.4 ± 8.9	9.8 ± 2.8
Paper IV	W	Disability	1	22	70	8.3
	M	Disability	14	27.4 ± 9.1	71.3 ± 8.3	11.8 ± 4.4

* The sex of 12 athletes included in parts of the analyses of Paper I was not specified.
Women (W), men (M)

Methods of Paper I

Paper I

PubMed, CINAHL, SPORTDiscus^{TM} and EMBASE were systematically searched in October 2016 using relevant medical subject headings, keywords and a Boolean search string. Study titles were first screened, after eligible abstracts and thereafter full-text articles were read. Studies that assessed VO_{2peak} values in sitting sports athletes with a physical disability in a laboratory setting were included. Data was extracted and analysed, and the quality of the included studies was assessed with a modified version of the Downs and Black checklist[98] by two independent reviewers. A meta-analysis was performed in Microsoft Excel (Version 2010, Microsoft Cooperation, The Microsoft Network, LLC, Richmond, USA) with the random-effects approach of DerSimonian and Laird[99] by pooling together studies that investigated athletes of the same sports discipline. A random-effects model was chosen to allow the true effect size to differ from study to study[100]. Participants of the studies included in Paper I in each sports discipline were heterogeneous with regards to age, sex, type of disability and training status, which will explain some of the variation in VO_{2peak} between studies. The alternative would have been to use a fixed effects model, where one assumes that the variation of effect sizes between studies can be explained by sampling error alone[100]. The quality of the studies included for each sports discipline was used to determine the level of evidence from unknown to strong[101]. A meta-regression was performed in Stata 14.2 (StataCorp LLC, Texas, USA) and a pooled-data multiple regression was performed in SPSS 22.0 (Software for Windows, SPSS Inc., Chicago, IL, USA) to investigate the effect of age, sex, body mass, disability and test mode on absolute and body-mass normalized VO_{2peak}.

Methods of Paper II – IV

Test set-up

All testing was conducted in a sitting position in different modified seat constructions in front of a Concept2 ski-ergometer (Concept2 Inc., Morrisville, Vermont, USA) in Paper II – IV (Figure 4). Furthermore, a modified ACE made from a road-bike (White, XXL Sport & Villmark AS, Norway) was employed in Paper II. In Paper II (Figure 4A and 4B) the upper-body was restricted, whereas in Paper III (Figure 4C) and in Paper IV (Figure 4D) the upper-body was unrestricted.

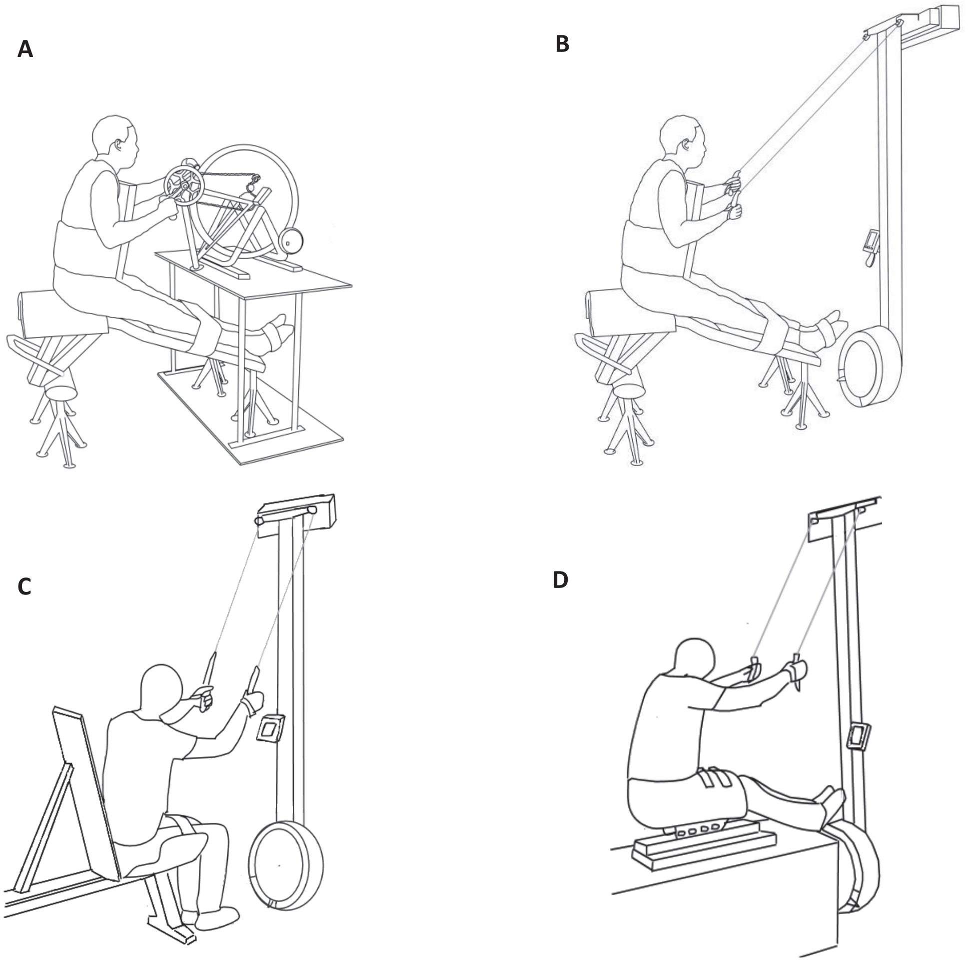


Figure 4. Test set-up of a modified weight-lifting bench with upper-body fixation with a chest-rest in front of a modified arm crank ergometer (A) and the Concept2 ski-ergometer (B) used in Paper II, a modified weight-lifting bench in front of the Concept2 ski-ergometer (C) used in Paper III, and of a Para ice hockey seat mounted on a platform in front of the Concept2 ski-ergometer (D) used in Paper IV.

Instruments and materials

In Paper II – IV, PO in the UBP was measured per stroke with the ski-ergometer's in-built software, which was previously validated with force and velocity measurements using a force cell (Noraxon USA inc., Scottsdal, AZ, USA) and the Oqus cameras of the Qualisys motion capture system (Qualisys AB, Gothenburg, Sweden)[46]. In Paper II, the modified ACE was attached to an electronical brake system for indoor cycling (CompuTrainer™, Racermate®, Inc, Seattle, USA), which was previously found to be valid compared to a Lode ergometer[102], and the in-built software (PerfPRO Studio©, Dynastream Innovations Inc., Canada) continuously recorded PO.

Respiratory parameters were continuously measured using an Oxycon Pro ergospirometer (Jaeger, Viasys BV, Bilthoven, the Netherlands) using a mixing chamber in Paper II and in breath-by-breath mode in Paper III, as well as using a Metamax II ergospirometer (CORTEX Biophysik GmbH, Leipzig, Germany) in breath-by-breath mode in Paper IV. The ergospirometers were calibrated against a known mixture of gases (5% CO_2, 15% O_2 in Paper II and III, and 4% CO_2, 16% O_2 in Paper IV) and a known airflow using an automatic calibration procedure in Paper II versus a manual calibration with a 3 L syringe in Paper III and IV. Both, measuring with a mixing chamber[103] and in breath-by-breath mode[104] under a VO_2 of 5 $L \cdot min^{-1}$ were found to be reliable and valid methods for assessing respiratory parameters with the Oxycon Pro ergospirometer. Measuring in breath-by-breath mode with the Oxycon Pro in Paper III was done to be able to additionally investigate VO_2 kinetics with high time resolution and to be able to investigate the effect of different data processing methods on VO_{2peak}. In comparison, the portable Metamax II ergospirometer was used during the data collection of Paper IV, which took place at the national Olympic centre in Oslo, where most of the Para hockey players have their training base. The Metamax was previously shown to slightly overestimate VO_2 by ~5% at high exercise intensities but showed acceptable reliability and validity at lower exercise intensities[105]. When testing in an upper-body mode, VO_2 even at higher exercise intensities is considerably lower than when testing in a lower-body or whole-body mode.

HR was continuously measured using a Polar M400 heart rate monitor and belt (Polar Electro Inc., Port Washington, NY, USA) in Paper II – IV. A 20 µL blood sample was taken directly after the submaximal stages and 1 and 3 minutes after the peak tests and BLa was analysed using a

Biosen C-Line Sport lactate measurement system (EKF-diagnostic GmbH, Magdeburg, Germany) for Paper II and III, and a 5 µL blood sample was taken at the same instances and BLa analysed with a Lactate Pro device (Arkray Inc., Japan) for Paper IV. The Biosen lactate measurement system was shown to be reliable and valid as compared to the commonly used Yellow Springs Instruments 2300 lactate analyser system[106]. The Lactate Pro was shown to be reliable and valid compared to the Yellow Springs Instruments 2300 lactate analyser system[107].

Test protocols

The test protocols for Paper II – IV were similar in that they contained four submaximal stages at a rating of perceived exertion (RPE) of 9, 11, 13 and 15 and an incremental test to exhaustion (Figure 5). Participants of Paper IV performed an extra submaximal stage at RPE 7. Participants of Paper III performed a 1-min and a 3-min all-out test in addition to the incremental test, with counter-balanced order of the three tests.

Participants of Paper II performed the four submaximal stages, incremental test and verification stage in ACE and UBP on two different test days in a counter-balanced design with a maximum of 4 days between test sessions. Participants of Paper III performed the same four submaximal stages and the three peak tests in the same individual random order a second time 4 ± 3 days later. On test day 2, participants of Paper IV performed seven to eight 5-min stages at increasing percentage of the peak PO of the incremental test of test day 1 (Figure 6). Strong verbal encouragement was given during all closed-end all-out tests and the incremental tests to exhaustion.

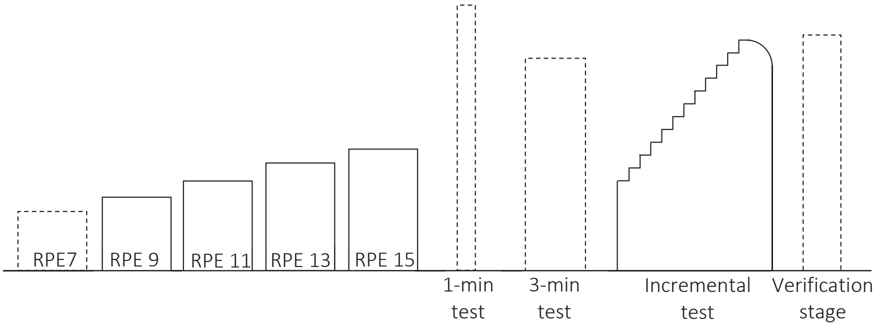


Figure 5. The test protocols for Paper II – IV contain four submaximal stages at RPE 9, 11, 13 and 15 and an incremental test to exhaustion (solid line). An extra submaximal stage at RPE 7 (Paper IV), the 1-min and the 3-min all-out test (Paper III), and the verification stage (Paper II, IV) were only part of one or two of the three papers (dotted lines).

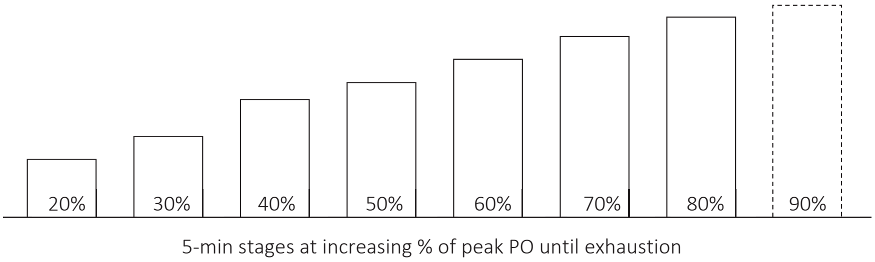


Figure 6. Seven to eight 5-min stages at increasing % of peak power output (PO) on the second test day of Paper IV.

Data processing and statistics

Data processing

Paper II – IV. The data processing of the three original papers was performed in Microsoft Excel (Version 2010, Microsoft Cooperation, The Microsoft Network, LLC, Richmond, USA) and MATLAB 8.1.0. (R2016a; Mathworks Inc., Natick, MA). For Paper II, MR was calculated from VO_2 and respiratory exchange ratio (RER) using a standard conversion table[108]. MR was then interpolated at a fixed PO of 40, 60 and 80 W to investigate gross efficiency and the entire PO-MR relationship was used to investigate exercise efficiency. For Paper III, the breath-by-breath data was interpolated at individually fitted sample frequencies and resampled at 1-s intervals. The single highest 30-s moving average of each test was then defined as the VO_{2peak} in accordance with Robergs et al.[109]. For Paper IV, two regression lines were fitted on the VO_2-VCO_2[29] and VE/VO_2 data[30] to determine the VT, and on the VCO_2-VE[29] and VE/VCO_2[30] data to determine the RCT. The LT1 was determined by the first rise in BLa of 0.4 $mmol \cdot L^{-1}$ and 1.0 $mmol \cdot L^{1}$[31, 32] and a breakpoint in the log-log transformed VO_2-BLa data[33], and the LT2 by a fixed BLa concentration of 4 $mmol \cdot L^{-1}$[36] and by employing the modified D_{max} method[31, 35]. PO, VO_2, HR, BLa and RPE were interpolated at each threshold. For comparing efficiency in Paper II – IV, MR was calculated from VO_2 and RER[108], and the PO-MR relationship was used to investigate exercise efficiency in each exercise mode and group.

Statistics

The statistics of Paper II to IV were performed in Microsoft Excel (Version 2010, Microsoft Cooperation, The Microsoft Network, LLC, Richmond, USA) and SPSS 22.0 (Software for Windows, SPSS Inc., Chicago, IL, USA). For all statistical tests an alpha level of 0.05 was used to indicate statistical significance.

Paper II. A linear mixed model with fixed coefficients and random intercept was employed to investigate the effect of exercise mode, group and intensity on PO, physiological and perceptual parameters during the submaximal stages and the incremental test. The same model was used to investigate the effect of exercise mode and group on exercise efficiency. The linear mixed model investigates the effect of each one of three factors (exercise mode, group and intensity), while adjusting for the effect of the two other factors. There is no consensus on whether to present the actual means or use the estimated marginal means[110]. Since p values are calculated on the basis of the estimated marginal means, to present the estimated marginal means in the results section was considered more consistent.

Paper III. Paired-samples T-tests were used to compare PO, the physiological and perceptual outcome parameters between the 1-min and 3-min closed-end and the open-end incremental test to exhaustion. *Absolute reliability.* Absolute reliability was assessed by the standard error of measurement (SEM) and the smallest detectable change (SDC). The SEM was calculated as $SD_{diff}/\sqrt{2}$[111], and the 80% and 95% SDC as $SEM \cdot 1.28 \cdot \sqrt{2}$ and $SEM \cdot 1.96 \cdot \sqrt{2}$, respectively[112]. *Relative reliability.* Intraclass correlation coefficients ($ICC_{1,1}$, $ICC_{2,1}$, and $ICC_{3,1}$) were calculated as a measure of relative reliability[113]. In addition to using the 80% SDC as a measure of absolute reliability and the $ICC_{2,1}$ as a measure of relative reliability in Paper III, 95% SDC and the $ICC_{1,1}$ and $ICC_{3,1}$ were calculated in this thesis. Whereas the $ICC_{2,1}$ allows the error to be partitioned between systematic and random error, systematic and random error are treated together in the calculations of the $ICC_{1,1}$ and variance associated with systematic error is not included in the $ICC_{3,1}$[113]. Accordingly, the $ICC_{1,1}$ and $ICC_{3,1}$ and $ICC_{2,1}$ are similar if the systematic error is small[113]. Ranges of 0.26-0.49, 0.50-0.69, 0.70-0.89 and 0.90-1.0 were classified as low, moderate, high and very high ICC according to Munro's criteria [114].

Paper IV. In the submitted version of Paper IV and the main results presented in this thesis, paired-samples T-tests were used to compare outcome parameters identified with the different methods within the VT, RCT, LT1, and LT2, and between all four different thresholds. Furthermore, in this thesis, a repeated-measures ANOVA was conducted followed by post hoc test with a Bonferroni correction to compare the % of VO_{2peak} identified with the different threshold methods. Pearson's r was used to investigate whether the different methods used within the VT, RCT, LT1, and LT2, and between all four different thresholds correlate. To investigate whether a breakpoint model or continuous linear and curvilinear (i.e. no-breakpoint) models fit the gas exchange data better, two regression lines (equation 1) versus a single linear regression line (equation 2), an exponential curve (equation 3), and a 3rd order polynomial curve (equation 4) were fitted on the data by linear least squares fitting.

$$y = \begin{cases} a_1 + b_1\, x, & t < k \\ a_2 + b_2\, x, & t \geq k \end{cases} \tag{1}$$

$$y = a + b\, x \tag{2}$$

$$y = a + c \cdot exp\left(\frac{x+g}{d}\right) \tag{3}$$

$$y = a + b_1\, x + b_2\, x^2 + b_3\, x^3 \tag{4}$$

y is the variable of interest, *a* the offset on the y-axis, *b* the slope coefficients, *c* and *d* the spreading coefficients, *g* the offset on the x-axis and *k* the point where the first and the second regression line of the piecewise function cross.

To compare the fit of the four models, the Akaike information criterion (AIC) (equation 5)[115] and the Akaike weights (w_i) (equation 7) for each model *i* relative to the set of *R* candidate models were calculated based on the delta AIC (Δ_i) (equation 6)[116, 117].

$$AIC = n \cdot \log\left(\frac{SS_{er}}{n}\right) + 2 \cdot K \tag{5}$$

$$Delta\ AIC = \Delta_i = AIC_i - AIC_{2reg} \tag{6}$$

$$AIC\ weight = w_i = \frac{exp^{(-\frac{\Delta i}{2})}}{\sum_{r=1}^{R} exp^{(-\frac{\Delta r}{2})}} \tag{7}$$

n is the number of data points, SS_{er} the error sums of squares, and *K* the number of parameters +1 of each model. Our rationale was that a better fit of the two regression lines (breakpoint model) compared to the single regression line, exponential or 3rd order polynomial curve (continuous no-breakpoint models) would indicate breakpoint presence.

RESULTS

Paper I – VO_{2peak} in Paralympic sitting sports disciplines

Included in this systematic literature review were 57 studies that provided absolute and body-mass normalized VO_{2peak} values in 771 athletes in 14 different sitting sports. Mean absolute and body-mass normalized VO_{2peak} ± standard error (SE) of the sports disciplines ranged from 2.9 ± 0.3 $L{\cdot}min^{-1}$ and 45.6 ± 5.1 $ml{\cdot}kg^{-1}{\cdot}min^{-1}$ in Nordic sit skiing to 1.4 ± 0.2-$L{\cdot}min^{-1}$ and 17.3 ± 3.5 $ml{\cdot}kg^{-1}{\cdot}min^{-1}$ in shooting and 1.3 ± 0.1 and 18.9 ± 1.6 in wheelchair rugby (Figure 7). Within-sports variations in absolute and body-mass normalized VO_{2peak} values, based on CI ranges, were relatively small in wheelchair basketball (0.4 $L{\cdot}min^{-1}$ and 7.2 $mL{\cdot}kg^{-1}{\cdot}min^{-1}$), wheelchair racing (0.6 $L{\cdot}min^{-1}$ and 7.4 $mL{\cdot}kg^{-1}{\cdot}min^{-1}$) and wheelchair rugby (0.4 $L{\cdot}min^{-1}$ and 6.1 $mL{\cdot}kg^{-1}{\cdot}min^{-1}$), but above 0.6 $L{\cdot}min^{-1}$ and 7.5 $mL{\cdot}kg^{-1}{\cdot}min^{-1}$ for the remaining sport disciplines.

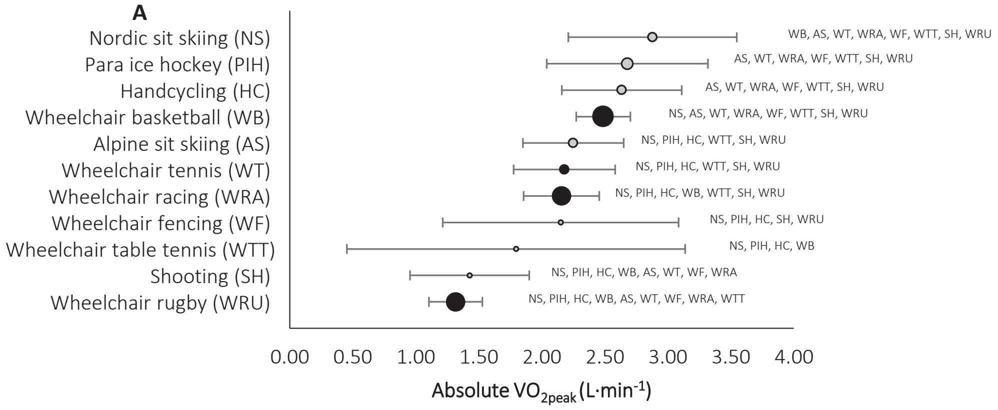


Figure 7. Overview of absolute (A) and body-mass normalized (B) peak oxygen uptake (VO_{2peak}) (mean ± 95% CI based on SE) within each of the sitting sports disciplines. Sports disciplines are presented in order of absolute VO_{2peak} values, from high to low. Sample size is indicated by the size of the dots. Black dots are sports disciplines with a strong level of evidence whereas grey dots denote sports disciplines with a limited or moderate level of evidence. Labels in superscript indicate significant differences to the respective sports discipline.
Note: several factors such as sex, age, body mass, disability, training status and test modes are grouped together in this overview table. Data of athletes with tetraplegia was excluded from the calculations of all sports discipline means except for wheelchair rugby.

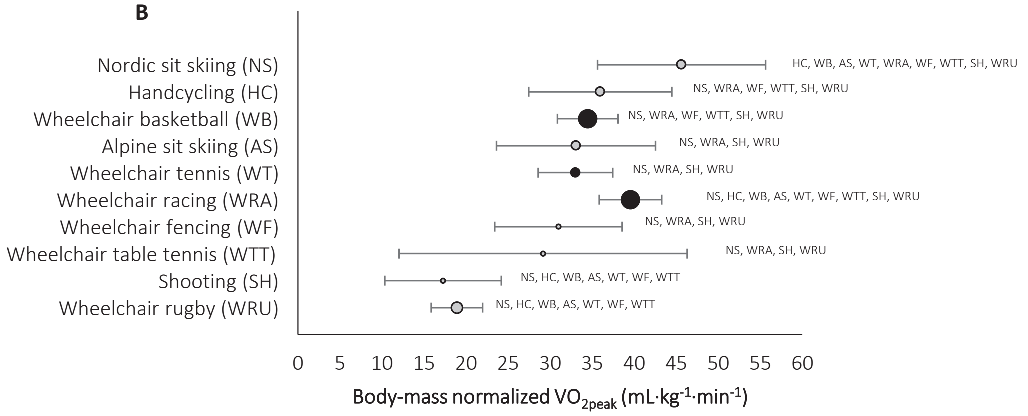


Figure 7 (continued). Please see figure legend of Figure 7.

The multiple regression analyses, based on 22 studies which provided individual data of 169 athletes in 4 different sports disciplines (wheelchair basketball, wheelchair racing, wheelchair tennis and wheelchair rugby), resulted in the following two equations as the best predictions of absolute (I) and body-mass normalized (II) VO_{2peak} values.

$$\text{Absolute } VO_{2peak} = 1.22 + \text{body mass}_i \cdot 0.02\ [0.25] + \text{female}_i \cdot -0.62\ [-0.25] + \text{TETRA}_i \cdot -1.09\ [-0.63] + \text{AMP}_i \cdot 0.29\ [0.10] + \text{WERG}_i \cdot 0.36\ [0.24] + \text{WTR}_i \cdot 0.32\ [0.20]$$
$$(F_{6,162} = 52.52,\ p = 0.00) \qquad \text{(I)}$$

$$\text{Body-mass normalized } VO_{2peak} = 49.11 + \text{body mass}_i \cdot -0.24\ [-0.26] + \text{female}_i \cdot -9.79\ [-0.24] + \text{TETRA}_i \cdot -16.52\ [-0.62] + \text{AMP}_i \cdot 5.38\ [0.12] + \text{WERG}_i \cdot 5.71\ [0.27] + \text{WTR}_i \cdot 4.54\ [0.18]$$
$$(F_{6,162} = 52.50,\ p = 0.00) \qquad \text{(II)}$$

Paper II – Comparison of VO_{2peak} and exercise efficiency between upper-body poling and arm crank ergometry

VO_{2peak} did not significantly differ between restricted UBP and restricted ACE (both groups pooled: 35.9 ± 7.8 vs 37.3 ± 7.0 mL·kg^{-1}·min^{-1}, $p = 0.112$), although peak PO was 19% lower in restricted UBP (both groups pooled: 118 ± 34 vs 145 ± 33 watt, $p < 0.001$). MR was ~24% higher in restricted UBP compared to restricted ACE ($p < 0.001$), i.e. exercise efficiency was lower in UBP (Figure 8A). In line with this, gross efficiency was lower in UBP (10.4 ± 0.9, 11.4 ± 0.8 and 12.0 ± 0.9%) compared to ACE (12.9 ± 1.8, 14.0 ± 1.8 and 14.7 ± 1.9%) at 40, 60 and 80 watt. PARA had 22% lower VO_{2peak} compared to AB (both exercise modes pooled: 31.5 ± 6.4 vs 39.7 ± 6.6 mL·kg^{-1}·min^{-1}, $p = 0.007$). However, there were no significant differences in exercise efficiency and gross efficiency between PARA and AB ($p = 0.323$ and $p > 0.489$) (Figure 8B, 8C).

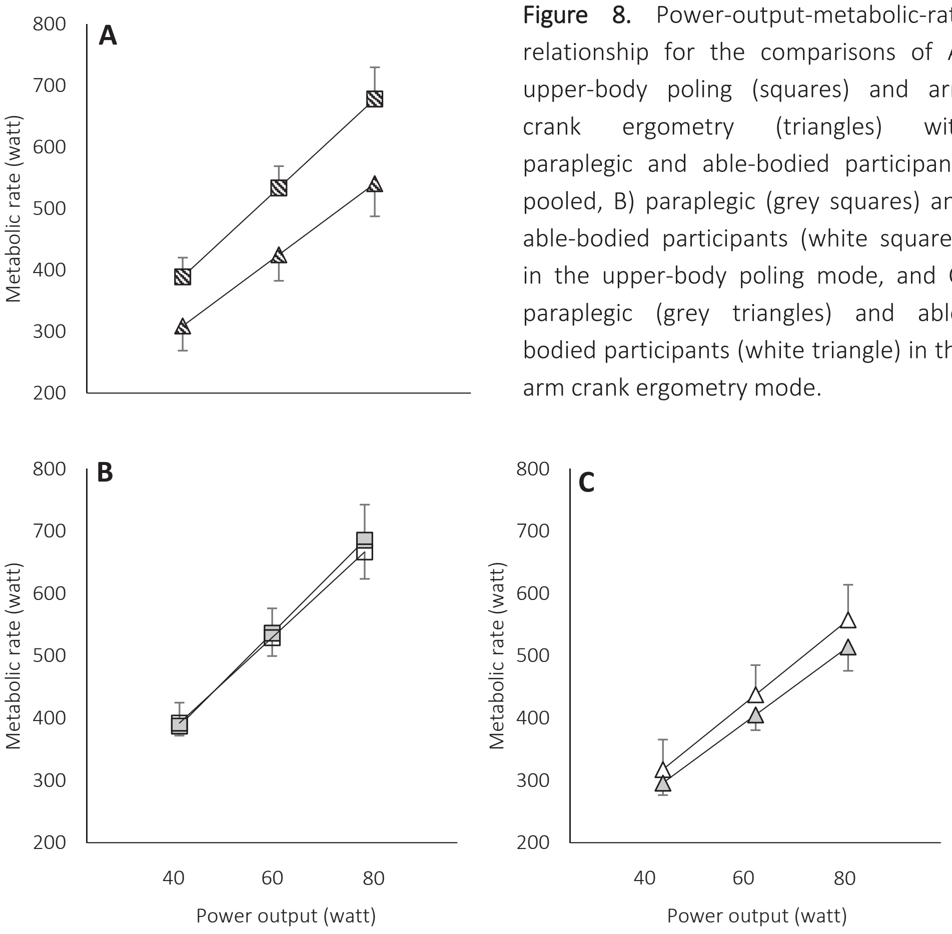


Figure 8. Power-output-metabolic-rate relationship for the comparisons of A) upper-body poling (squares) and arm crank ergometry (triangles) with paraplegic and able-bodied participants pooled, B) paraplegic (grey squares) and able-bodied participants (white squares) in the upper-body poling mode, and C) paraplegic (grey triangles) and able-bodied participants (white triangle) in the arm crank ergometry mode.

Paper III – Test-retest reliability of physiological parameters during upper-body poling

Comparison of the 1-min, the 3-min and the incremental test

Based on the average values of test day 1 and 2, the incremental (196 ± 28 watt, 45.4 ± 5.5 $mL \cdot kg^{-1} \cdot min^{-1}$, 169 ± 12 $beats \cdot min^{-1}$, 11.7 ± 2.3 $mmol \cdot L^{-1}$) and the 3-min test (201 ± 36 watt, 44.5 ± 5.5 $mL \cdot kg^{-1} \cdot min^{-1}$, 169 ± 12 $beats \cdot min^{-1}$, 11.7 ± 2.0 $mmol \cdot L^{-1}$) resulted in significantly lower peak PO, higher VO_{2peak}, higher peak HR and higher peak BLa compared to the 1-min test (256 ± 47 watt, 40.4 ± 5.0 $mL \cdot kg^{-1} \cdot min^{-1}$, 166 ± 12 $beats \cdot min^{-1}$, 10.9 ± 2.2 $mmol \cdot L^{-1}$) (all $p < 0.001$) (Figure 9). Additionally, the incremental test resulted in significantly higher VO_{2peak} compared to the 3-min test ($p = 0.03$).

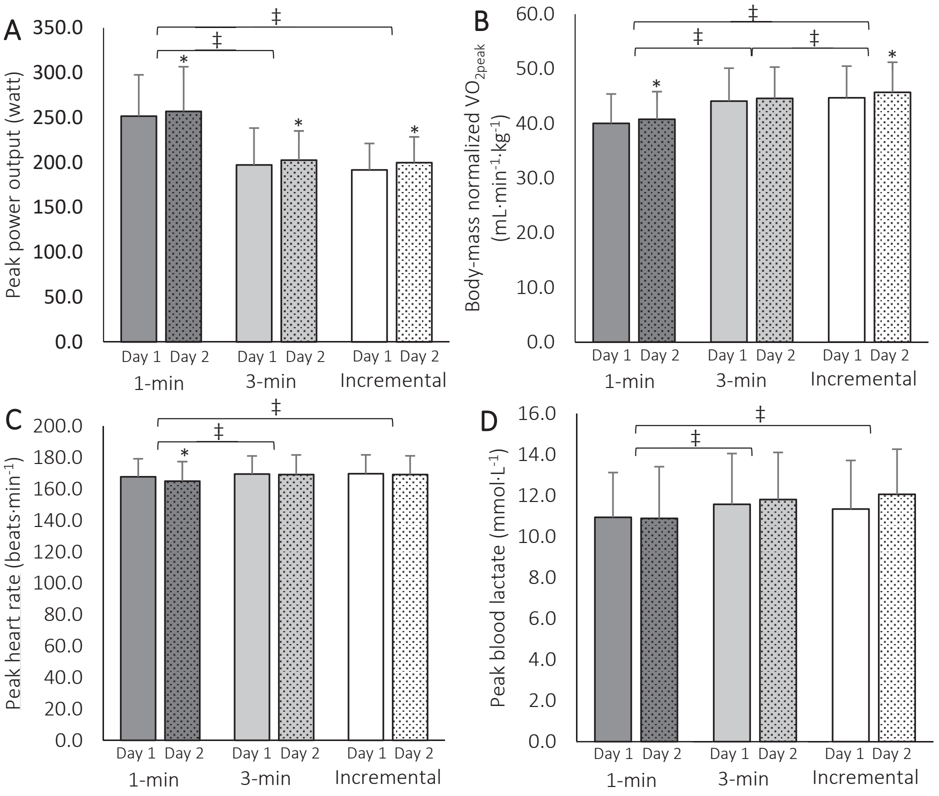


Figure 9. Peak power output (A), peak physiological parameters (C – D) of test day 1 and 2 in 22 cross-country skiers performing a 1-min and 3-min closed-end and an incremental upper-body poling test to exhaustion. Blank bars demark test day 1 and dotted bars test day 2.
** Significant differences between test day 1 and 2 at an alpha level of 0.05*
‡ Significant differences between the 1-min, 3min and incremental test at an alpha level of 0.05

Absolute and relative reliability

Irrespective of the type of ICC used, high and very high ICCs across all physiological outcome parameters and peak PO were found for the 1-min, the 3-min and the incremental test (Table 2). In all three tests, the 80% SDC was consistently small for HR_{peak} (1-min: 4%, 3-min: 4%, incremental: 3%), moderate for body-mass normalized VO_{2peak} (1-min: 5%, 3-min: 6%, incremental: 7%) as well as peak PO (1-min: 8%, 3-min: 9%, incremental: 6%) and large for peak BLa (1-min: 20%, 3-min: 12%, incremental: 22%). Higher differences between tests would have to be present to accept a true difference if the 95% SDC is employed.

Table 2. Interclass correlation coefficients ($ICC_{1,1}$, $ICC_{2,1}$, $ICC_{3,1}$) and 80% and 95% smallest detectable change (SDC) of power output and physiological parameters for a 1-min, a 3-min and an incremental upper-body poling test in able-bodied, upper-body trained participants. Calculations are based on data from 22 participants for the 1-min and the incremental test and 24 participants for the 3-min test.

	1-min					3-min					Incremental				
	$ICC_{1,1}$	$ICC_{2,1}$	$ICC_{3,1}$	80% SDC (%)	95% SDC (%)	$ICC_{1,1}$	$ICC_{2,1}$	$ICC_{3,1}$	80% SDC (%)	95% SDC (%)	$ICC_{1,1}$	$ICC_{2,1}$	$ICC_{3,1}$	80% SDC (%)	95% SDC (%)
Peak PO (watt)	0.937	0.946	0.941	7.6	11.7	0.887	0.923	0.891	9.4	14.4	0.920	0.955	0.954	5.7	8.7
VO_{2peak} ($mL\cdot kg^{-1}\cdot min^{-1}$)	0.942	0.956	0.954	4.9	7.5	0.925	0.942	0.925	6.2	9.5	0.901	0.933	0.910	6.7	10.3
Peak HR ($beats\cdot min^{-1}$)	0.872	0.926	0.898	4.1	6.2	0.914	0.935	0.911	4.1	6.3	0.955	0.956	0.954	3.1	4.8
Peak BLa ($mmol\cdot L^{-1}$)	0.738	0.827	0.728	19.9	30.5	0.878	0.916	0.876	12.4	19.1	0.573	0.728	0.588	22.2	34.1

Power output (PO), peak oxygen uptake (VO_{2peak}), heart rate (HR), blood lactate (BLa)

Paper IV – Gas exchange and blood lactate thresholds in Paralympic athletes during upper-body poling

At the aerobic threshold, all outcome parameters identified with breakpoints in the VO_2-VCO_2 or VE/VO_2-time relationship at the VT were higher than the ones identified with a fixed increase in BLa of 0.4 or 1.0 mmol·L^{-1} at LT1 (all $p < 0.001$). At the anaerobic threshold, the outcome parameters identified with breakpoints in the VCO_2-VE relationship at the RCT and with the modified D_{max} method at the LT2 were higher compared to parameters identified with the fixed BLa value of 4 mmol·L^{-1} at the LT2 (all $p < 0.03$) (Figure 10). When using Bonferroni corrections, there were no significant differences between the VO_2-VCO_2, VE/VO_2, and log-log transformed VO_2-BLa method used to determine the aerobic threshold and the VE/VCO_2, VCO_2-VE and D_{max} method used to determine the anaerobic threshold (all $p > 0.08$) (Figure 10).

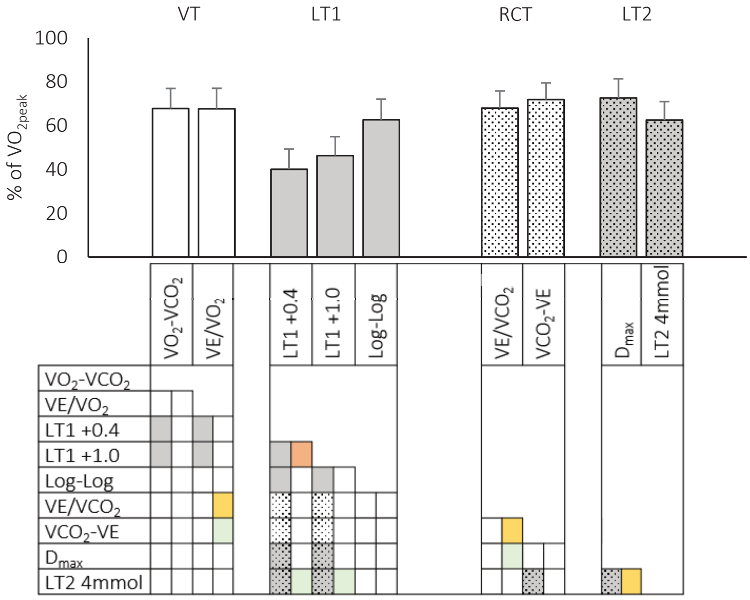


Figure 10. Percentage of peak oxygen uptake (% of VO_{2peak}) identified with different methods at the ventilatory threshold, the lactate threshold 1, the respiratory compensation threshold and the lactate threshold 2. The first column for each method indicates significant differences to the other methods at an alpha level of 0.05 (indicated by pattern-filled squares) after using Bonferroni corrections. The second column for each method indicates significant correlations with the other methods at an alpha level of 0.05 (green: moderate correlation, yellow: high correlation, orange: very high correlation).

Ventilatory threshold (VT), first lactate threshold (LT1), respiratory compensation threshold (RCT), second lactate threshold (LT2)

All methods used to determine thresholds in the gas exchange data are based on identifying a breakpoint. To investigate whether clear breakpoints exist in the gas exchange data, the fit of a two regression line (i.e. breakpoint) model was compared to continuous linear and curvilinear (i.e. no-breakpoint) models. The determination of the VT and RCT with breakpoint models was possible with good fit in all 15 participants (all $R^2 > 0.931$). However, the 3rd order polynomial (continuous no-breakpoint model) fit the data of 10 athletes better than the breakpoint model for both the VO_2-VCO_2 (VT) (Figure 11) as well as the VCO_2-VE relationship (RCT) (Figure 12). This corresponds with 71 and 69% probability that continuous no-breakpoint models fit the VO_2-VCO_2 and VCO_2-VE relationship better compared to the breakpoint model. In comparison, the breakpoint model fit the VE/VO_2 (VT) and the VE/VCO_2 data (RCT) better in 6 and 7 participants compared to continuous no-breakpoint models. This corresponds with 41 and 47% probability for the breakpoint model to fit the data better than continuous no-breakpoint models.

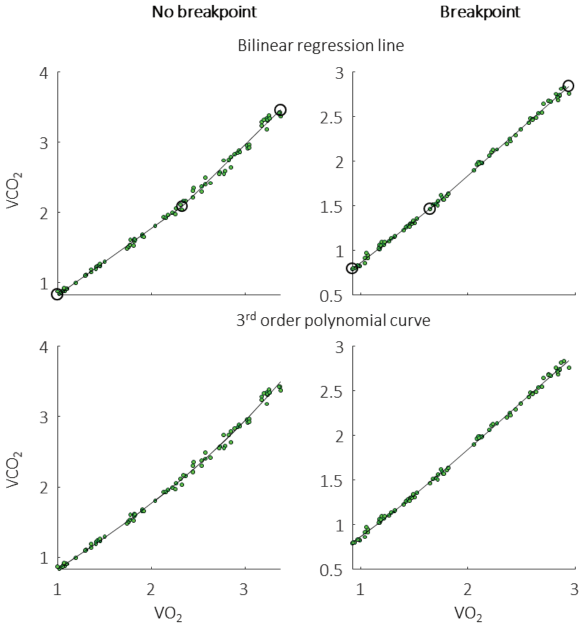


Figure 11. Exemplary VO$_2$-VCO$_2$ plots fitted with a bilinear regression line and a 3rd order polynomial curve for an athlete without breakpoint (plots to the left) and with suggested breakpoint presence (plots to the right). (Note that the plots of all five athletes with a suggested breakpoint also show a rather linear increase in the VO$_2$-VCO$_2$ relationship.)

Oxygen uptake (VO$_2$), carbon dioxide production (VCO$_2$)

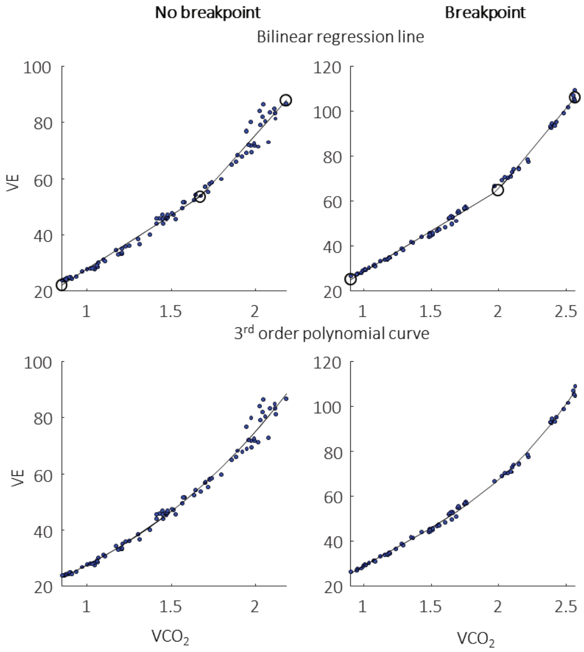


Figure 12. Exemplary VCO$_2$-VE plots fitted with a bilinear regression line, and a third order polynomial curve for an athlete without breakpoint (plots to the left) and with suggested breakpoint presence (plots to the right). (Note that the plots of all five athletes with a suggested breakpoint show a rather curvilinear increase in the VCO$_2$-VE relationship.)

Carbon dioxide production (VCO$_2$), minute ventilation (VE)

Overview of Paper II – IV: VO$_{2peak}$ and exercise efficiency in upper-body poling and arm crank ergometry

Overview of Paper II – IV: Peak responses during the incremental tests to exhaustion

Figure 13 provides an overview over peak PO and VO$_{2peak}$ during the incremental tests performed in Paper II – IV. Peak PO and VO$_{2peak}$ were generally higher in the unrestricted UBP mode (Paper III and IV) compared to the restricted UBP and restricted ACE mode (Paper II). Peak PO and VO$_{2peak}$ were generally lower in participants with a disability (blank bars) compared to AB (dotted bars). This is confirmed by the significantly lower VO$_{2peak}$ in PARA compared to AB Paper II (31.5 ± 6.4 vs 39.7 ± 6.6 mL·kg^{-1}·min^{-1}, respectively, $p < 0.007$). Statistics were only performed to compare UBP and ACE in PARA and AB within Paper II, and not across papers. This since the participants and the test set-up were not the same across studies.

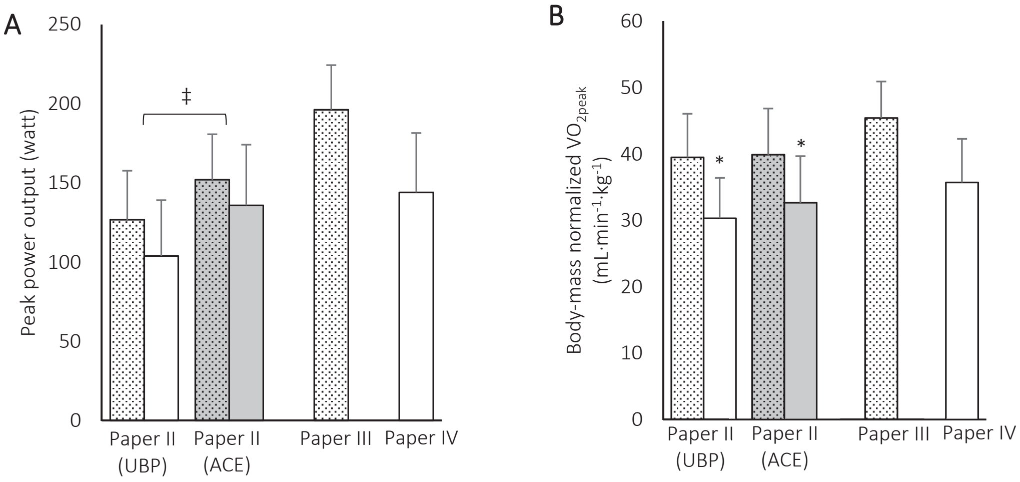


Figure 13. Peak power output (A) and body-mass normalized VO$_{2peak}$ (B) during restricted upper-body poling and arm crank ergometry in able-bodied cross-country skiers and paraplegic athletes (Paper II), during unrestricted upper-poling in able-bodied cross-country skiers (Paper III), and during unrestricted upper-body poling in Para ice hockey players (Paper IV).
Dotted bars demark data from able-bodied participants. White bars are data from upper-body poling, whereas grey bars are from arm crank ergometry.
** Significant differences between able-bodied and paraplegic participants at an alpha level of 0.05*
‡ Significant differences between upper-body poling and arm crank ergometry at an alpha level of 0.05

Overview of Paper II – IV: Exercise efficiency

The actual PO-MR relationships of restricted and unrestricted UBP and restricted ACE in PARA and AB are illustrated in Figure 14 and the recalculated PO-MR relationships at fixed POs are illustrated in Figure 15. At fixed POs, MR was 19% higher in UBP compared to ACE, indicating lower exercise efficiency in UBP ($p < 0.001$) (Figure 15). In addition, in UBP, there was an interaction between restricted UBP and unrestricted UBP ($p = 0.048$).

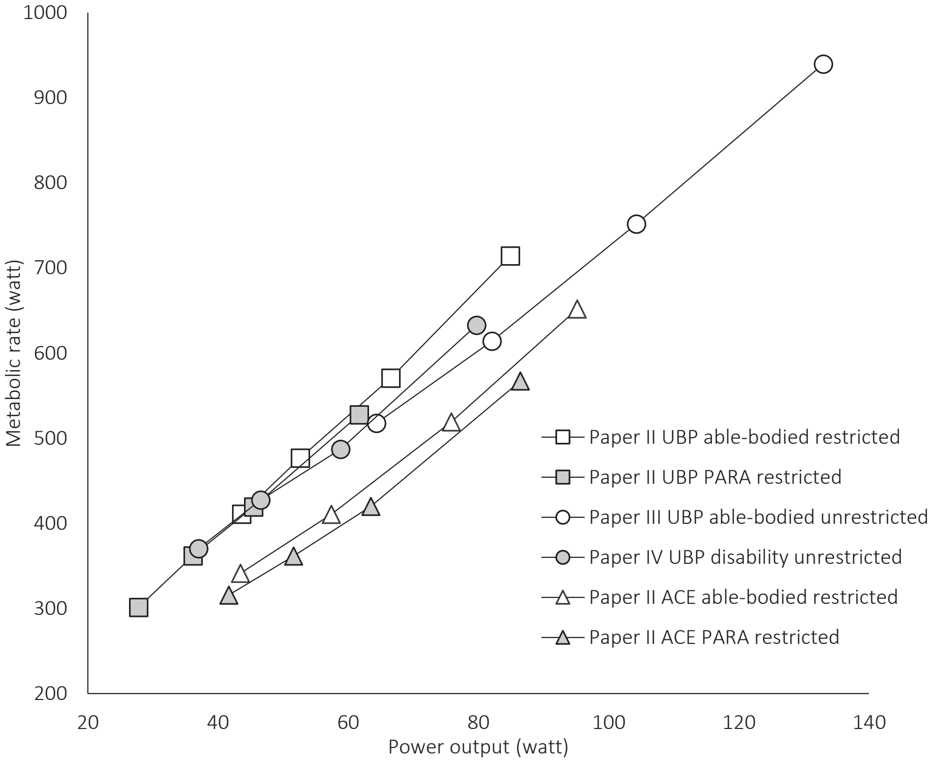


Figure 14. Power-output–Metabolic-rate relationship during restricted upper-body poling and arm crank ergometry in able-bodied cross-country skiers and athletes with a disability, during unrestricted upper-poling in able-bodied cross-country skiers, and during unrestricted upper-body poling in Para ice hockey players.

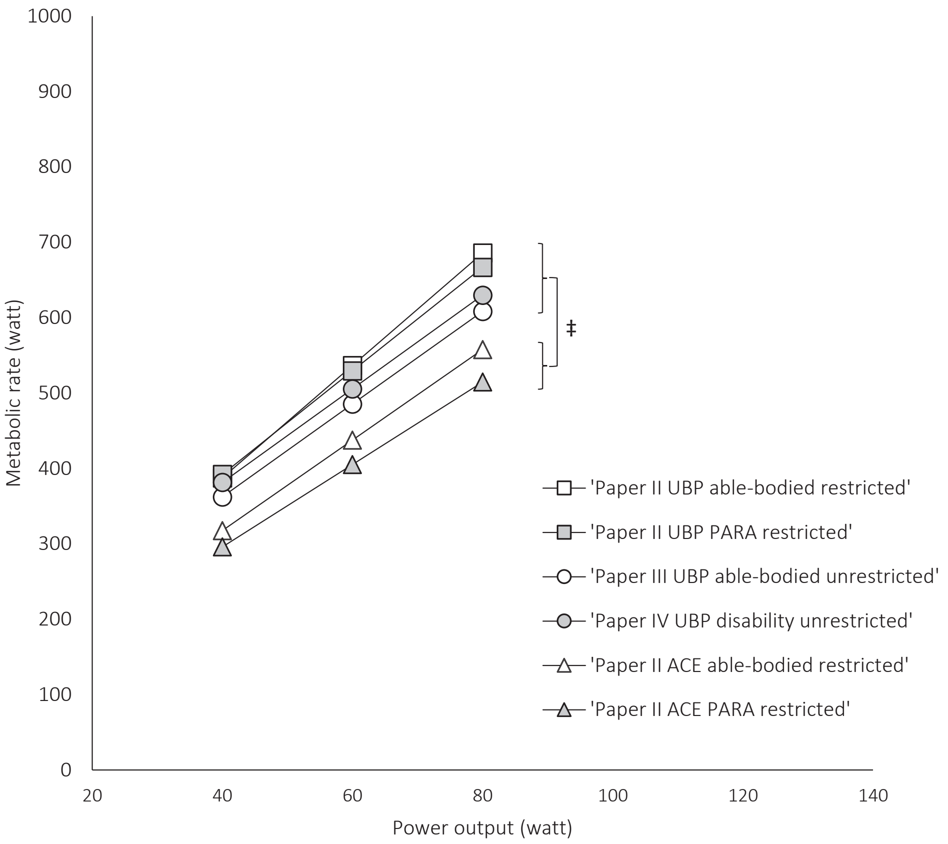


Figure 15. Power-output–Metabolic-rate relationship at fixed power outputs of 40, 60 and 80 watt during restricted upper-body poling and arm crank ergometry in able-bodied cross-country skiers and athletes with a disability, during unrestricted upper-poling in able-bodied cross-country skiers, and during unrestricted upper-body poling in Para ice hockey players.
‡ significant differences between upper-body poling and arm crank ergometry at an alpha level of 0.05

DISCUSSION

The main findings of the four papers included in this PhD were that: 1) VO_{2peak} values between Paralympic sitting sports were fairly well reflected by the sport-specific demands and, therefore, highest in sports with continuously high physical efforts such as Nordic sit skiing. In wheelchair athletes, being male, not being TETRA or having an amputation and testing in a WERG or WTR mode was favourable for high VO_{2peak} (Paper I). VO_{2peak} was not different between UBP and ACE but higher in AB compared to PARA. However, exercise efficiency was lower in UBP compared to ACE, but not different between AB and PARA (Paper II). In UBP, the relative test-retest reliability of VO_{2peak} during different closed-end and an open-end incremental upper-body poling test was high. However, VO_{2peak} was significantly higher during a 3-min closed-end test and an open-end incremental test compared to a 1-min closed-end test (Paper III). Furthermore, the outcome parameters identified with breakpoint methods at the VT and with fixed methods at the LT1, used to determine the AT, did significantly differ. The RCT and LT2 used to determine the ANT, and the VT were closely located. In addition, continuous linear and curvilinear (i.e. no-breakpoint) models fitted most of the gas exchange data obtained from Paralympic athletes during UBP better than breakpoint models (Paper IV).

VO_{2peak} during upper-body exercise

Influence of sports discipline. In line with the hypothesis of Paper I, VO_{2peak} was highest in Nordic sit skiing (Paper I). High absolute VO_{2peak} values were also found in Para ice hockey players and high body-mass normalized VO_{2peak} values in wheelchair racing. This might in part be explained by the sport-specific demands in these three sports. Nordic sit skiing and wheelchair-racing athletes perform at continuously high physical efforts[5, 6]. In Para ice hockey, athletes require maintenance of sprint ability during games that last three times 15 min, which was highly correlated with VO_{2peak}[89]. In contrast, shooting athletes display low VO_{2peak} values, which might be related to their low levels of displacement. However, the representativeness of the low VO_{2peak} values for the wheelchair shooting athlete population remains reduced due to a limited number of studies and participating athletes included for this sports discipline in the meta-analysis. VO_{2peak} values were also found to be low in wheelchair rugby athletes, which might be surprising given that long distances need to be covered during games that last approximately 70 minutes[118]. However, only athletes with TETRA are eligible to wheelchair rugby, who were found to have low VO_{2peak} values due to SCI-specific limitations[119, 120].

Influence of type of disability. The finding of low VO_{2peak} values in TETRA is in line with the hypothesis and the results of the regression analyses conducted in Paper I, which show that compared to PARA, TETRA athletes have lower and AMP higher absolute and body-mass normalized VO_{2peak} values. Both TETRA and PARA lack sympathetic control to the paralyzed lower limbs, and depending on the level of injury to the paralyzed trunk, which limits the amount of active muscle mass during testing[25]. Furthermore, blood redistribution is impaired below the level of injury, including a lack of innervation to the splanchnic vascular bed in individuals with complete SCI injuries above the Th6 level[24]. The splanchnic vascular bed is the body's largest blood reservoir and lacking innervation of this area further negatively influences blood redistribution to the active muscles[24]. Moreover, TETRA athletes with autonomic completeness above Th1 lack innervation to the heart, which reduces peak HR and consequently leads to lower cardiac output and lower VO_{2peak}[67]. In a study by West et al.[120], athletes competing in wheelchair rugby on an international level were tested on autonomic completeness. All were found to have incomplete autonomic injuries, which may indicate that high performance on an international level in wheelchair rugby is only possible when part of the autonomic system is still intact. The TETRA athletes included in wheelchair rugby in the systematic literature review (Paper I) might, therefore, mostly be athletes with an incomplete autonomic injury and have higher VO_{2peak} values than would be expected in athletes with complete autonomic injuries.

There were no significant differences in VO_{2peak} between PARA and athletes with spina bifida (Paper I), suggesting that, irrespective of the different test modes and protocols employed in the included studies, non-traumatic versus traumatic spinal cord injuries lead to similar limitations in the cardiorespiratory system. Even though this has not yet been looked into in detail, several studies that investigated VO_{2peak} have pooled participants with spina bifida and SCI into one group before[9, 121-127] and the absence of this being discussed in further detail leads us to believe that similar responses were seen on an individual level. Furthermore, and in line with the higher peak PO during an incremental test in sitting athletes with a single leg amputation reported by Hutzler et al. [68], VO_{2peak} values in Paper I were significantly higher in

AMP compared to PARA. This is in contrast to Coutts et al.[69], who found similar absolute VO_{2peak} in AMP and PARA. However, the AMP included in the latter study were all double above knee amputees, and likely had lower body mass compared to PARA. In line with this, low body mass was shown to be related to low absolute VO_{2peak}[128]. In comparison, the AMP included in Paper I were mostly single leg amputees and body mass did not significantly differ between AMP and PARA. The 5.3 $mL \cdot kg^{-1} \cdot min^{-1}$ higher body-mass normalized VO_{2peak} in AMP compared to PARA in Paper I was lower than the 8.2 $mL \cdot kg^{-1} \cdot min^{-1}$ difference between AB and PARA found in Paper II, and lower than the 9.4 - 17 $mL \cdot kg^{-1} \cdot min^{-1}$ differences in a range of studies that tested AB and PARA in the same upper-body mode and protocol[71, 129, 130]. Even though this needs to be investigated in future studies, it is speculated that VO_{2peak} during upper-body exercise is lower in AMP compared to AB, which might be related to AMP not being able to use both legs for stabilization and therewith recruit less active muscle mass during incremental testing. Overall, future studies are needed to look into how other common disabilities of Paralympic sitting athletes, such as an amputation or spina bifida, influence VO_{2peak}.

Influence of test mode. In addition to sport-specific demands and disability-related limitations, the test mode is important when assessing VO_{2peak}. In wheelchair athletes, the WTR and WERG mode resulted in higher VO_{2peak} values compared to ACE in the regression analyses of the systematic literature review (Paper I). The clear differences between WTR and WERG compared to ACE might be related to that only data of wheelchair athletes was included in these analyses for whom the wheelchair test modes may be more sports-specific than ACE. However, the results of these regression analyses need to be interpreted with caution since modes were not compared within the same studies in the same participants. Therefore, another meta-analysis was conducted (Supplementary data 1, Appendix 1). In the latter analysis, the difference in VO_{2peak} was investigated between test modes in athletes and non-athlete participants both with and without a disability, who were tested in at least two of the modes within the same study. In the latter analysis, no differences in VO_{2peak} were found between ACE and WERG. The reasons for this need to be further investigated but might be due to less of an effect of sport-specificity of the test mode when also non-athlete participants with a disability and AB participants are included. Furthermore, ACE and WERG might be similar in that trunk oscillations and shifts in centre of gravity do not contribute to power production as much as during WTR[131].

In line with no difference in VO_{2peak} in ACE and WERG, no difference was found in VO_{2peak} between UBP and ACE, despite a lower peak PO in UBP, in neither specifically upper-body trained able-bodied cross-country skiers nor non-specifically upper-body trained PARA, when the upper-body was restricted (Paper II). This suggests that neither being specifically trained for the UBP in AB, nor the continuous PO production in ACE leads to higher VO_{2peak}. The lower PO in UBP may hence be solely explained by lower efficiency in UBP compared to ACE, which will be discussed in the next section. Caution is required when comparing VO_{2peak} between the experimental studies conducted in this PhD. However, in unrestricted UBP, higher VO_{2peak} values in AB participants (Paper III) and athletes with different disabilities (Paper IV), were found compared to restricted UBP in AB and PARA, respectively (Paper II). This is likely due to an increased active trunk muscle mass in unrestricted UBP, which is supported by Supplementary data 1 (Appendix I) that consistently shows higher VO_{2peak} in WTR compared to ACE in all four studies included in the meta-analysis[79, 132-134]. Trunk oscillations and shifts in centre of gravity contribute to wheelchair speed in the WTR mode[131] and likely also to power production in UBP, whereas trunk movement is limited in ACE. Overall, differences in continuous and discontinuous power production, respectively, in ACE compared to WERG, WTR and UBP, do not have an impact on VO_{2peak}. In comparison, higher VO_{2peak} values were found in modes and test-setups that involve trunk movement and thereby increase active muscle mass, such as the WTR and the unrestricted UBP mode. Therefore, the WTR and the unrestricted UBP mode may be the recommended exercise modes to increase endurance in individuals with different disabilities with sufficient trunk control.

Influence of test protocol. Next to the test mode, the test protocol employed is important for obtaining the highest possible VO_{2peak} values. The incremental test led to 1 $mL \cdot kg^{-1} \cdot min^{-1}$ higher VO_{2peak} compared to the 3-min test (Paper III), which is in line with similar differences in a comparable study in cross-country skiers[135]. The practical significance of this slight difference in terms of endurance performance during cardiorespiratory exercise testing may be questioned. In this context it has to be noted that all participants performed a thorough warm-up of four 5-min submaximal stages (RPE 9 – RPE 15) initially and an additional standardized 5-min warm-up before each of the three VO_{2peak} tests (1-min, 3-min and incremental). Larger differences in VO_{2peak} between the 3-min and the incremental test might be expected in case no or a much shorter warm-up protocol was performed. A warm-up accelerates VO_2 kinetics and decreases the slow component of VO_2[136]. Given that participants perform a thorough warm-up, bigger differences between the 3-min and the incremental test are not expected with a different incremental protocol. This since a short (20 watt/30 s), two moderate (20 watt/min and 10 watt/30 s) and one long duration (10 watt/min) incremental UBP protocol led to similar VO_{2peak} values despite significant differences in peak PO (Supplementary data 2, Appendix 2). In line with this, Smith et al.[86] and Castro et al.[137] found no differences in VO_{2peak} between a protocol with 6 watt/min and 12 watt/min increases and between a short (2 watt/6 s) compared to a long (1 watt/6 s) protocol, respectively, in incremental tests to exhaustion in AB participants during ACE.

The high and very high ICCs (0.887 - 0.956) in Paper III indicate high relative test-retest reliability for peak PO, VO_{2peak} and peak HR during the closed-end, all-out 1-min and the 3-min test as well as the incremental UBP test to exhaustion. This is in line with several studies that found high relative test-retest reliability of peak parameters during 3-min all-out and incremental tests to exhaustion in not specifically upper-body trained AB during exercise in the ACE or handcycle ergometry mode[84, 88, 138]. In comparison, no studies had yet investigated the relative test-retest reliability of the 1-min test. Whereas the 1-min all-out test also displays high relative test-retest reliability, it is not recommended as a VO_{2peak} test due to the 4 - 5 $mL \cdot kg^{-1} \cdot min^{-1}$ lower VO_{2peak} values compared to the 3-min and incremental test. The ICC is a measure of between-subjects to within-subjects variation. High ICCs might hence be the result of sample heterogeneity and corresponding large variation in outcome parameters[111, 139]. The coefficient of variation for VO_{2peak} in the participants of Paper III was 12%, which indicates heterogeneous physiological responses in even homogeneous, upper-body trained AB. Furthermore, there was some systematic error reflected in consistently higher peak outcome for some of the tests on day 2, which is reflected in the $ICC_{1,1}$ and $ICC_{3,1}$ being lower than the $ICC_{2,1}$. As such the interpretability of the ICC as a measure of relative reliability remains limited and absolute reliability measures, such as the SDC also need to be taken into consideration. The 80% SDC was 12 - 19 watt for peak PO, 2 - 3 $mL \cdot kg^{-1} \cdot min^{-1}$ for VO_{2peak} and 5 - 7 $beats \cdot min^{-1}$ for peak HR for the three VO_{2peak} tests. Even though significant, the 1 $mL \cdot kg^{-1} \cdot min^{-1}$ difference between the 3-min and the incremental test is below the SDC for VO_{2peak}. Concluding from the above, both the 3-min and incremental tests with different increment duration and workload increases can be used as reliable VO_{2peak} tests, given that participants performed a thorough warm-up prior to testing.

Efficiency during upper-body exercise

Influence of test mode. Combining the results of Paper II to IV confirmed that irrespective of restricting or not restricting the upper body, exercise in the UBP is less efficient than in the ACE mode (gross efficiency: 10 - 13 vs 13 - 15%, respectively). This is in line with studies that found the discontinuous movement during WERG exercise to be less efficient that exercising in the ACE mode (gross efficiency: 6 - 12 vs 10 - 16%, respectively)[9, 39, 40]. The lower gross efficiency in discontinuous movement modes may be attributed to that the instantaneous power produced during a shorter part of the cycle needs to be higher in WERG, WTR and UBP in order to produce a given PO compared to a more continuous power production in ACE[39, 140]. Furthermore, in UBP, participants need to lift their arms up against gravity before pulling down on the ropes. This is likely more energy consuming than having the arms supported through the crank throughout the whole movement cycle, where kinetic energy of the downwards movement of one arm might be reutilized during the upwards movement of the contralateral arm.

In UBP, data was collected with both an unrestricted (Paper III and IV) and a restricted trunk movement (Paper II), and an interaction was found in the PO-MR relationship between unrestricted and restricted UBP. The interaction indicates that compared to unrestricted UBP, restricted UBP gets less efficient with increasing PO in both PARA and AB. This is likely due to movement technique being increasingly disturbed when producing higher POs in UBP with a restricted upper body. The PO-MR relationship of restricted ACE does not interact with that of unrestricted UBP. Although not tested, a potential explanation might be that trunk movement is less pronounced in unrestricted ACE, and therefore movement technique is not as influenced in restricted ACE compared to restricted UBP.

Influence of disability. There were no significant differences in gross efficiency between non-specifically upper-body trained PARA and specifically UBP trained AB in neither restricted UBP (10 - 12 vs 10 - 12%, respectively) nor restricted ACE (14 - 16 vs 13 - 15%, respectively) (Paper

II). This finding is in contrast to higher efficiency in PARA compared to AB in a study of Croft et al.[74], employing the WTR mode (6 - 8% vs 4 - 6%, respectively) and Lenton et al.[75] employing the WERG mode (8 vs 6%, respectively). This may be explained by differences in experience with upper-body exercise in AB in Paper II compared to the latter two studies. Whereas the AB in the latter two studies had limited or no experience with wheelchair propulsion, the AB in Paper II, who were all cross-country skiers, were well accustomed to upper-body exercise since double-poling requires significant contribution from the upper-body to forward propulsion[141]. The higher or similar efficiency in PARA compared to AB may indicate that disability-related physiological limitations do not negatively affect efficiency. This is supported by Coutts et al.[142] who found no differences in efficiency between PARA and TETRA. Overall, the lower efficiency in UBP compared to ACE may be attributed to the discontinuous nature of UBP and to the movement of the arms against gravity, but not to differences between PARA and AB. For our findings to be fully applicable to practice, future studies are needed to investigate the differences in VO_{2peak} and efficiency between restricted and unrestricted UBP and ACE.

Gas exchange and blood lactate thresholds during upper-body exercise

Comparison of gas exchange and BLa threshold methods. Although the outcome parameters identified at the VT were not significantly different from the ones identified with the log-log transformed VO_2-BLa at the LT1, the outcome parameters identified with a fixed first rise in BLa at the LT1 were significantly lower. Furthermore, the methods used to identify the VT and the LT1 did not highly correlate with each other indicating that these two thresholds cannot be used interchangeably to determine the AT. The early occurrence of a rise in BLa in upper-body exercise compared to lower-body exercise is in accordance with Beneke et al. [143], who found BLa to be higher at a given workload in activities involving smaller muscle mass, where PO per kg of active muscle mass and, thus, local metabolic stress is increased. In addition, BLa accumulation after cessation of exercise was shown to be faster in individuals with a SCI as compared to able-bodied individuals[11].

The outcome parameters identified with breakpoints in the VCO_2-VE data and in the VE/VCO_2 data at the RCT were closely located to the ones identified with the D_{max} method at the LT2. This indicates that the exercise intensity where a disproportionate increase in VE and in BLa occurs is relatively similar. There was a close location between outcome parameters at the VT, RCT and LT2 despite small significant differences in the submitted version of Paper IV. These differences became insignificant when using Bonferroni corrections. In the submitted version of Paper IV, no adjustment for multiple comparisons were made since this study is considered one that might generate new hypotheses for future research but not necessarily lead to policy changes, which is in accordance with the recommendations of Rothman[144]. Furthermore, most of the outcome parameters identified with the different methods at the RCT and LT2 are low to moderately correlated, coinciding with high individual variation in the outcome parameters within each of the methods used. This indicates that an individual with a high LT2 does not necessarily display a high RCT. The high individual variation may be explained by disability-related differences in the cardio-respiratory system that might affect physiological responses to upper-body exercise. For example, athletes with a SCI exercising in an upper-body mode were shown to vary considerably in their VO_{2peak} depending on their level of injury [119], which might also reflect differences in the % of VO_{2peak} that can be sustained during exercise. Furthermore, individual variation in % of VO_{2peak} may be higher in upper-body exercise compared to lower-body exercise.

The outcome parameters identified by the breakpoints in the VE/VO_2 at the VT and in VE/VCO_2 at the RCT did not significantly differ and were highly correlated. This, together with the rather linear increase in the VO_2-VCO_2 relationship suggests that it is solely the disproportionate rise in VE that leads to a rather rapid increase in the data of the VE/VO_2, and the VE/VCO_2 plots, and to discernible breakpoints in approximately half of the participants. Together with the close location of the breakpoints identified in the VO_2-VCO_2 data at the VT and the VCO_2-VE data at the RCT, this indicates that a two-phase (low-high) rather than a three-phase (low-moderate-high) intensity zone model could be applicable in athletes with a disability who exercise in an upper-body mode. This is in line with a study of Pires et al.[145] who also found one rather than two thresholds in the gas exchange data in upper-body trained able-bodied participants during exercise in the arm crank ergometry mode. However, significant differences between the VT and the RCT in were found in a study of Dekerle et al.[62], who tested able-bodied participants in the arm crank ergometry mode and of Leicht et al.[97], who tested wheelchair athletes in the wheelchair treadmill mode. Follow-up studies are needed to look into whether outcome measures identified with different gas exchange and BLa threshold methods to determine the AT and ANT differ during upper-body exercise in AB and participants with disability and between upper- and lower-body exercise in AB participants.

Breakpoints in the gas exchange data at VT and RCT. All gas exchange threshold methods have in common that there is an a priori assumption of the presence of a breakpoint, defined as "a place where an interruption or change occurs"[146]. However, the presence or absence of breakpoints in the gas exchange data is a debated topic[37, 147]. Thus, in addition to the breakpoint models used to identify the VT and the RCT in Paper IV, continuous no-breakpoint models were fitted to the gas exchange data to investigate if there are clear breakpoints. In most participants, continuous no-breakpoint models fit the gas exchange data better. Furthermore, also in the participants with suggested breakpoint presence, the breakpoint models only fit marginally better ($\Delta R^2 < 0.002$) and the practical consequences of these differences are debatable. Overall, it can be questioned if clear breakpoints really exist in the gas exchange data of athletes with disabilities in an upper-body exercise mode.

% of VO_{2peak} at the VT and RCT. The interpretation of the % of VO_{2peak} at the VT and the RCT as two separate thresholds might not be valid since they are closely located in Paper IV. The close location of the VT and RCT might be a consequence of the absence of clear breakpoints in the gas exchange data, where the two regression lines identify a breakpoint that is located somewhere in the middle of the data irrespective of the method used. However, the 68% of VO_{2peak} identified at the VT in the current study is relatively similar compared to the 65 - 74% in athletes of different Paralympic sitting sports being tested in different upper-body mode[5, 69, 97]. In comparison, the VT in the Paper IV occurs at a higher % of VO_{2peak} as compared to the VT in most other studies that investigated able-bodied participants and participants with a disability in different upper-body exercise modes (in all studies, % of VO_{2peak} < 60%)[60-62, 72, 148-150]. This is likely related to the fact that participants in these studies were less physically active compared to the Para ice hockey players in Paper IV. Overall, despite a lack of clear breakpoints in the gas exchange data in Paper IV and the close location of the VT and RCT, the % of VO_{2peak} at the VT is similar to other studies who test upper-body trained participants with a disability in an upper-body exercise mode.

Compared to the 68 - 72% of VO_{2peak} at the RCT in our study, higher values were found in paraplegic wheelchair basketball players tested in the WTR mode (78%)[97] and in able-bodied participants tested in the ACE mode (83%)[62]. Furthermore, the % of VO_{2peak} at the RCT in Paper IV was also lower compared to the 85 - 89% in elite able-bodied athletes during lower-body[151] and whole-body exercise[152]. This suggests that the latter wheelchair basketball players and able-bodied athletes are able to maintain performance at a higher % of VO_{2peak}. This may be attributed to them being better trained with respect to cardio-respiratory fitness, compared to the Para ice hockey players investigated in Paper IV, who next to training endurance also spend quite some time training strength and power to improve their sprint ability.

Methodological considerations

Sample size, sample characteristics and sample heterogeneity. The challenge in elite sports research – and even more so in the Paralympic field – is a scarcity of athletes and their heterogeneity in physical characteristics and disabilities, as was the case in Paper I – IV included in this PhD. Combining the results of several studies by means of a meta-analysis constitutes a good, if currently not the only, opportunity to provide knowledge based on a larger sample size. However, the numbers of included studies and participants remained low even in the meta-analyses conducted in Paper I, in all sports other than wheelchair racing, wheelchair basketball and wheelchair rugby. This reduced the generalizability of VO_{2peak} values to the population of athletes participating in most sports disciplines. A further limitation of Paper I was that detailed information on training status was missing in most and detailed information on sex, age, body mass and type of disability was missing in some studies. Some of the differences in VO_{2peak} between sports disciplines and some of the variation within each sport discipline may be attributed to the differences in anthropometrics, type of disability and training status. Therefore, the meta-analyses of Paper I are not able to separately investigate the effect participating in a certain sports discipline on VO_{2peak}.

When individuals with a disability exercise in an upper-body mode, two factors mainly influence physiological outcome parameters: 1) the upper-body exercise mode itself and 2) the impact of an individual’s disability on movement function and physiological capacity. To be able to investigate the influence of the upper-body exercise mode versus disability-related limitations, the inclusion of AB in addition to participants with a disability is be necessary. In Paper III, it was decided to only include male upper-body trained AB participants to achieve a sufficient sample size with relatively homogeneous participants since especially relative reliability is affected by sample heterogeneity. However, due to this, the generalizability of our findings to individuals with a disability remains limited. In Paper IV, next to investigating athletes with different disabilities, the addition of an AB control group would have been desirable, since as of now it is not known whether the findings may be attributed to limitations related to upper-body exercise mode or disability-related limitations. Furthermore, the Para ice hockey players included in the latter study contained three athletes with athrogryposis multiplex congenita, motor cerebral palsy and a single leg amputation in addition to the athletes with SCI and spina bifida with different levels of paraplegia. Some of the variation found within each of the threshold methods in this Paper IV may, therefore, be attributed to differences in disabilities.

In comparison in Paper II, both PARA and AB were included. However, all AB were specifically trained in the UBP mode, and most athletes with disabilities were not specifically trained for either mode. The inclusion of a specifically trained group for UBP and ACE within AB (e.g. cross-country skiers vs kayakers, respectively) and PARA (e.g. Para ice hockey players, Para cross country skiers or Para biathletes vs hand-cyclists, respectively) was initially aimed for. This would have allowed us to investigate more thoroughly if differences between UBP and ACE could be attributed to being specifically trained for each mode.

Test mode and protocol. A further limitation of Paper I was that different test modes, test equipment and test protocols were used in the studies included for each sports discipline. Some of the differences in VO_{2peak} between sports disciplines and some of the variation within each sport discipline might additionally be explained by the different test modes and protocols used. This further limited the possibility to investigate the effect of participating in a certain sports discipline on VO_{2peak}. In Paper II, only the restricted UBP and the restricted ACE mode were used. Especially in UBP, the trunk is dynamically used during training and competition. As such, our findings are not applicable to a practical setting and future studies need to investigate the difference between restricted and unrestricted UBP. In paper III, no familiarization session with each of the three VO_{2peak} tests (1-min, 3-min and incremental test to exhaustion) was performed. In hindsight a familiarization session should be performed for all three tests if the main outcome parameter is peak PO and for the 1-min and the incremental test if the main outcome parameter is VO_{2peak}, since there were significant day-to-day differences in these parameters.

In Paper IV, an intermittent incremental protocol was employed to be able to take a BLa sample from the fingertip in between stages. This was done in order to be able to investigate if gas exchange and BLa thresholds differ in determining the aerobic and the anaerobic threshold. In comparison to the continuous incremental protocol, which is most commonly used during gas exchange threshold analysis[29, 33, 93, 97], the intermittent protocol led to gaps in the gas exchange data. This might have affected the accuracy with which breakpoints in the gas exchange data were detected, especially if breakpoints occurred between stages. In addition, the thresholds identified in Paper IV were not validated against the ones identified with a maximal lactate steady state protocol. Such a protocol consists of 30-min stages at increasing intensity and is considered the gold standard for determining the highest constant exercise intensity[153].

Scaling. In the absence of valid scaling methods for VO_{2peak} to account for differences in body size in athletes with different disabilities, absolute and body-mass normalized VO_{2peak} values are provided in the studies included in this PhD. Absolute VO_{2peak} values are - next to the endurance capacity - influenced by body size, with higher VO_{2peak} values in participants with higher body size. To be able to compare the VO_{2peak} of individuals whilst reducing the effect of different body sizes, VO_{2peak} values are normalized by total body-mass. However, when testing in an upper-body mode, only a limited part of the total muscle mass is active, which is further reduced in participants with a disability. In addition, participants with disabilities are even more heterogeneous in their distribution of body mass between the upper- and lower-body compared to AB. Therefore, the validity of normalizing VO_{2peak} by total body mass when testing participants with different disabilities in an upper-body mode is questionable.

CONCLUSIONS

Participating in sports with continuously high movement demands, being a man, not being tetraplegic or having an amputation compared to being paraplegic, as well as testing in a wheelchair treadmill or unrestricted upper-body poling mode compared to an arm crank ergometry, wheelchair ergometry or restricted upper-body poling mode are favourable for high absolute and body-mass VO_{2peak} values (Paper I – IV). Movement differences between the upper-body poling and arm crank ergometry mode do not seem to have an impact on VO_{2peak} when the upper-body is restricted, but the discontinuous power production in upper-body poling leads to lower efficiency compared to arm crank ergometry. In addition, compared to able-bodied participants, spinal-cord injury related limitations negatively influence VO_{2peak} but not efficiency in paraplegic participants (Paper II). In upper-body poling, both a 3-min and an incremental test of moderate duration are highly reliable VO_{2peak} tests (Paper III). Furthermore, the breakpoint methods used to identify the ventilatory threshold and the fixed methods used to identify the first lactate threshold cannot be used interchangeably. In addition, the close location of the ventilatory threshold, the respiratory compensation threshold and the second lactate threshold does not allow us to distinguish the aerobic and anaerobic threshold, indicating the presence of only one threshold in athletes with a disability exercising in an upper-body mode. Since continuous no-breakpoint models fit the gas exchange data better in athletes with a disability during upper-body poling, it is questionable if clear breakpoints exist in these participants in an upper-body test mode (Paper IV).

ACKNOWLEDGEMENTS

I am immensely grateful to the people that have helped me grow both personally as well as professionally over the past years and during the course of the PhD and would therefore like to thank:

My main supervisor, Professor Øyvind Sandbakk, for finding a great balance of guiding me through this PhD, being very available for discussions day and night, while at the same time giving me freedom to take responsibility in projects and to make my own mistakes.

My co-supervisor, Professor Gertjan Ettema, for the many discussions on and off topic and for being a living example of what I believe true research ethics are.

My co-supervisor, Dr Berit Brurok, for contributing to the team with a strong clinical background. Thank you also for the hours spent on personal guidance in life and for being a sounding board to ideas, thoughts and challenges that needed reflection.

My colleagues and friends for the many great discussions on research and other topics in life. Thank you for being there when life got challenging! And thanks for your acceptance of my lack of cross-country skiing knowledge and teaching me a thing or two - way to go! Thanks also to my cube-buddy and PhD partner in crime, Jørgen Danielsen, who surprisingly seemed to have survived the past years in the same office just fine!

Yvette Hoel at the Olympic centre of Mid-Norway here in Trondheim, Cato Zahl Petersen and Espen Tønnessen at the National Olympic centre in Oslo for being positive towards our research and for acting as a bridge between the Paralympic athletes and our research.

The many brains and hands-on of the Master students Katrin Daetwyler, Laura Gürtler, Mirjam Mellema, Maaike Moes, and Lilli Marie Myhre Ofstad, who were involved in the data collections of the studies included in this PhD. Without you collecting this much data would not have been possible!

The crazy, mostly Norwegian participants involved in the data collections, who spent many hours of their free time with us in the laboratory. Thank you for being patient when – more often than we had wished – one or another part of the laboratory equipment bailed on us. Thank you for your willingness to push yourself to exhaustion, which provided us with the data of the studies included in this PhD!

Finally, to my beloved family – mum, dad, Maria, Martin and my aunt Kathrin – and to my close friends, especially Lea, whom I spent far too little time with the past couple of years due to either working hard here in Trondheim or spending time and getting some head space on a mountain somewhere in the world. Thank you for your endless support, for accepting and supporting me for who I am and what I do!

REFERENCES

1. Webborn AD. Fifty years of competitive sport for athletes with disabilities: 1948-1998. Br J Sports Med. 1999;33(2):138.
2. Paralympic sports: International Paralympic Committee; 2018 [cited 2018 22.05.]. Available from: https://www.paralympic.org/sports.
3. Classification: International Paralympic Committee; 2018 [cited 2018 22.05.]. Available from: https://www.paralympic.org/classification.
4. Perret C. Elite-adapted wheelchair sports performance: a systematic review. Disabil Rehabil. 2015:1-9. doi:10.3109/09638288.2015.1095951
5. Bernardi M, Guerra E, Di Giacinto B, Di Cesare A, Castellano V, Bhambhani Y. Field evaluation of paralympic athletes in selected sports: implications for training. Med Sci Sports Exerc. 2010;42(6):1200-1208. doi:10.1249/MSS.0b013e3181c67d82
6. Bernardi M, Carucci S, Faiola F, Egidi F, Marini C, Castellano V, et al. Physical fitness evaluation of paralympic winter sports sitting athletes. Clin J Sport Med. 2012;22(1):26-30. doi:10.1097/JSM.0b013e31824237b5
7. Hagberg JM, Hickson RC, Ehsani AA, Holloszy JO. Faster adjustment to and recovery from submaximal exercise in the trained state. J Appl Physiol Respir Environ Exerc Physiol. 1980;48(2):218-224. doi:10.1152/jappl.1980.48.2.218
8. Drory Y, Ohry A, Brooks ME, Dolphin D, Kellermann JJ. Arm crank ergometry in chronic spinal cord injured patients. Arch Phys Med Rehabil. 1990;71(6):389-392.
9. Glaser R, Sawka M, Brune M, Wilde S. Physiological responses to maximal effort wheelchair and arm crank ergometry. J Appl Physiol. 1980;48(6):1060-1064.
10. 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. doi:10.1080/02640410600702883
11. Leicht C, Perret C. Comparison of Blood Lactate Elimination in Individuals With Paraplegia and Able-Bodied Individuals During Active Recovery From Exhaustive Exercise. The Journal of Spinal Cord Medicine. 2008;31(1):60-64. doi:10.1080/10790268.2008.11753982
12. Price MJ, Bottoms L, Smith PM, Nicholettos A. The effects of an increasing versus constant crank rate on peak physiological responses during incremental arm crank ergometry. J Sports Sci. 2011;29(3):263-269. doi:10.1080/02640414.2010.525520
13. Tropp H, Samuelsson K, Jorfeldt L. Power output for wheelchair driving on a treadmill compared with arm crank ergometry. Br J Sports Med. 1997;31(1):41-44.
14. Bassett DR, Jr., Howley ET. Limiting factors for maximum oxygen uptake and determinants of endurance performance. Med Sci Sports Exerc. 2000;32(1):70-84.
15. Coyle EF. Integration of the physiological factors determining endurance performance ability. Exerc Sport Sci Rev. 1995;23:25-63.
16. Joyner MJ. Modeling: optimal marathon performance on the basis of physiological factors. J Appl Physiol (1985). 1991;70(2):683-687. doi:10.1152/jappl.1991.70.2.683
17. Joyner MJ, Coyle EF. Endurance exercise performance: the physiology of champions. J Physiol. 2008;586(1):35-44. doi:10.1113/jphysiol.2007.143834
18. Herbst R. Der Gasstoffwechsel als Mass der koerperlichen Leistungsfaehigkeit. I. Mitteilung: die Bestimmung des Sauerstoffaufnahmevermoegens bei Gesunden. . Deu Arch Klin Med. 1928;162:33-50.
19. Hill AV, Long C, Lupton H. Muscular exercise, lactic acid, and the supply and utilisation of oxygen. Proceedings of the Royal Society of London Series B, Containing Papers of a Biological Character. 1924;97(681):84-138.

20. Fick A. Ueber die Messung des Blutquantum in den Herzventrikeln. Sb Phys Med Ges Worzburg. 1870:16-17.
21. Wasserman K, Hansen JE, Sue DY, Stringer WW, Whipp BJ. Principles of exercise testing and interpretation: including pathophysiology and clinical applications. Med Sci Sports Exerc. 2005;37(7):1249.
22. Hopman MT, Verheijen PH, Binkhorst RA. Volume changes in the legs of paraplegic subjects during arm exercise. J Appl Physiol (1985). 1993;75(5):2079-2083. doi:10.1152/jappl.1993.75.5.2079
23. Theisen D. Cardiovascular determinants of exercise capacity in the Paralympic athlete with spinal cord injury. Exp Physiol. 2012;97(3):319-324. doi:10.1113/expphysiol.2011.063016
24. Thijssen DH, Steendijk S, Hopman MT. Blood redistribution during exercise in subjects with spinal cord injury and controls. Med Sci Sports Exerc. 2009;41(6):1249-1254. doi:10.1249/MSS.0b013e318196c902
25. Hopman MT, Oeseburg B, Binkhorst RA. Cardiovascular responses in paraplegic subjects during arm exercise. Eur J Appl Physiol Occup Physiol. 1992;65(1):73-78.
26. Shephard RJ, Bouhlel E, Vandewalle H, Monod H. Muscle mass as a factor limiting physical work. J Appl Physiol (1985). 1988;64(4):1472-1479.
27. Arena R, Myers J, Williams MA, Gulati M, Kligfield P, Balady GJ, et al. Assessment of functional capacity in clinical and research settings: a scientific statement from the American Heart Association Committee on Exercise, Rehabilitation, and Prevention of the Council on Clinical Cardiology and the Council on Cardiovascular Nursing. Circulation. 2007;116(3):329-343. doi:10.1161/circulationaha.106.184461
28. Binder RK, Wonisch M, Corra U, Cohen-Solal A, Vanhees L, Saner H, 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. doi:10.1097/HJR.0b013e328304fed4
29. Beaver WL, Wasserman K, Whipp BJ. A new method for detecting anaerobic threshold by gas exchange. J Appl Physiol (1985). 1986;60(6):2020-2027.
30. Reinhard U, Muller PH, Schmulling RM. Determination of anaerobic threshold by the ventilation equivalent in normal individuals. Respiration. 1979;38(1):36-42.
31. Bishop D, Jenkins DG, Mackinnon LT. The relationship between plasma lactate parameters, Wpeak and 1-h cycling performance in women. Med Sci Sports Exerc. 1998;30(8):1270-1275.
32. Yoshida T, Chida M, Ichioka M, Suda Y. Blood lactate parameters related to aerobic capacity and endurance performance. Eur J Appl Physiol Occup Physiol. 1987;56(1):7-11.
33. Beaver WL, Wasserman K, Whipp BJ. Improved detection of lactate threshold during exercise using a log-log transformation. J Appl Physiol (1985). 1985;59(6):1936-1940.
34. Wasserman K. The anaerobic threshold: definition, physiological significance and identification. Adv Cardiol. 1986;35:1-23.
35. Cheng B, Kuipers H, Snyder AC, Keizer HA, Jeukendrup A, Hesselink M. A new approach for the determination of ventilatory and lactate thresholds. Int J Sports Med. 1992;13(7):518-522. doi:10.1055/s-2007-1021309
36. Stegmann H, Kindermann W. Comparison of prolonged exercise tests at the individual anaerobic threshold and the fixed anaerobic threshold of 4 mmol.l(-1) lactate. Int J Sports Med. 1982;3(2):105-110. doi:10.1055/s-2008-1026072
37. Myers J, Ashley E. Dangerous curves. A perspective on exercise, lactate, and the anaerobic threshold. Chest. 1997;111(3):787-795.

38. Stainbsy WN, Gladden LB, Barclay JK, Wilson BA. Exercise efficiency: validity of base-line subtractions. J Appl Physiol. 1980;48(3):518-522.
39. Hintzy F, Tordi N, Perrey S. Muscular efficiency during arm cranking and wheelchair exercise: a comparison. Int J Sports Med. 2002;23(6):408-414. doi:10.1055/s-2002-33734
40. Mukherjee G, Samanta A. Physiological response to the ambulatory performance of hand-rim and arm-crank propulsion systems. J Rehabil Res Dev. 2001;38(4):391.
41. Minetti AE, Moia C, Roi GS, Susta D, Ferretti G. Energy cost of walking and running at extreme uphill and downhill slopes. J Appl Physiol (1985). 2002;93(3):1039-1046. doi:10.1152/japplphysiol.01177.2001
42. Gaesser GA, Tucker WJ, Sawyer BJ, Bhammar DM, Angadi SS. Cycling efficiency and energy cost of walking in young and older adults. J Appl Physiol (1985). 2018;124(2):414-420. doi:10.1152/japplphysiol.00789.2017
43. Hopker J, Jobson S, Carter H, Passfield L. Cycling efficiency in trained male and female competitive cyclists. J Sports Sci Med. 2010;9(2):332-337.
44. Moseley L, Achten J, Martin JC, Jeukendrup AE. No differences in cycling efficiency between world-class and recreational cyclists. Int J Sports Med. 2004;25(5):374-379. doi:10.1055/s-2004-815848
45. Bijker KE, De Groot G, Hollander AP. Delta efficiencies of running and cycling. Med Sci Sports Exerc. 2001;33(9):1546-1551.
46. Hegge AM, Bucher E, Ettema G, Faude O, Holmberg HC, Sandbakk O. Gender differences in power production, energetic capacity and efficiency of elite cross-country skiers during whole-body, upper-body, and arm poling. Eur J Appl Physiol. 2015. doi:10.1007/s00421-015-3281-y
47. Gaesser GA, Brooks GA. Muscular efficiency during steady-rate exercise: effects of speed and work rate. J Appl Physiol. 1975;38(6):1132-1139. doi:10.1152/jappl.1975.38.6.1132
48. Ettema G, Loras HW. Efficiency in cycling: a review. Eur J Appl Physiol. 2009;106(1):1-14. doi:10.1007/s00421-009-1008-7
49. Pendergast D, Cerretelli P, Rennie DW. Aerobic and glycolytic metabolism in arm exercise. J Appl Physiol Respir Environ Exerc Physiol. 1979;47(4):754-760. doi:10.1152/jappl.1979.47.4.754
50. Reybrouck T, Heigenhauser GF, Faulkner JA. Limitations to maximum oxygen uptake in arms, leg, and combined arm-leg ergometry. J Appl Physiol. 1975;38(5):774-779. doi:10.1152/jappl.1975.38.5.774
51. Bergh U, Kanstrup I, Ekblom B. Maximal oxygen uptake during exercise with various combinations of arm and leg work. J Appl Physiol. 1976;41(2):191-196.
52. Pandolf KB, Billings DS, Drolet LL, Pimental NA, Sawka MN. Differentiated ratings of perceived exertion and various physiological responses during prolonged upper and lower body exercise. Eur J Appl Physiol Occup Physiol. 1984;53(1):5-11. doi:10.1007/bf00964681
53. Bevegard S, Freyschuss U, Strandell T. Circulatory adaptation to arm and leg exercise in supine and sitting position. J Appl Physiol. 1966;21(1):37-46. doi:10.1152/jappl.1966.21.1.37
54. Freyschuss U. Comparison between arm and leg exercise in women and men. Scand J Clin Lab Invest. 1975;35(8):795-800.
55. Sawka MN, Miles DS, Petrofsky JS, Wilde SW, Glaser RM. Ventilation and acid-base equilibrium for upper body and lower body exercise. Aviat Space Environ Med. 1982;53(4):354-359.

56. Eston RG, Brodie DA. Responses to arm and leg ergometry. Br J Sports Med. 1986;20(1):4-6.
57. Glaser RM, Sawka MN, Laubach LL, Suryaprasad AG. Metabolic and cardiopulmonary responses to wheelchair and bicycle ergometry. J Appl Physiol Respir Environ Exerc Physiol. 1979;46(6):1066-1070. doi:10.1152/jappl.1979.46.6.1066
58. Marais G, Dupont L, Maillet M, Weissland T, Vanvelcenaher J, Pelayo P. Cardiorespiratory and efficiency responses during arm and leg exercises with spontaneously chosen crank and pedal rates. Ergonomics. 2002;45(9):631-639. doi:10.1080/00140130210151821
59. Kang J, Robertson RJ, Goss FL, Dasilva SG, Suminski RR, Utter AC, et al. Metabolic efficiency during arm and leg exercise at the same relative intensities. Med Sci Sports Exerc. 1997;29(3):377-382.
60. Davis JA, Vodak P, Wilmore JH, Vodak J, Kurtz P. Anaerobic threshold and maximal aerobic power for three modes of exercise. J Appl Physiol. 1976;41(4):544-550.
61. Yasuda N, Gaskill SE, Ruby BC. No gender-specific differences in mechanical efficiency during arm or leg exercise relative to ventilatory threshold. Scand J Med Sci Sports. 2008;18(2):205-212. doi:10.1111/j.1600-0838.2007.00637.x
62. Dekerle J, Dupont L, Caby I, Marais G, Vanvelcenaher J, Lavoie JM, et al. Ventilatory thresholds in arm and leg exercises with spontaneously chosen crank and pedal rates. Percept Mot Skills. 2002;95(3 Pt 2):1035-1046. doi:10.2466/pms.2002.95.3f.1035
63. Keyser RE, Mor D, Andres FF. Cardiovascular responses and anaerobic threshold for bicycle and arm ergometer exercise. Arch Phys Med Rehabil. 1989;70(9):687-691.
64. Price M, Campbell I. Thermoregulatory responses of spinal cord injured and able-bodied athletes to prolonged upper body exercise and recovery. Spinal Cord. 1999;37(11).
65. Hopman MT, Houtman S, Groothuis JT, Folgering HT. The effect of varied fractional inspired oxygen on arm exercise performance in spinal cord injury and able-bodied persons. Arch Phys Med Rehabil. 2004;85(2):319-323.
66. Van Loan MD, McCluer S, Loftin JM, Boileau R. Comparison of physiological responses to maximal arm exercise among able-bodied, paraplegics and quadriplegics. Spinal Cord. 1987;25(5):397.
67. West CR, Gee CM, Voss C, Hubli M, Currie KD, Schmid J, et al. Cardiovascular control, autonomic function, and elite endurance performance in spinal cord injury. Scand J Med Sci Sports. 2015;25(4):476-485. doi:10.1111/sms.12308
68. 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.
69. Coutts KD, McKenzie DC. Ventilatory thresholds during wheelchair exercise in individuals with spinal cord injuries. Paraplegia. 1995;33(7):419-422. doi:10.1038/sc.1995.85
70. Hopman MT, Pistorius M, Kamerbeek IC, Binkhorst RA. Cardiac output in paraplegic subjects at high exercise intensities. Eur J Appl Physiol Occup Physiol. 1993;66(6):531-535.
71. Lin KH, Lai JS, Kao MJ, Lien IN. Anaerobic threshold and maximal oxygen consumption during arm cranking exercise in paraplegia. Arch Phys Med Rehabil. 1993;74(5):515-520.
72. Schneider DA, Sedlock DA, Gass E, Gass G. VO2peak 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.
73. Brown DD, Knowlton RG, Hamill J, Schneider TL, Hetzler RK. Physiological and biomechanical differences between wheelchair-dependent and able-bodied subjects during wheelchair ergometry. Eur J Appl Physiol Occup Physiol. 1990;60(3):179-182.

74. Croft L, Lenton J, Tolfrey K, Goosey-Tolfrey V. The effects of experience on the energy cost of wheelchair propulsion. Eur J Phys Rehabil Med. 2013;49(6):865-873.
75. Lenton JP, Fowler NE, van der Woude L, Goosey-Tolfrey VL. Wheelchair propulsion: effects of experience and push strategy on efficiency and perceived exertion. Appl Physiol Nutr Metab. 2008;33(5):870-879. doi:10.1139/h08-072
76. McCafferty WB, Horvath SM. Specificity of exercise and specificity of training: a subcellular review. Res Q. 1977;48(2):358-371.
77. Bloemen MA, de Groot JF, Backx FJ, Westerveld RA, Takken T. Arm cranking versus wheelchair propulsion for testing aerobic fitness in children with spina bifida who are wheelchair dependent. J Rehabil Med. 2015;47(5):432-437. doi:10.2340/16501977-1944
78. Gass E, Harvey L, Gass G. Maximal physiological responses during arm cranking and treadmill wheelchair propulsion in T4–T6 paraplegic men. Spinal Cord. 1995;33(5):267-270.
79. Arabi H, Vandewalle H, Pitor P, de Lattre J, Monod H. Relationship between maximal oxygen uptake on different ergometers, lean arm volume and strength in paraplegic subjects. Eur J Appl Physiol Occup Physiol. 1997;76(2):122-127. doi:10.1007/s004210050223
80. Gayle GW, Pohlman RL, Glaser RM, Davis GM. Cardiorespiratory and perceptual responses to arm crank and wheelchair exercise using various handrims in male paraplegics. Res Q Exerc Sport. 1990;61(3):224-232. doi:10.1080/02701367.1990.10608683
81. Martel G, Noreau L, Jobin J. Physiological responses to maximal exercise on arm cranking and wheelchair ergometer with paraplegics. Spinal Cord. 1991;29(7):447-456.
82. Bar-Or O, Zwiren LD. Maximal oxygen consumption test during arm exercise--reliability and validity. J Appl Physiol. 1975;38(3):424-426.
83. Bhambhani YN, Eriksson P, Steadward RD. Reliability of peak physiological responses during wheelchair ergometry in persons with spinal cord injury. Arch Phys Med Rehabil. 1991;72(8):559-562.
84. Leicht AS, Sealey RM, Sinclair WH. The reliability of VO2(peak) determination in healthy females during an incremental arm ergometry test. Int J Sports Med. 2009;30(7):509-515. doi:10.1055/s-0029-1202351
85. Leicht CA, Tolfrey K, Lenton JP, Bishop NC, Goosey-Tolfrey VL. The verification phase and reliability of physiological parameters in peak testing of elite wheelchair athletes. Eur J Appl Physiol. 2013;113(2):337-345. doi:10.1007/s00421-012-2441-6
86. Smith PM, Amaral I, Doherty M, Price MJ, Jones AM. The influence of ramp rate on VO2peak and "excess" VO2 during arm crank ergometry. Int J Sports Med. 2006;27(8):610-616. doi:10.1055/s-2005-865857
87. 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. doi:10.1055/s-2004-817880
88. Flueck JL, Lienert M, Schaufelberger F, Perret C. Reliability of a 3-min all-out Arm Crank Ergometer Exercise Test. Int J Sports Med. 2015;36(10):809-813. doi:10.1055/s-0035-1548811
89. Baumgart JK, Sandbakk O. Laboratory Determinants of Repeated-Sprint and Sport-Specific-Technique Ability in World-Class Ice Sledge Hockey Players. Int J Sports Physiol Perform. 2016;11(2):182-190. doi:10.1123/ijspp.2014-0516
90. Sperlich B, Haegele M, Thissen A, Mester J, Holmberg HC. Are peak oxygen uptake and power output at maximal lactate steady state obtained from a 3-min all-out cycle test? Int J Sports Med. 2011;32(6):433-437. doi:10.1055/s-0031-1271770

91. Wasserman K. The anaerobic threshold measurement to evaluate exercise performance. Am Rev Respir Dis. 1984;129(2 Pt 2):S35-40. doi:10.1164/arrd.1984.129.2P2.S35
92. Wasserman K. Determinants and detection of anaerobic threshold and consequences of exercise above it. Circulation. 1987;76(6 Pt 2):Vi29-39.
93. 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 Sports Exerc. 2001;33(11):1841-1848.
94. Maffulli N, Testa V, Capasso G. Anaerobic threshold determination in master endurance runners. J Sports Med Phys Fitness. 1994;34(3):242-249.
95. Chicharro JL, Perez M, Vaquero AF, Lucia A, Legido JC. Lactic threshold vs ventilatory threshold during a ramp test on a cycle ergometer. J Sports Med Phys Fitness. 1997;37(2):117-121.
96. Ribeiro J, Figueiredo P, Sousa M, De Jesus K, Keskinen K, Vilas-Boas JP, et al. Metabolic and ventilatory thresholds assessment in front crawl swimming. J Sports Med Phys Fitness. 2015;55(7-8):701-707.
97. 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. doi:10.1007/s00421-014-2886-x
98. Downs SH, Black N. The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. J Epidemiol Community Health. 1998;52(6):377-384.
99. DerSimonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials. 1986;7(3):177-188.
100. Pigott T. Advances in meta-analysis: Springer Science & Business Media; 2012.
101. van Tulder M, Furlan A, Bombardier C, Bouter L. Updated method guidelines for systematic reviews in the cochrane collaboration back review group. Spine (Phila Pa 1976). 2003;28(12):1290-1299. doi:10.1097/01.brs.0000065484.95996.af
102. Cane J, Seidman B, Sowash J, Otto R, Wygand J. A Comparison Of The Computrainer Load Simulator And Traditional Cycle Ergometry 1239. Med Sci Sports Exerc. 1996;28(5):208.
103. Foss O, Hallen J. Validity and stability of a computerized metabolic system with mixing chamber. Int J Sports Med. 2005;26(7):569-575. doi:10.1055/s-2004-821317
104. Carter J, Jeukendrup AE. Validity and reliability of three commercially available breath-by-breath respiratory systems. Eur J Appl Physiol. 2002;86(5):435-441. doi:10.1007/s00421-001-0572-2
105. Medbo JI, Mamen A, Holt Olsen O, Evertsen F. Examination of four different instruments for measuring blood lactate concentration. Scand J Clin Lab Invest. 2000;60(5):367-380.
106. Davison RR, Coleman D, Balmer J, Nunn M, Theakston S, Burrows M, et al. Assessment of blood lactate: practical evaluation of the Biosen 5030 lactate analyzer. Med Sci Sports Exerc. 2000;32(1):243-247.
107. Pyne DB, Boston T, Martin DT, Logan A. Evaluation of the Lactate Pro blood lactate analyser. Eur J Appl Physiol. 2000;82(1-2):112-116. doi:10.1007/s004210050659
108. Peronnet F, Massicotte D. Table of nonprotein respiratory quotient: an update. Can J Sport Sci. 1991;16(1):23-29.
109. Robergs RA, Dwyer D, Astorino T. Recommendations for improved data processing from expired gas analysis indirect calorimetry. Sports Med. 2010;40(2):95-111. doi:10.2165/11319670-000000000-00000

110. Field A. Discovering statistics using IBM SPSS statistics: Sage; 2013.
111. Hopkins WG. Measures of reliability in sports medicine and science. Sports Med. 2000;30(1):1-15.
112. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet. 1986;1(8476):307-310.
113. Weir JP. Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. J Strength Cond Res. 2005;19(1):231-240. doi:10.1519/15184.1
114. Plichta SB, Kelvin EA, Munro BH. Munro's statistical methods for health care research: Wolters Kluwer Health/Lippincott Williams & Wilkins; 2013.
115. Bozdogan H. Model selection and Akaike's information criterion (AIC): The general theory and its analytical extensions. Psychometrika. 1987;52(3):345-370.
116. Burnham KP, Anderson DR. Model selection and multimodel inference: a practical information-theoretic approach: Springer Science & Business Media; 2003.
117. Wagenmakers E-J, Farrell S. AIC model selection using Akaike weights. Psychonomic bulletin & review. 2004;11(1):192-196.
118. Sarro KJ, Misuta MS, Burkett B, Malone LA, Barros RML. Tracking of wheelchair rugby players in the 2008 Demolition Derby final. J Sports Sci. 2010;28(2):193-200. doi:10.1080/02640410903428541
119. Bhambhani Y. Physiology of wheelchair racing in athletes with spinal cord injury. Sports Med. 2002;32(1):23-51.
120. West CR, Romer LM, Krassioukov A. Autonomic function and exercise performance in elite athletes with cervical spinal cord injury. Med Sci Sports Exerc. 2013;45(2):261-267. doi:10.1249/MSS.0b013e31826f5099
121. Croft L, Dybrus S, Lenton J, Goosey-Tolfrey V. A comparison of the physiological demands of wheelchair basketball and wheelchair tennis. Int J Sports Physiol Perform. 2010;5(3):301-315.
122. Goosey VL, Campbell IG. Pushing economy and propulsion technique of wheelchair racers at three speeds. Adapt Phys Activ Q. 1998;15(1):36-50 15p.
123. Goosey VL, Campbell IG, Fowler NE. Effect of push frequency on the economy of wheelchair racers. Med Sci Sports Exerc. 2000;32(1):174-181.
124. Goosey-Tolfrey VL, Tolfrey K. The oxygen uptake-heart rate relationship in trained female wheelchair athletes. J Rehabil Res Dev. 2004;41(3b):415-420.
125. Knechtle B, Muller G, Willmann F, Eser P, Knecht H. Fat oxidation at different intensities in wheelchair racing. Spinal Cord. 2004;42(1):24-28. doi:10.1038/sj.sc.3101548
126. Tolfrey K, Goosey-Tolfrey VL, Campbell IG. Oxygen uptake-heart rate relationship in elite wheelchair racers. Eur J Appl Physiol. 2001;86(2):174-178.
127. Vinet A, Bernard PL, Poulain M, Varray A, Le Gallais D, Micallef JP. Validation of an incremental field test for the direct assessment of peak oxygen uptake in wheelchair-dependent athletes. Spinal Cord. 1996;34(5):288-293.
128. Bergh U, Sjodin B, Forsberg A, Svedenhag J. The relationship between body mass and oxygen uptake during running in humans. Med Sci Sports Exerc. 1991;23(2):205-211.
129. Jacobs KA, Burns P, Kressler J, Nash MS. Heavy reliance on carbohydrate across a wide range of exercise intensities during voluntary arm ergometry in persons with paraplegia. J Spinal Cord Med. 2013;36(5):427-435. doi:10.1179/2045772313y.0000000123
130. Alves ES, Santos RV, Ruiz FS, Lira FS, Almeida AA, Lima G, et al. Physiological and lipid profile response to acute exercise at different intensities in individuals with spinal cord injury. Spinal Cord Ser Cases. 2017;3:17037. doi:10.1038/scsandc.2017.37

131. Vanlandewijck Y, Theisen D, Daly D. Wheelchair propulsion biomechanics: implications for wheelchair sports. Sports Med. 2001;31(5):339-367.
132. Gass GC, Camp EM. The maximum physiological responses during incremental wheelchair and arm cranking exercise in male paraplegics. Med Sci Sports Exerc. 1984;16(4):355-359.
133. McConnell TJ, Horvat MA, Beutel-Horvat TA, Golding LA. Arm crank versus wheelchair treadmill ergometry to evaluate the performance of paraplegics. Paraplegia. 1989;27(4):307-313. doi:10.1038/sc.1989.46
134. Pitetti KH, Snell PG, Stray-Gundersen J. Maximal response of wheelchair-confined subjects to four types of arm exercise. Arch Phys Med Rehabil. 1987;68(1):10-13.
135. McGawley K. The Reliability and Validity of a Four-Minute Running Time-Trial in Assessing [Formula: see text]max and Performance. Front Physiol. 2017;8:270. doi:10.3389/fphys.2017.00270
136. Gerbino A, Ward SA, Whipp BJ. Effects of prior exercise on pulmonary gas-exchange kinetics during high-intensity exercise in humans. J Appl Physiol (1985). 1996;80(1):99-107. doi:10.1152/jappl.1996.80.1.99
137. Castro RR, Pedrosa S, Chabalgoity F, Sousa EB, Nobrega AC. The influence of a fast ramp rate on peak cardiopulmonary parameters during arm crank ergometry. Clin Physiol Funct Imaging. 2010;30(6):420-425. doi:10.1111/j.1475-097X.2010.00958.x
138. Hutchinson MJ, Paulson TAW, Eston R, Goosey-Tolfrey VL. Assessment of peak oxygen uptake during handcycling: Test-retest reliability and comparison of a ramp-incremented and perceptually-regulated exercise test. PLoS One. 2017;12(7):e0181008. doi:10.1371/journal.pone.0181008
139. Atkinson G, Nevill AM. Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. Sports Med. 1998;26(4):217-238.
140. Tropp H, Samuelsson K, Jorfeldt L. Power output for wheelchair driving on a treadmill compared with arm crank ergometry. Br J Sports Med. 1997;31(1):41-44.
141. Danielsen J, Sandbakk O, McGhie D, Ettema G. The effect of exercise intensity on joint power and dynamics in ergometer double-poling performed by cross-country skiers. Hum Mov Sci. 2018;57:83-93. doi:10.1016/j.humov.2017.11.010
142. 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. doi:10.1152/jappl.1983.55.2.479
143. Beneke R, Leithäuser R, Hütler M. Dependence of the maximal lactate steady state on the motor pattern of exercise. Br J Sports Med. 2001;35(3):192-196.
144. Rothman KJ. No adjustments are needed for multiple comparisons. Epidemiology. 1990;1(1):43-46.
145. Pires FO, Hammond J, Lima-Silva AE, Bertuzzi RC, Kiss MA. Ventilation behavior during upper-body incremental exercise. J Strength Cond Res. 2011;25(1):225-230. doi:10.1519/JSC.0b013e3181b2b895
146. English Oxford Dictionaries 2017 [Available from: https://en.oxforddictionaries.com/definition/break_point.
147. Hopker JG, Jobson SA, Pandit J. Controversies in the physiological basis of the 'anaerobic threshold'and their implications for clinical cardiopulmonary exercise testing. Anaesthesia. 2011;66(2):111-123.
148. Martin TW, Zeballos RJ, Weisman IM. Gas exchange during maximal upper extremity exercise. Chest. 1991;99(2):420-425.

149. Orr JL, Williamson P, Anderson W, Ross R, McCafferty S, Fettes P. Cardiopulmonary exercise testing: arm crank vs cycle ergometry. Anaesthesia. 2013;68(5):497-501. doi:10.1111/anae.12195

150. Vinet A, Le Gallais D, Bernard PL, Poulain M, Varray A, Mercier J, et al. Aerobic metabolism and cardioventilatory responses in paraplegic athletes during an incremental wheelchair exercise. Eur J Appl Physiol Occup Physiol. 1997;76(5):455-461. doi:10.1007/s004210050275

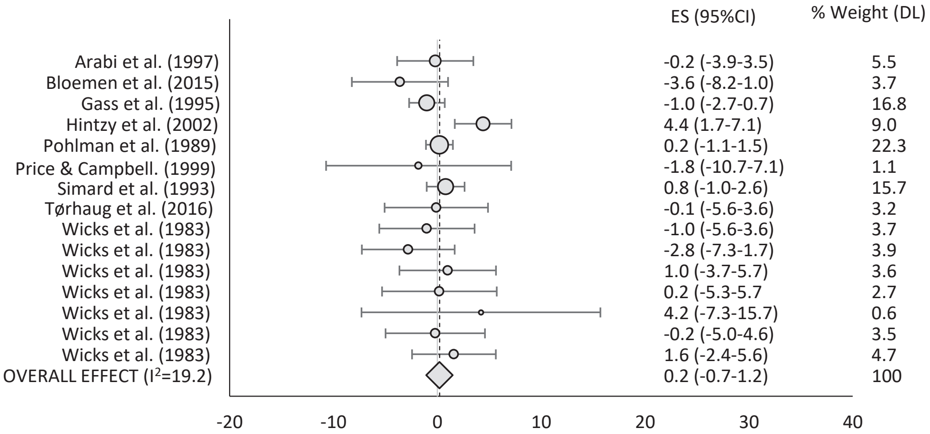
151. Esteve-Lanao J, San Juan AF, Earnest CP, Foster C, Lucia A. How do endurance runners actually train? Relationship with competition performance. Med Sci Sports Exerc. 2005;37(3):496-504.

152. Seiler KS, Kjerland GO. Quantifying training intensity distribution in elite endurance athletes: is there evidence for an "optimal" distribution? Scand J Med Sci Sports. 2006;16(1):49-56. doi:10.1111/j.1600-0838.2004.00418.x

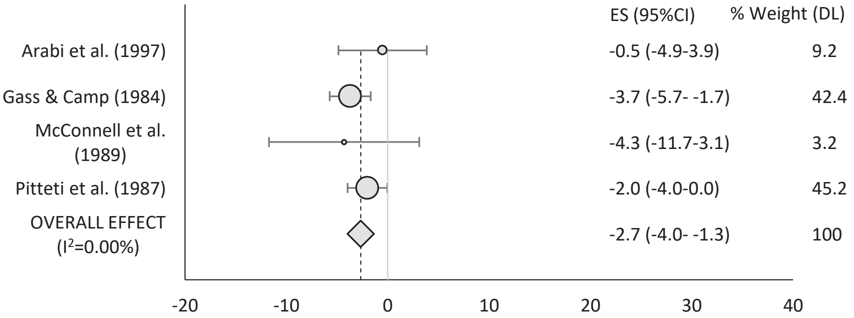
153. Faude O, Kindermann W, Meyer T. Lactate threshold concepts: how valid are they? Sports Med. 2009;39(6):469-490. doi:10.2165/00007256-200939060-00003

Appendix 1: Supplementary data 1 – VO_{2peak} in different upper-body exercise modes

Included in this preliminary systematic literature review were 12 studies, of which 9 studies[1-9] compared body-mass normalized VO_{2peak} values between the ACE and the WERG mode in 199 participants, and 4 studies[1, 10-12] between the WERG and the wheelchair treadmill mode in 42 participants. There was no significant difference in body-mass normalized VO_{2peak} between testing in the ACE and WERG mode (overall effect ± 95% CI: 0.2 ± 0.9, $p>0.05$) (Supplementary Figure A). However, testing in the ACE mode resulted in significantly lower body-mass normalized VO_{2peak} values compared to the wheelchair treadmill mode (-2.7 ± 1.3, $p<0.05$) (Supplementary Figure B). Only one study compared body-mass normalized VO_{2peak} between the WERG to the wheelchair treadmill mode in 13 participants and found no differences[1]. No studies have yet compared UBP to neither ACE, WERG nor WTR. The results of the comparison of absolute VO_{2peak} values between testing in the ACE, WERG and the wheelchair treadmill mode are similar to results presented for the body-mass normalized VO_{2peak}, and are, therefore, not further presented in this dissertation.



Supplementary Figure A. Effect size (ES) (95% CI range) of the difference in body-mass normalized VO_{2peak} between participants being tested in an arm crank ergometry versus a wheelchair ergometry mode. The dot size indicates the relative weight of each study in determining the overall effect size.



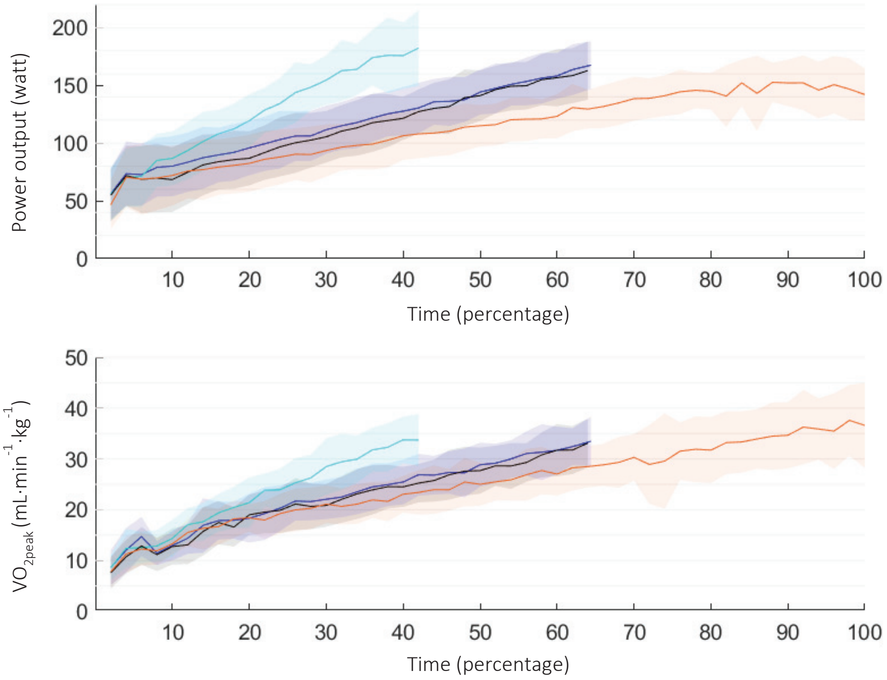
Supplementary Figure B. Effect size (ES) (95% CI range) of the difference in body-mass normalized VO_{2peak} between participants being tested in an arm crank ergometry versus a wheelchair treadmill mode. The dot size indicates the relative weight of each study in determining the overall effect size.

References

1. Arabi H, Vandewalle H, Pitor P, de Lattre J, Monod H. Relationship between maximal oxygen uptake on different ergometers, lean arm volume and strength in paraplegic subjects. Eur J Appl Physiol Occup Physiol. 1997;76(2):122-127. doi:10.1007/s004210050223
2. Bloemen MA, de Groot JF, Backx FJ, Westerveld RA, Takken T. Arm cranking versus wheelchair propulsion for testing aerobic fitness in children with spina bifida who are wheelchair dependent. J Rehabil Med. 2015;47(5):432-437. doi:10.2340/16501977-1944
3. Gass E, Harvey L, Gass G. Maximal physiological responses during arm cranking and treadmill wheelchair propulsion in T4–T6 paraplegic men. Spinal Cord. 1995;33(5):267-270.
4. Hintzy F, Tordi N, Perrey S. Muscular efficiency during arm cranking and wheelchair exercise: a comparison. Int J Sports Med. 2002;23(6):408-414. doi:10.1055/s-2002-33734
5. Pohlman RL, Gayle GW, Davis GM, Glaser RM. Metabolic responses to maximal arm crank and wheelchair ergometry in male paraplegics. Clinical Kinesiology. 1989;43(4):89-95.
6. Price MJ, Campbell IG. Thermoregulatory and physiological responses of wheelchair athletes to prolonged arm crank and wheelchair exercise. Int J Sports Med. 1999;20(7):457-463. doi:10.1055/s-1999-8831
7. Simard C, Noreau L, Paré G, Pomerleau P. Réponses physiologiques maximales lors d'un effort chez des sujets quadriplégiques. Can J Appl Physiol. 1993;18(2):163-174.
8. Torhaug T, Brurok B, Hoff J, Helgerud J, Leivseth G. Arm Crank and Wheelchair Ergometry Produce Similar Peak Oxygen Uptake but Different Work Economy Values in Individuals with Spinal Cord Injury. Biomed Res Int. 2016;2016:5481843. doi:10.1155/2016/5481843
9. Wicks JR, Oldridge NB, Cameron BJ, Jones NL. Arm cranking and wheelchair ergometry in elite spinal cord-injured athletes. Med Sci Sports Exerc. 1983;15(3):224.
10. Gass GC, Camp EM. The maximum physiological responses during incremental wheelchair and arm cranking exercise in male paraplegics. Med Sci Sports Exerc. 1984;16(4):355-359.
11. McConnell TJ, Horvat MA, Beutel-Horvat TA, Golding LA. Arm crank versus wheelchair treadmill ergometry to evaluate the performance of paraplegics. Paraplegia. 1989;27(4):307-313. doi:10.1038/sc.1989.46
12. Pitetti KH, Snell PG, Stray-Gundersen J. Maximal response of wheelchair-confined subjects to four types of arm exercise. Arch Phys Med Rehabil. 1987;68(1):10-13.

Appendix 2: Supplementary data 2 – Influence of increment duration and workload increases on VO_{2peak}

Peak PO was significantly higher in the short incremental test (20W/30s: 189 ± 30.3 watt) (all comparisons, $p < 0.001$) and significantly lower in the long incremental test (10W/min: 152 ± 21 watt) (all comparisons, $p < 0.001$) compared to the two incremental tests of moderate duration (10 watt/30s and 20 watt /min: 169 ± 26.7 and 175 ± 25 watt, respectively). Despite this, VO_{2peak} did not significantly differ between the four tests (20 watt/30 s: 36.3 ± 5.0, 10 watt/30 s: 37.0 ± 4.9, 20 watt/min: 37.2 ± 5.3, 10 watt/min: 38.2 ± 6.1 $mL·kg^{-1}·min^{-1}$, $p > 0.11$) (Supplementary Figure C). Furthermore, VO_2 at peak PO was significantly lower compared to VO_{2peak} in all four tests (20 watt/30 s: 32.7 ± 5.8, 10 watt/30 s: 35.2 ± 5.0, 20 watt/min: 35.3 ± 4.5, 10 watt/min: 37.2 ± 6.6 $mL·kg^{-1}·min^{-1}$, $p < 0.01$).



Supplementary Figure C. Power output and VO_2 kinetics during incremental test a (20 watt/30 s) (cyan), test b (10 watt/30 s) (black), test c (20 watt/1 min) (blue) and test d (10 watt/1 min) (orange). To be able to calculate the mean (solid lines), the individual data of each participant was divided into an equal number of steps for each of the four tests. The shaded areas are ± one standard deviation. This figure is duplicated in a manuscript that will be submitted for publication to Frontiers in Physiology[1].

References

1. Brurok B, Baumgart JK, Mellema M, Sandbakk O. The effect of increment rate and duration on the attainment of peak oxygen uptake during seated upper-body double poling (to be submitted to Front Physiol). 2018.

Paper I

PLOS ONE

RESEARCH ARTICLE

Peak oxygen uptake in Paralympic sitting sports: A systematic literature review, meta- and pooled-data analysis

Julia Kathrin Baumgart[1]*, **Berit Brurok**[1,2], **Øyvind Sandbakk**[1]

1 Centre for Elite Sports Research, Department of Neuromedicine and Movement Science, Faculty of Medicine and Health Science, Norwegian University of Science and Technology, Trondheim, Norway,
2 Department of Physical Medicine and Rehabilitation, St. Olav's University Hospital, Trondheim, Norway

* jk.baumgart@gmail.com

Citation: Baumgart JK, Brurok B, Sandbakk Ø (2018) Peak oxygen uptake in Paralympic sitting sports: A systematic literature review, meta- and pooled-data analysis. PLoS ONE 13(2): e0192903. https://doi.org/10.1371/journal.pone.0192903

Editor: Nicola Bragazzi, University of Genoa, School of Public Health, ITALY

Received: October 18, 2016

Accepted: January 12, 2018

Published: February 23, 2018

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: The review was funded by the Centre for Elite Sports Research. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

Abstract

Background

Peak oxygen uptake (VO_{2peak}) in Paralympic sitting sports athletes represents their maximal ability to deliver energy aerobically in an upper-body mode, with values being influenced by sex, disability-related physiological limitations, sport-specific demands, training status and how they are tested.

Objectives

To identify VO_{2peak} values in Paralympic sitting sports, examine between-sports differences and within-sports variations in VO_{2peak} and determine the influence of sex, age, body-mass, disability and test-mode on VO_{2peak}.

Design

Systematic literature review and meta-analysis.

Data sources

PubMed, CINAHL, SPORTDiscus^{TM} and EMBASE were systematically searched in October 2016 using relevant medical subject headings, keywords and a Boolean.

Eligibility criteria

Studies that assessed VO_{2peak} values in sitting sports athletes with a disability in a laboratory setting were included.

Data synthesis

Data was extracted and pooled in the different sports disciplines, weighted by the Dersimonian and Laird random effects approach. Quality of the included studies was assessed with a modified version of the Downs and Black checklist by two independent reviewers. Meta-regression and pooled-data multiple regression analyses were performed to assess the influence of sex, age, body-mass, disability, test mode and study quality on VO_{2peak}.

Results

Of 6542 retrieved articles, 57 studies reporting VO_{2peak} values in 14 different sitting sports were included in this review. VO_{2peak} values from 771 athletes were used in the data analysis, of which 30% participated in wheelchair basketball, 27% in wheelchair racing, 15% in wheelchair rugby and the remaining 28% in the 11 other disciplines. Fifty-six percent of the athletes had a spinal cord injury and 87% were men. Sports-discipline-averaged VO_{2peak} values ranged from 2.9 $L{\cdot}min^{-1}$ and 45.6 $mL{\cdot}kg^{-1}{\cdot}min^{-1}$ in Nordic sit skiing to 1.4 $L{\cdot}min^{-1}$ and 17.3 $mL{\cdot}kg^{-1}{\cdot}min^{-1}$ in shooting and 1.3 $L{\cdot}min^{-1}$ and 18.9 $mL{\cdot}kg^{-1}{\cdot}min^{-1}$ in wheelchair rugby. Large within-sports variation was found in sports with few included studies and corresponding low sample sizes. The meta-regression and pooled-data multiple regression analyses showed that being a man, having an amputation, not being tetraplegic, testing in a wheelchair ergometer and treadmill mode, were found to be favorable for high absolute and body-mass normalized VO_{2peak} values. Furthermore, high body mass was favourable for high absolute VO_{2peak} values and low body mass for high body-mass normalized VO_{2peak} values.

Conclusion

The highest VO_{2peak} values were found in Nordic sit skiing, an endurance sport with continuously high physical efforts, and the lowest values in shooting, a sport with low levels of displacement, and in wheelchair rugby where mainly athletes with tetraplegia compete. However, VO_{2peak} values need to be interpreted carefully in sports-disciplines with few included studies and large within-sports variation. Future studies should include detailed information on training status, sex, age, test mode, as well as the type and extent of disability in order to more precisely evaluate the effect of these factors on VO_{2peak}.

1. Introduction

The Paralympic Games are the world's second largest sporting event, and athletes with 10 different eligible physical impairments [1] participated in 23 summer disciplines in Rio 2016 and will participate in 6 winter disciplines in Pyoengchang 2018 (https://www.paralympic.org/sports). Of these, 16 of the summer sports and 5 of the winter sports disciplines have at least one sitting class. Depending on the eligibility criteria of each sitting sports discipline, athletes with impaired muscle power, impaired passive range of movement, limb deficiency, leg length difference, hypertonia, ataxia and athetosis are allowed to compete (https://www.paralympic.org/sports). Even though performance in all Paralympic sitting sports disciplines is mainly dependent on the work done by the upper body, the physical demands vary within a spectrum from typical endurance sports requiring high aerobic energy delivery over sustained periods to those performed with relatively low levels of displacement and corresponding low aerobic demands [2].

As an indicator of the humans' maximal ability to deliver energy aerobically, the measurement of maximal oxygen uptake (VO_{2max}) is regarded as the "gold standard" [3]. However, during exercise employing relatively low muscle mass, like in upper-body modes, the cardiorespiratory system is not fully taxed and VO_{2max} is rarely reached even in able-bodied participants [4, 5]. In such cases, peak oxygen uptake (VO_{2peak}) denotes the highest oxygen uptake reached during exercise to voluntary exhaustion [3] and is a common indicator of peak aerobic energy delivery capacity during upper-body exercise.

In sitting endurance sports with a continuously high physical effort, VO_{2peak} is suggested to be a paramount determinant of performance [6]. Whereas VO_{2max} values are available for elite athletes in a wide range of Olympic sports disciplines [7–10], only one study by Bhambhani et al. [11] provides a general overview of VO_{2peak} values in trained male wheelchair athletes. However, the latter study does not systematically report VO_{2peak} values for the individual Paralympic sitting sports disciplines. A systematic literature review on VO_{2peak} in Paralympic sports disciplines may, therefore, improve the scientific understanding of sport-specific aerobic demands, which is of importance for scientists as well as coaches and athletes. Furthermore, VO_{2peak} values of sitting sport athletes provide clinicians with a framework of what is possible to achieve in terms of peak aerobic capacity when exercising with a given modality and disability. This might be of relevance for providing feedback to their patients once they start engaging in a particular sitting sport activity.

In addition to the sport-specific demands, disability-related physiological limitations also influence VO_{2peak} in athletes with a disability. One study provided absolute VO_{2peak} in well-trained spinal cord injured (SCI) individuals (1.0–1.2 vs. 2.0–2.3 $L \cdot min^{-1}$ for tetraplegic (TETRA) vs. paraplegic (PARA), respectively) [12]. In the latter study, large differences in VO_{2peak} were found even within the well-trained individuals with different levels of SCI [12]. Whereas the focus in the few previous studies is on the influence of the different levels of SCI on VO_{2peak} [12, 13], there is lack of knowledge on how VO_{2peak} is influenced in Paralympic sitting sports athletes with other common disabilities, such as amputations, spina bifida and poliomyelitis. Furthermore, in the studies that focus on individuals with SCI, an inverse relationship between level of SCI and VO_{2peak} has been shown [14]. One may therefore expect high within-sports variation in VO_{2peak} in Paralympic sitting sports, since they include athletes with a large heterogeneity in disabilities.

Therefore, the purpose of this systematic literature review and meta-analysis was to (i) identify VO_{2peak} values for Paralympic sitting sports, (ii) examine between-sports differences and within-sports variations in VO_{2peak} and (iii) determine the influence of sex, age, body-mass, disability, test-mode and study-quality on VO_{2peak}. We hypothesized that VO_{2peak} values would be highest in Paralympic endurance sports with continuously high physical efforts over sustained periods. The lowest VO_{2peak} values were expected in sports with low levels of displacement and sports where athletes with large disability-related physiological limitations, such as athletes with tetraplegia, participate. Furthermore we expected that within-sports variation would be highest in sitting sports disciplines where athletes with a wide range of disabilities are included.

2. Methods

We conducted a systematic literature review and meta-analysis in accordance with the Preferred Reporting Items for Systematic Review and Meta-Analysis (PRISMA) guidelines [15]. Additionally, we registered the study protocol a priori in the International Prospective Register of Systematic Literature Reviews (PROSPERO) under the following registration number: CRD42015025134.

2.1 Eligibility criteria

Athletes with a physical disability above the age of 15, who were participating in sitting sports, were eligible for inclusion. An athlete was defined as a person who participates "[. . .] in an organized team or individual sport requiring systematic training and regular competition against others [. . .]"[16] at least on a national level. This rather broad definition may have resulted in the inclusion of some athletes that cannot be considered "elite". Athletes with a cognitive impairment were not included, since we would have not been able the separate the influence of the cognitive versus the physical disability on VO_{2peak}. Studies were included if absolute or body-mass normalized VO_{2peak} values were directly measured in a standardized laboratory setting. Studies that measured VO_{2peak} in a field setting were excluded due to lack of standardization. Only full-text, cross-sectional and intervention studies published in peer-reviewed journals in English, German or French were considered. Abstracts and conference proceedings were not eligible due to lack of detailed reporting of methods and results.

2.2 Data sources and search strategy

PubMed, CINAHL (through EBSCOhost), SPORTDiscus^{TM} (through EBSCOhost) and EMBASE were systematically and independently searched by JKB and BB in October 2016 using relevant medical subject headings, keywords and a Boolean search string. The search string combined synonyms and MeSH terms (the latter only relevant for our search in PubMed) of the two parts of the research question: peak oxygen uptake (outcome measure) and sitting athletes with a disability (population) (see S1 Fig). We decided to construct a broad search string to limit the potential of missing out on studies meeting our inclusion criteria. References of the included studies were searched manually and main research groups in the field were contacted for further identification of studies relevant to the research question.

2.3 Study selection

After eliminating duplicates articles, the titles were screened by JKB and BB. We only excluded titles that we were certain not to fit in the area of our review topic (e.g. the title being off topic, the title clearly stating that patients/able-bodied participants were investigated, etc.). Studies that did not directly mention VO_{2peak} in their title but were likely to have included it as a secondary outcome measure, were also included. In a second step, the abstracts of studies deemed relevant by title were read. Articles considered relevant by abstract, were then read in full-text. Details on the studies that were included or excluded based on abstract and full-text, and reasons for the excluded studies can be found in attachment S1 Excel file, sheet "study selection". All disagreements in the selection process were resolved by discussion between JKB and BB. The two reviewers were not blinded to the names of the authors of the included studies. If multiple studies from the same research group included the same data, only the first published study or the study with the most comprehensive information was included.

2.4 Data extraction

Data on the sports discipline competed in, the characteristics of the participants (number of participants, sex, age, body mass, type of disability and training status), test mode and peak oxygen uptake (absolute and body-mass normalized VO_{2peak} values) was extracted from the included studies by JKB with BB cross-checking all the data. Where necessary the unit of the training data was converted from minutes to hours and from miles to kilometers.

In the absence of a valid allometric scaling method that is generalizable to athletes with different disabilities [17], we chose to extract and report absolute and body-mass normalized VO_{2peak} values. When studies did not report absolute VO_{2peak} values ($L \cdot min^{-1}$), these were calculated by multiplying the individual body-mass normalized VO_{2peak} values (converted from mL to L) by the respective participants' body mass. When body-mass normalized VO_{2peak} values ($mL \cdot kg^{-1} \cdot min^{-1}$) values were not reported, these were calculated by dividing the individual absolute VO_{2peak} values (converted from L to mL) by the individual body mass (in kg^{-1}). When body-mass was not provided, this was calculated by dividing the individual absolute VO_{2peak} values (converted from L to mL) by the individual body-mass normalized VO_{2peak} values. In case of missing individual data, these calculations were not possible and data are not reported accordingly.

2.5 Assessment of methodological quality

The quality of the included studies was assessed by JKB and BB with a modified version of the Downs and Black checklist [18]. Modified versions of this checklist have been employed in several reviews in the field of sports science, which also mainly used cross-sectional studies for data retrieval [19–21]. The original checklist comprises 27 items, which are distributed over five sub-scales: reporting (item 1–10), external validity (item 11–13), bias (item 14–20), confounding (items 21–26) and power (item 27) [18]. For the purpose of the present review the following 12 items were included: 1–3, 5–7, 11, 12, 20–22 and 25. The other items were excluded since our review did not focus on interventions or differences between groups, where statistical considerations needed to be made and significance values or power would have been important. The term 'patient' was replaced by participant and 'treatment' was interpreted in the context of testing as described by Hebert-Losier et al [21]. The 'source population' was defined as all athletes with a disability within the respective sports discipline. All items, except item number 5, were rated as 'Yes' (1 point), 'No' (0 points) or 'Unknown' (0 points). For item 5, sex, age, weight, type of disability and training status were considered to be core confounders [17]. Test mode as well as the time of testing within the season were determined to be secondary confounders. Item 5 was scored with 2 points if all core confounders were mentioned. 1 point was scored if 4 out of the 5 core confounders and 1 secondary confounder were explained. 'No' or 'Unknown' were scored with 0, as described above. As we regarded the core confounders to be sufficiently assessed in item 5, we chose to in more detail address the determination criteria for VO_{2peak} in item 25. As no uniform criteria for the determination of maximal effort exist in a VO_{2peak} test in an upper-body mode, we defined our own minimum criteria. In accordance with Leicht et al. [22], these criteria should be viewed as a way to exclude studies in which maximal effort was clearly not reached rather than to confirm that VO_{2peak} was reached. In case studies 'Not applicable' (N/A) was added as a fourth option for items 7, 11, 12, 21 and 22; and items rated as such were excluded from the analysis. The modified version of the Downs and Black checklist used in this literature review can be found in the S1 Table. Quality cut-off points were decided on retrospectively and studies were ranked to be of low (0–5 points), moderate (6–8 points) or good (9–13 points) methodological quality. The level of evidence for each sports discipline was ranked from unknown to strong by combining the quality scores of each of the studies included in the respective discipline (see Table 1). The case studies were excluded from the analysis on level of evidence.

Table 1. Criteria for reporting methodological quality and consistency (adjusted from the criteria provided by van Tulder et al.[23]).

Level	Criteria
Strong	Data provided in multiple studies of good methodological quality
Moderate	Data provided in multiple studies of moderate methodological quality OR in one study of good methodological quality
Limited	Data provided in one study of moderate methodological quality
Very limited	Data provided in one study of low quality

https://doi.org/10.1371/journal.pone.0192903.t001

2.6 Statistics

All data are presented as means ± standard error (SE) and 95% confidence intervals (CI) unless specified otherwise. A meta-analysis, which is defined as "[. . .] the use of statistical techniques to integrate and summarize the results of included studies."[15], was performed by grouping together studies that determined VO_{2peak} in the same sports discipline. Sports discipline means were calculated in Microsoft Excel 2016 (Microsoft Cooperation, Washington, USA) by pooling study means by the random effects approach described more in detail by DerSimonian and Laird [24]. In connection to this, TETRA athletes were previously shown to display significantly lower VO_{2peak} values compared to athletes with other disabilities [25, 26]. Therefore, to lower the variation around the mean and to increase the sensitivity of the statistical tests, only the studies where it was possible to remove the VO_{2peak} data from TETRA athletes were included in the pooling procedure. The only exception was wheelchair rugby where all athletes included had TETRA and all studies in this sports discipline were pooled.

Between-sports differences were analyzed in Microsoft Excel by a one-way ANOVA with Tukey-Kramer Q tests to localize pair-wise differences based on study means and pooled study variances. An α level of 0.05 was employed to indicate statistical significance. To investigate the influence of each of the included studies on the VO_{2peak} values presented for the different sports disciplines, leave-one-out sensitivity analyses were performed in Stata 14.2 (StataCorp LLC, Texas, USA). Furthermore, cumulative meta-analyses were conducted to investigate possible VO_{2peak} changes as a function of time for each of the sport disciplines.

A meta-regression was performed in Stata 14.2 to investigate the relationship between absolute and body-mass normalized VO_{2peak} values, respectively, and the following 11 factors (levels of categorical factors are presented in brackets): age, body mass, percentage of men in each study (%Men), percentage of athletes with tetraplegia (%TETRA), paraplegia (%PARA), an amputation (%AMP), spina bifida (%SB), poliomyelitis (%PM) and athletes with other disabilities (%LA), test mode (arm crank ergometry (ACE), wheelchair ergometry (WERG) and wheelchair treadmill (treadmill) and study quality (moderate, good). Studies that provided information on all factors either as group or individual athlete data were used in the meta-regression. Because of too few studies with complete information, individual athlete data was included where the standard error was replaced by the standard deviation of all participants within each respective study. The levels "poling" and "handbiking" for the factor test mode and the level "low" for the factor study quality were excluded from the meta-regression. This is due to these levels providing only few data points for each factor. Baseline levels for dummy coding the two categorical factors test mode and study quality were "ACE" and "good", respectively. Only factors that significantly contributed to the model and decreased the Tau^2 estimate were included in the final meta-regression model. Before performing the meta-regression analyses, the variables were checked for multicollinearity.

A pooled-data multiple regression analysis was performed in IBM SPSS Statistics 24.0 (SPSS Inc., Chicago, USA) to investigate the relationship between absolute and body-mass normalized VO_{2peak} values, respectively, and the following six factors (levels of the categorical factors are presented in brackets): age, body mass, sex (male, female), disability (TETRA, PARA, amputation (AMP), spina bifida), test mode (ACE, WERG, treadmill) and study quality (low, moderate, good). Pooled data of studies that provided individual athlete data on all factors was used in the multiple regression analysis. Excluded from the regression analysis were the levels Les Autres and poliomyelitis for the factor disability, and poling and handcycling for the factor test mode. This is due to these levels comprising less than five percent of the data points of these two factors. Study quality was not entered in the multiple regression analysis as a factor due to too few data points with the level "low" and "good". Baseline levels for dummy coding the three categorical factors disability, test mode and study quality were "PARA", "ACE" and "good", respectively [27]. Only factors that significantly contributed to the model and increased the adjusted R^2 were included in the final regression model. Before performing the regression analyses, the data set was checked for outliers and multicollinearity, and each variable was tested for normality and homoscedasticity of residuals.

Sports discipline was not included in the meta-regression and multiple regression analyses due to multicollinearity with several of the other included factors. Furthermore, only data in the sports disciplines wheelchair basketball, wheelchair tennis, wheelchair racing and wheelchair rugby was included due to too few data points in other sports disciplines.

All figures and tables including information on VO_{2peak} values are arranged according to absolute VO_{2peak} values from highest to lowest values.

3. Results

3.1 Study selection and characteristics of included athletes

The systematic search resulted in 6542 studies. After removal of duplicate articles and the subsequent screening process, 57 full text studies were included. These 57 studies reported VO_{2peak} values in 771 athletes from 14 different Paralympic sitting sports disciplines (Fig 1). Athletics was divided into its two sub-disciplines, throwing disciplines and wheelchair racing due to the distinct differences in movement demands. No VO_{2peak} values were reported for wheelchair boccia, para-canoeing, para-equestrian, para-rowing, para-sailing, sitting volleyball, para-triathlon, and para-biathlon.

3.2 Methodological quality

Agreement on all assessed quality items was reached by JKB and BB. Four studies were ranked as having low and 6 studies as having good methodological quality (S2 Table). No quality label was attached to the 2 included case-studies. The remaining 45 studies were regarded to have moderate methodological quality. The quality of the studies that are included in each sports discipline determines the level of evidence of the VO_{2peak} values.

3.3 Between-sports differences

Mean absolute and body-mass normalized VO_{2peak} ± standard error (SE) of the sports disciplines ranged from 2.9 ± 0.3 $L \cdot min^{-1}$ and 45.6 ± 5.1 $ml \cdot kg^{-1} \cdot min^{-1}$ in Nordic sit skiing to 1.4 ± 0.2 $L \cdot min^{-1}$ and 17.3 ± 3.5 $ml \cdot kg^{-1} \cdot min^{-1}$ in shooting and 1.3 ± 0.1 and 18.9 ± 1.6 in wheelchair rugby. In Table 2 an overview of absolute and body-mass normalized VO_{2peak} values of all sports disciplines with more than one study with at least 3 participants is provided. In this overview, several factors, such as sex, age, body mass, type of disability, training status and test modes are grouped together. Table 3 and the regression analyses provide details on the influence of these factors on absolute and body-mass normalized VO_{2peak}. In the sports with a strong level of evidence and a large number of included studies (wheelchair basketball, wheelchair racing and wheelchair rugby), leave-one-out analyses, examining the effect of each of the included studies, did not have a great impact on neither absolute nor body-mass normalized VO_{2peak} values (S1 Excel file, sheet "MetaInf Output"). However, in sports with a low level of evidence and few included studies, omitting some of the studies had a larger impact on the VO_{2peak} values. With regards to the cumulative meta-analysis, wheelchair basketball and wheelchair racing showed a relatively stable VO_{2peak} over time, whereas wheelchair rugby showed a trend towards an increase in VO_{2peak} (S1 Excel file, sheet "MetaCum Output"). For all other sports, changes over time could not be investigated due to the few number of included studies.

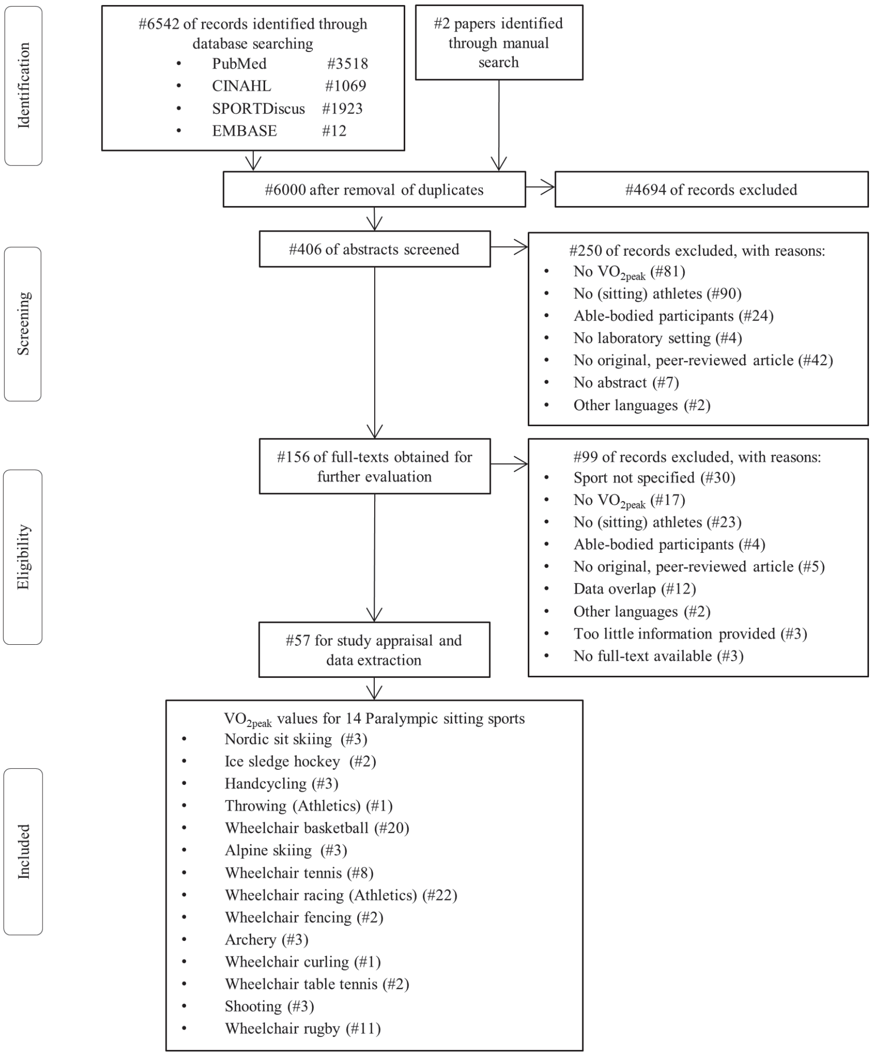


Fig 1. Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) flowchart depicting the study identification, screening, eligibility and inclusion process. The sports disciplines presented in the box at the bottom are ranked according to their absolute peak oxygen uptake (VO_{2peak}) values, from highest to lowest. * Note that 1) some of the studies provide values for more than one sports discipline and 2) athletics was divided into throwing events and wheelchair racing due to the distinct differences in movement demands between these two sub-disciplines.

https://doi.org/10.1371/journal.pone.0192903.g001

3.4 Within-sports variations

Within-sports variations in absolute and body-mass normalized VO_{2peak} values, based on CI ranges (Table 2), were relatively small in wheelchair basketball (0.4 $L \cdot min^{-1}$ and 7.2 $mL \cdot kg^{-1} \cdot min^{-1}$), wheelchair racing (0.6 $L \cdot min^{-1}$ and 7.4 $mL \cdot kg^{-1} \cdot min^{-1}$) and wheelchair rugby (0.4

Table 2. Overview of absolute and body-mass normalized peak oxygen uptake (VO_{2peak}) (mean ± SE [95% CI]) and level of evidence within the separate sitting sports disciplines. Sports disciplines are presented in order of absolute VO_{2peak} values, from high to low.

		Number of athletes	Absolute VO_{2peak} ± SE (L·min^{-1}) [95% CI]	Number of athletes	Body-mass normalized VO_{2peak} ± SE (mL·kg^{-1}·min^{-1}) [95% CI]	Level of evidence
1	Nordic sit skiing	24	2.9 ± 0.3 [2.2–3.5]$^{\text{WB, AS, WT, WRA, WF, WTT, SH, WRU}}$	24	45.6 ± 5.1 [35.6–55.6]$^{\text{HC, WB, AS, WT, WRA, WF, WTT, SH, WRU}}$	moderate
2	Para ice hockey	46	2.7 ± 0.3 [2.0–3.3]$^{\text{AS, WT, WRA, WF, WTT, SH, WRU}}$	-	-	limited
3	Hand cycling	30	2.6 ± 0.2 [2.2–3.1]$^{\text{AS, WT, WRA, WF, WTT, SH, WRU}}$	30	36.0 ± 4.3 [27.4–44.5]$^{\text{NS, WRA, WF, WTT, SH, WRU}}$	moderate
4	Wheelchair basketball	209	2.5 ± 0.1 [2.3–2.7]$^{\text{NS, AS, WT, WRA, WF, WTT, SH, WRU}}$	158	34.5 ± 1.8 [30.9–38.1]$^{\text{NS, WRA, WF, WTT, SH, WRU}}$	strong
5	Alpine sit skiing	21	2.3 ± 0.2 [1.9–2.7]$^{\text{NS, PIH, HC, WTT, SH, WRU}}$	21	33.1 ± 4.8 [23.6–42.5]$^{\text{NS, WRA, SH, WRU}}$	moderate
6	Wheelchair tennis	23	2.2 ± 0.2 [1.8–2.6]$^{\text{NS, PIH, HC, WTT, SH, WRU}}$	23	33.0 ± 2.3 [28.6–37.4]$^{\text{NS, WRA, SH, WRU}}$	strong
7	Wheelchair racing	179	2.2 ± 0.2 [1.9–2.5]$^{\text{NS, PIH, HC, WB, WTT, SH, WRU}}$	110	39.6 ± 1.9 [35.9–43.3]$^{\text{NS, HC, WB, AS, WT, WF, WTT, SH, WRU}}$	strong
8	Wheelchair fencing	10	2.2 ± 0.5 [1.2–3.1]$^{\text{NS, PIH, HC, SH, WRU}}$	10	31.0 ± 3.8 [23.4–38.6]$^{\text{NS, WRA, SH, WRU}}$	moderate
9	Wheelchair table tennis	7	1.8 ± 0.7 [0.5–3.1]$^{\text{NS, PIH, HC, WB}}$	7	29.2 ± 8.7 [12.0–46.3]$^{\text{NS, WRA, SH, WRU}}$	moderate
10	Shooting	8	1.4 ± 0.2 [1.0–1.9]$^{\text{NS, PIH, HC, WB, AS, WT, WF, WRA}}$	8	17.3 ± 3.5 [10.3–24.2]$^{\text{NS, HC, WB, AS, WT, WF, WTT}}$	moderate
11	Wheelchair rugby	114	1.3 ± 0.1 [1.1–1.5]$^{\text{NS, PIH, HC, WB, AS, WT, WF, WRA, WTT}}$	95	18.9 ± 1.6 [15.9–22.0]$^{\text{NS, HC, WB, AS, WT, WF, WTT}}$	strong/ moderate

Labels in superscript indicate significant differences to the respective sports discipline

The level of evidence with two attributes refers to absolute/body-mass normalized mean values, respectively. The results of the assessment of methodological quality need to be considered cautiously given the lack of empirical evidence that supports these. Note: several factors such as sex, age, body mass, disabilities, training status and test modes are grouped together in this overview table. Data of athletes with TETRA was excluded from the calculations of all sports discipline means except for wheelchair rugby.

https://doi.org/10.1371/journal.pone.0192903.t002

L·min^{-1} and 6.1 mL·kg^{-1}·min^{-1}), but above 0.6 L·min^{-1} and 7.5 mL·kg^{-1}·min^{-1} for the remaining sport disciplines. CI's for absolute and body-mass normalized VO_{2peak} values could not be reported for throwing, wheelchair curling and archery, and for body-mass normalized values in Para ice hockey, as only one study with a sample size of more than two athletes was included for each of these sports disciplines.

3.5 Meta-regression analyses

The meta-regression analyses, based on 35 studies that provided data of 26 sub-groups and 171 individual athletes in 4 different sports disciplines (wheelchair basketball, wheelchair racing, wheelchair tennis and wheelchair rugby), resulted in the following two equations as the best predictions of absolute (1) and body-mass normalized (2) VO_{2peak} values.

$$\begin{aligned}\text{Absolute VO}_{2\text{peak}} &= 0.93 + \text{body mass}_i \cdot 0.01 + \%\text{Men}_i \cdot 0.01 + \%\text{TETRA}_i \cdot -0.01 + \text{WERG}_i \cdot 0.29 \\ &\quad + \text{treadmill}_i \cdot 0.22\end{aligned} \tag{1}$$

The factors included in Eq (1) all significantly contribute to the model (all $p < 0.001$) and explain 77% of the variance in absolute VO_{2peak}. The coefficients presented in the model are

Table 3. Data extraction of number of male and female participants, absolute and body-normalized VO$_{2peak}$ values, age, body mass, type of disability, training status, exercise mode, and methodological quality of each of the studies included in this systematic literature review on peak aerobic capacity between and within Paralympic sitting sports. Mean age and body mass ± SE are presented of each sports discipline are presented in the grey lines.

Author and year of publication	Total number of athletes	Male athletes	Female athletes	Absolute VO$_{2peak}$ ± SD (L·min^{-1})	Body-mass normalized VO$_{2peak}$ ± SD (mL·kg^{-1}·min^{-1})	Age ± SD (grey lines: ± SE)	Body mass ± SD (grey lines: ± SE)	Disability	Training status	Test mode and protocol	Methodological quality
NORDIC SIT SKIING	24	23	1			41.2 ± 6.3	64.8 ± 5.2				
Bernardi et al. (2010) [2][‡]	5	5	0	3.3 ± 0.3	51.9 ± 6.9	39.6 ± 7.0	64.6 ± 4.8	3 PARA, 2 PM	*ns*	ACE (R)	good
Bernardi et al. (2012) [6]	16	16	0	2.9 ± 0.5	46 ± 9.8	41 ± 6.7	63.6 ± 6.3	3 AMP, 4 PM, 9 *ns*	*ns*	ACE (R)	low
Bhambhani et al. (2012) [28][‡]	3	2	1	2.3 ± 0.4	34.7 ± 9.3	44 ± 10.5	67.1 ± 8.8	3 PARA	*ns*	Poling (R)	moderate
PARA ICE HOCKEY	46	46	0			34.1 ± 6.0	75.9 ± 10.5				
Bernardi et al. (2012) [6]	34	34	0	2.5 ± 0.4	32.4 ± 6.1	38 ± 6.8	78 ± 11.4	20 AMP, 2 SB, 2 PM, 1 LA, 9 *ns*	*ns*	ACE (R)	low
Sandbakk et al. (2014) [29]	12	12	0	2.8 ± 0.3	-	28 ± 9.0	74.0 ± 10.0	12 *ns*	491 ± 112 hrs/year	Poling (R)	moderate
HANDCYCLING	30	20	2			41.3 ± 3.8	70.2 ± 3.7				
Fischer et al. (2014) [30][‡]	12	6	1	2.2 ± 0.6	31.7 ± 8.2	42.4 ± 5.1	68.1 ± 7.5	7 PARA	6.3 ± 2.9 hrs/week	Handbike (R)	moderate
		4	1	2.1 ± 0.6	32 ± 7.1	42.8 ± 4.5	64.4 ± 5.8	5 PARA	6.6 ± 2.6 hrs/week		
Knechtle et al. (2004b) [31]	8	*ns*	*ns*	2.6 ± 0.4	37.5 ± 7.3	38.6 ± 5.9	71.4 ± 8.4	6 PARA, 2 AMP	*ns*	Handbike (S)	moderate
Lovell et al. (2012) [32][‡]	10	10	0	3.2 ± 0.4	40.4 ± 5.5	40.8 ± 7.6	80.3 ± 7.8	9 PARA, 1 SB	230 ± 57 km/week	ACE (R)	moderate
THROWING (Athletics)	4	4	0								
Gass & Camp (1979) [33]	4	4	0	2.6 ± 0.5	30.1 ± 4.2	-	85.5 ± 9.98	4 PARA	8 ± 4 hrs/week	Treadmill (S-I)	good
WHEELCHAIR BASKETBALL	234	198	36			29.0 ± 1.7	69.9 ± 3.0				
Bernardi et al. (2010) [2][‡]	13	13	0	2.7 ± 0.5	36.9 ± 3.7	30.8 ± 7.2	73.5 ± 9.3	7 PARA, 4 AMP, 2 PM	*ns*	ACE (R)	good
Bloxham et al. (2001) [34][‡,†]	6	6	0	2.6 ± 0.6	37.6 ± 6.7	26 ± 5.9	69.1 ± 9.5	3 AMP, 3 SB	*ns*	WERG (S)	low
Coutts et al. (1990) [35][‡,†]	3	3	0	2.6 ± 0.4	34.6 ± 3.9	32 ± 9.5	74.5 ± 15.0	2 PARA, 1 PM	*ns*	WERG (R)	low
Croft et al. (2010) [36][‡,†]	6	4	2	3.0 ± 0.9	39.8 ± 5.4	26.7 ± 5.5	74.1 ± 18.2	3 PARA, 1 SB, 2 LA	15.8 ± 3.7 hrs/week	Treadmill (R)	moderate
de Lira et al. (2010) [37][‡]	17	17	0	1.9 ± 0.4	30.8 ± 6.1	25.4 ± 4.4	63.9 ± 15.4	7 PARA, 2 AMP, 8 PM	*ns*	Treadmill (S)	moderate
Dwyer & Davis (1997) [38]	13	0	13	1.7 ± 0.4	26.8 ± 5.3	26 ± 6.0	62.5 ± 9.5	13 *ns*	*ns*	ACE (R)	low
Goosey-Tolfrey & Tolfrey. (2004) [39][‡,†]	1	0	1	1.6	-	22.0	60.0	1 PARA	*ns*	WERG (S)	moderate
Goosey-Tolfrey et al. (2005) [40][‡]	12	12	0	2.8 ± 0.5	-	32.3 ± 4.6	74.7 ± 14.4	7 PARA, 1 AMP, 2 SB, 2 PM	20 hrs/week	WERG (S)	good
Goosey-Tolfrey & Tolfrey (2008) [41]	24	2	0	2.2 ± 0.2	-	35 ± 1.0	75.8 ± 14.9	2 *ns*	*ns*	WERG (S)	moderate
	-	11	0	2.5 ± 0.2	-	28 ± 5.0	71 ± 8.7	11 *ns*	*ns*	WERG (S)	
	-	4	0	2.3 ± 0.1	-	32 ± 3.0	70.7 ± 5.8	4 *ns*	*ns*	WERG (S)	
	-	7	0	3.3 ± 0.3	-	28 ± 7.0	79.2 ± 10.0	7 *ns*	*ns*	WERG (S)	
Goosey-Tolfrey et al. (2014) [42]	17	9	0	2.7 ± 0.5	-	29 ± 9.0	70.3 ± 12.6	9 *ns*	14.9 ± 1 hrs/week	Treadmill (I)	good

(*Continued*)

Table 3. (Continued)

Author and year of publication	Total number of athletes	Male athletes	Female athletes	Absolute VO$_{2peak}$ ± SD (L·min^{-1})	Body-mass normalized VO$_{2peak}$ ± SD (mL·kg^{-1}·min^{-1})	Age ± SD (grey lines: ± SE)	Body mass ± SD (grey lines: ± SE)	Disability	Training status	Test mode and protocol	Methodological quality
	-	8	0	3.8 ± 0.3	-	27 ± 8.0	84.8 ± 10.7	6 AMP, 2 LA	14 ± 3 hrs/ week	Treadmill (I)	
Griggs et al. (2015) [43][‡]	8	7	1	1.9 ± 0.5	-	27.8 ± 6.2	67.7 ± 13.1	8 PARA	16 ± 2 hrs/ week	Treadmill (S)	low
Knechtle & Knopfli. (2001) [44][P,‡,†]	10	10	0	2.5 ± 0.4	35.4 ± 4.5	29.4 ± 6.3	72.8 ± 16.9	7 PARA, 1 AMP, 1 PM, 1 LA	*ns*	Treadmill (I)	moderate
	1	1	0	2.8	38.3	21	84	1 TETRA	*ns*	Treadmill (I)	
Leicht et al. (2012) [45][‡]	9	9	0	2.5 ± 0.3	34.9 ± 5.1	30.6 ± 9.0	71.9 ± 12.6	9 PARA	11.6 ± 4.1 hrs/week	Treadmill (I)	good
Leicht et al. (2014) [46]	9	8	1	2.1 ± 0.5	32.8 ± 10.3	26.2 ± 5.6	64.1 ± 10.4	8 PARA, 1 LA	10.6 ± 5.5 hrs/week	Treadmill (S)	moderate
Rotstein et al. (1994) [47][‡,†]	8	8	0	2.0 ± 0.7	26.3 ± 7.5	31.3 ± 9.5	76.1 ± 20.4	4 PARA, 2 AMP, 1 PM, 1 LA	*ns*	ACE (R) /Treadmill (I)	moderate
Schmid et al. (1998) [48]	13	0	13	-	33.7 ± 5.2	27.8 ± 5.6	56.5 ± 6.8	9 PARA, 4 *ns*	7.6 ± 2.1 hrs/week	WERG (R)	moderate
vd Woude et al. (2002) [26]	5	0	5	1.5 ± 0.7	-	30.8 ± 6.3	67.6 ± 18.4	5 LA	8.4 ± 5.5 hrs/week	WERG (R)	moderate
Vanlandewijk et al. (1994) [49][‡]	40	13	0	1.9 ± 0.5	29.7 ± 8.6	29.6 ± 4.8	65.5 ± 12.6	12 PARA, 1 PM	4.5 ± 1.7 hrs/week	Treadmill (S)	moderate
	-	14	0	2.4 ± 0.4	36.3 ± 9.3	32.9 ± 8.4	70.7 ± 12.4	8 PARA, 1 SB, 5 PM	6.4 ± 3.4 hrs/week	Treadmill (S)	
	-	13	0	2.6 ± 0.3	37.9 ± 5.2	32.8 ± 7.2	67.9 ± 12.2	2 PARA, 3 AMP, 1 SB, 7 PM	5.5 ± 1.7 hrs/week	Treadmill (S)	
Veeger et al. (1991) [50]	11	11	0	2.7 ± 0.6	37.9 ± 6.9	29 ± 3.5	72 ± 9.4	11 *ns*	*ns*	Treadmill (S-I)	moderate
Zacharakis et al. (2012) [51][E,‡]	8	8	0	1.7 ± 0.1	-	31.4 ± 8.4	72.8 ± 8.5	1 TETRA, 7 PARA	*ns*	WERG (R)	moderate
ALPINE SIT SKIING	23	21	2			32.2 ± 5.0	61.6 ± 7.3				
Bernardi et al. (2012) [2]	15	15	0	2.3 ± 0.4	31.3 ± 6.7	33.1 ± 4.2	75.9 ± 15.4	1 SB, 14 *ns*	*ns*	ACE (R)	low
Gass & Camp (1979) [33]	3	3	0	1.6 ± 0.5	30.6 ± 9.7	-	52.4 ± 5.3	2 PARA, 1 PM	3.5 ± 2.2 hrs/week	Treadmill (S-I)	good
Goll et al. (2015) [52]	5	0	2	1.8 ± 0.2	44.5 ± 4.9	18.5 ± 0.7	40 ± 0.0	2 *ns*	*ns*	ACE (R)	moderate
	-	3	0	2.4 ± 0.2	35 ± 3.6	31 ± 5.9	69.0 ± 10.0	3 *ns*	*ns*	ACE (R)	
WHEELCHAIR TENNIS	36	29	7			30.0 ± 3.7	64.7 ± 4.9				
Bernardi et al. (2010) [2][‡]	4	4	0	2.3 ± 0.3	33.1 ± 2.9	38.5 ± 10.3	68.5 ± 8.4	4 PARA	*ns*	ACE (R)	good
Croft et al. (2010) [36][‡,†]	6	4	2	2.1 ± 0.7	31 ± 6.6	23 ± 8.2	65.8 ± 18.1	3 PARA, 3 LA	14.7 ± 7.8 hrs/week	Treadmill (R)	moderate
Diaper & Goosey-Tolfrey (2009) [53]	1	0	1	2.0	39.5	33.0	50.1	1 PARA	*ns*	WERG (S)	-
Goosey-Tolfrey & Tolfrey (2004) [39][‡,†]	3	0	3	1.7 ± 0.5	32.4 ± 7.0	28.7 ± 5.9	51.0 ± 8.4	3 PARA	*ns*	WERG (S)	moderate
Goosey-Tolfrey et al. (2006) [54][E,‡,†]	4	4	0	1.0 ± 0.3	14.9 ± 2.6	30 ± 4.3	68.3 ± 7.9	4 TETRA	*ns*	ACE (R)	moderate
Goosey-Tolfrey et al. (2008) [55][E,‡]	8	7	1	1.9 ± 0.7	-	27.2 ± 6.9	68.3 ± 17.9	2 TETRA, 3 PARA, 1 SB, 2 LA	*ns*	WERG (S)	moderate
Roy et al. (2006) [56][‡,†]	6	6	0	2.1 ± 0.5	27.5 ± 6.5	40.2 ± 9.8	77.5 ± 15.5	5 PARA, 1 AMP	8.7 ± 3.3 hrs/week	ACE (R)	moderate

(*Continued*)

Table 3. (Continued)

Author and year of publication	**Total number of athletes**	**Male athletes**	**Female athletes**	**Absolute VO_{2peak} ± SD ($L \cdot min^{-1}$)**	**Body-mass normalized VO_{2peak} ± SD ($mL \cdot kg^{-1} \cdot min^{-1}$)**	**Age ± SD (grey lines: ± SE)**	**Body mass ± SD (grey lines: ± SE)**	**Disability**	**Training status**	**Test mode and protocol**	**Methodological quality**
Vinet et al. (1996) [57]‡,†	4	4	0	2.4 ± 0.2	34.9 ± 1.8	28 ± 5.0	67.8 ± 5.7	4 PARA	4.8 ± 1.0 hrs/week	Treadmill (S)	moderate
WHEELCHAIR FENCING	11	10	1			31.9 ± 4.4	69.0 ± 8.6				
Bernardi et al. (2010) [2]‡	6	6	0	2.4 ± 0.7	34.4 ± 5.8	31.8 ± 5.4	68.3 ± 7.0	4 PARA, 1 AMP, 1 PM	*ns*	ACE (R)	good
Veeger et al. (1991) [50]	5	4	0	2.0 ± 0.4	29.2 ± 3.6	32 ± 3.3	70.0 ± 10.2	4 *ns*	*ns*	Treadmill (S-I)	moderate
	-	0	1	1.2	-	-	-	1 *ns*	*ns*	Treadmill (S-I)	
WHEELCHAIR RACING (athletics)	205	177	24			29.0 ± 1.4	61.4 ± 1.8				
Bernardi et al. (2010) [2]‡	6	6	0	3.1 ± 0.3	48.1 ± 6.4	30.2 ± 7.0	64.0 ± 7.2	5 PARA, 1 AMP	*ns*	ACE (R)	good
Bhambhani et al. (1995) [58]E,‡,†	8	8	0	1.4 ± 0.4	19.8 ± 4.3	31.8 ± 6.9	72.1 ± 6.9	8 TETRA	*ns*	WERG (S)	moderate
Campbell et al. (2004) [25]P,‡	20	3	0	1.3 ± 0.2	-	34 ± 8.0	67.5 ± 3.2	3 TETRA	5.4 hrs/week	Treadmill (I)	moderate
	-	8	0	2.1 ± 0.6	-	32 ± 6.0	67.8 ± 7.6	8 PARA	5.4 hrs/week	Treadmill (I)	
	-	9	0	2.2 ± 0.5	-	30 ± 8.0	62.8 ± 10.9	9 PARA	6.0 hrs/week	Treadmill (I)	
Cooper et al. (1992) [59]‡,†	11	11	0	2.6 ± 0.3	39.8 ± 4.2	30.9 ± 6.1	66.0 ± 6.4	10 PARA, 1 SB	7.9 ± hrs/ week	WERG (S/R)	moderate
Cooper et al. (1999) [60]P,‡,†	7	6	1	2.8 ± 0.8	41.0 ± 11.9	31.7 ± 4.9	68.8 ± 6.2	7 PARA	*ns*	ACE (R) /WERG (R)	moderate
	3	1	2	1.4 ± 0.6	22.4 ± 7.6	28.3 ± 2.5	61.2 ± 12.4	3 TETRA	*ns*	WERG (R)	
Coutts & Stogryn. (1987) [61]‡,†	4	4	0	2.7 ± 0.9	41 ± 9.9	26.8 ± 4.4	71.2 ± 16.7	3 PARA, 1 PM	*ns*	WERG (R)	moderate
	2	2	0	1.0 ± 0.02	17.1 ± 0.3	25.0 ± 1.4	59.4 ± 0.3	2 TETRA	*ns*		
Coutts et al. (1990) [35]‡,†	6	6	0	3.1 ± 0.5	52.7 ± 7.8	25.7 ± 4.0	58.5 ± 8.0	2 PARA, 3 AMP, 1 PM	*ns*	WERG (R)	low
Crews et al. (1982) [62]‡,†	4	4	0	2.2 ± 0.1	30.9 ± 7.1	28.8 ± 3.7	73.3 ± 3.7	3 PARA, 1 AMP	72.4 ± 33.8 km/week	Treadmill (S)	moderate
Gass et al. (1979) [33]P	4	4	0	2.3 ± 0.6	38.4 ± 9.5	-	61.3 ± 6.5	4 PARA	4.1 ± 1.8 hrs/week	Treadmill (S-I)	good
	1	1	0	1.1	19.4	-	54.6	1 TETRA	1.5 hrs/week		
Gass et al. (2002) [63]E,‡,†	4	4	0	1.1 ± 0.3	16.7 ± 3.5	38 ± 4.6	68.7 ± 12.0	4 TETRA	*ns*	Treadmill (S-I)	moderate
Goosey-Tolfrey & Campbell (1998) [64]‡,†	8	7	1	2.5 ± 0.5	37.8 ± 7.9	29.9 ± 8.0	68.0 ± 11.4	7 PARA, 1 SB	*ns*	Treadmill (I)	moderate
Goosey et al. (2000) [65]‡,†	8	8	0	2.6 ± 0.4	43.0 ± 10.6	26 ± 8.3	61.7 ± 11.6	3 PARA, 5 SB	*ns*	WERG (S)	moderate
Goosey-Tolfrey & Tolfrey (2004) [39]‡,†	5	0	5	1.8 ± 0.5	33.9 ± 3.8	29 ± 7.6	52.5 ± 14.5	1 PARA, 1 AMP, 3 SB	*ns*	WERG (S)	moderate
Hooker & Wells (1992) [66]‡	7	6	1	2.7 ± 0.5	43.1 ± 7.4	35 ± 6.1	61.6 ± 5.7	7 PARA	specified in article	ACE (R)	moderate
Knechtle et al. (2004a) [67]‡,†	8	6	2	2.5 ± 0.5	41.2 ± 6.5	34.8 ± 6.3	59.6 ± 5.5	5 PARA, 2 SB, 1 PM	*ns*	Treadmill (S)	moderate
	1	1	0	1.8	32.7	51.0	56.0	1 TETRA	*ns*		
Morris (1986) [68]	1	0	1	1.4	21.1	25.0	65.5	1 PARA	80.5 km/ week	ACE (R)	-
O'Connor et al. (1998) [69]‡,†	6	6	0	2.3 ± 0.2	36.2 ± 5.5	27.5 ± 4.9	64.1 ± 8.0	6 PARA	*ns*	WERG (S)	moderate

(*Continued*)

Table 3. (Continued)

Author and year of publication	Total number of athletes	Male athletes	Female athletes	Absolute VO$_{2peak}$ ± SD (L·min^{-1})	Body-mass normalized VO$_{2peak}$ ± SD (mL·kg^{-1}·min^{-1})	Age ± SD (grey lines: ± SE)	Body mass ± SD (grey lines: ± SE)	Disability	Training status	Test mode and protocol	Methodological quality
Perret et al. (2012) [70]‡	8	7	1	2.8 ± 0.7	-	32.8 ± 12.2	59.1 ± 11.0	6 PARA, 2 SB	*ns*	Treadmill (S)	moderate
Shiba et al. (2010) [71]	4	*ns*	*ns*	1.9 ± 0.4	36.9 ± 4.3	31.5 ± 9.0	52.0 ± 8.2	4 PARA	*ns*	WERG (I)	moderate
Tolfrey et al. (2001) [72]‡,†	16	16	0	2.4 ± 0.5	41.2 ± 9.2	28.1 ± 8.3	60.6 ± 11.0	8 PARA, 1 AMP, 6 SB, 1 PM	*ns*	WERG (S)	moderate
vd Woude et al. (2002) [26]P	48	3	0	0.7 ± 0.4	-	30.7 ± 8.3	61.7 ± 7.6	3 TETRA	11.3 ± 1.2 hrs/week	WERG (R)	moderate
	-	4	0	1.3 ± 0.3	-	29.5 ± 7.2	66.0 ± 9.3	4 *ns*	15.8 ± 7.2 hrs/week	WERG (R)	
	-	8	0	2.0 ± 0.3	-	31.4 ± 2.8	62.1 ± 8.8	8 *ns*	13.5 ± 3 hrs/week	WERG (R)	
	-	23	0	2.3 ± 0.4	-	27 ± 5.4	59.9 ± 11.8	23 *ns*	15.9 ± 7.3 hrs/week	WERG (R)	
	-	0	4	0.7 ± 0.2	-	29 ± 2.9	46.0 ± 4.1	4 *ns*	13.4 ± 3.3 hrs/week	WERG (R)	
	-	0	3	1.3 ± 0.1	-	26 ± 5.6	52.3 ± 10.1	3 *ns*	13.8 ± 6.8 hrs/week	WERG (R)	
	-	0	3	1.2 ± 0.2	-	23 ± 3.5	51.3 ± 8.1	3 *ns*	12.5 ± 2.5 hrs/week	WERG (R)	
Vinet et al. (1996) [57]‡,†	5	5	0	2.7 ± 0.6	43.6 ± 7.9	29.6 ± 4.0	63.2 ± 12.1	3 PARA, 1 SB, 1 PM	7.8 ± 3.5 hrs/week	Treadmill (S)	moderate
ARCHERY	8	7	1			-	-				
Cooper et al. (1999) [60]‡	4	4	0	1.9 ± 0.2	22.2 ± 2.1	44.0 ± 11.3	88.0 ± 12.9	3 PARA	*ns*	WERG (R)	moderate
	1	1	0	0.9	14.9	35	61.7	1 TETRA	*ns*		
Gass & Camp (1979) [33]	2	2	0	1.6 ± 0.3	38.3 ± 4.1	41.9 ± 2.5	-	2 PARA	11 ± 9.2 hrs/week	Treadmill (S-I)	good
Veeger et al. (1991) [50]	2	1	0	1.4	17.5	47.0	80.0	1 *ns*	*ns*	Treadmill (S-I)	moderate
	-	0	1	1.2	-	-	-	1 *ns*	*ns*	Treadmill (S-I)	
WHEELCHAIR CURLING						-	-				
Bernardi et al. (2012) [6]	10	10	0	1.8 ± 0.4	23.4 ± 7.6	42 ± 8.6	82.3 ± 29.3	1 PM, 1 LA, 8 *ns*	*ns*	ACE (R)	low
WHEELCHAIR TABLE TENNIS	8	6	2			30.1 ± 11.4	60.8 ± 8.9				
Cooper et al. (1999) [60]P,‡	3	1	2	1.7 ± 1.1	26.5 ± 11.5	26 ± 12.2	62.4 ± 12.0	3 PARA	*ns*	ACE (R) /WERG (R)	moderate
	1	1	0	0.96	12.2	38	78.69	1 TETRA	*ns*	WERG (R)	
Veeger et al. (1991) [50]	4	4	0	1.8 ± 0.2	30.7 ± 6.5	34 ± 11.9	60.0 ± 5.9	4 *ns*	*ns*	Treadmill (S-I)	moderate
SHOOTING	18	9	9			39.6 ± 5.5	84.8 ± 19.6				
Castle et al. (2013) [73]E,‡	5	3	2	1.2 ± 0.4	-	40.2 ± 1.8	69.7 ± 7.4	1 TETRA, 2 PARA, 1 SB, 1 PM	*ns*	ACE (R)	moderate
Cooper et al. (1999) [60]P,‡	4	2	2	1.5 ± 0.2	17.9 ± 2.7	44.8 ± 8.5	86.8 ± 20.9	4 PARA	*ns*	ACE (R) /WERG (R)	moderate
	1	0	1	0.72	7.6	52	94.7	1 TETRA	*ns*	WERG (R)	
Veeger et al. (1991) [50]	8	4	0	1.3 ± 0.3	16.3 ± 4.5	37 ± 5.1	83.0 ± 18.4	4 *ns*	*ns*	Treadmill (S-I)	moderate
	-	0	4	1.3	-	-	-	4 *ns*	*ns*	Treadmill (S-I)	

(*Continued*)

Table 3. (Continued)

Author and year of publication	Total number of athletes	Male athletes	Female athletes	Absolute VO_{2peak} ± SD ($L\cdot min^{-1}$)	Body-mass normalized VO_{2peak} ± SD ($mL\cdot kg^{-1}\cdot min^{-1}$)	Age ± SD (grey lines: ± SE)	Body mass ± SD (grey lines: ± SE)	Disability	Training status	Test mode and protocol	Methodological quality
WHEELCHAIR RUGBY	114	110	4			29.9 ± 1.7	70.4 ± 3.5				
Barfield et al. (2010) [74]†,‡	9	9	0	1.1 ± 0.4	16.0 ± 4.7	32.7 ± 7.8	68.2 ± 15.2	9 TETRA	11.4 ± 10.4 hrs/week	ACE (R)	moderate
Domaszewska et al. (2013) [75]‡	14	14	0	1.3 ± 0.3	17.8 ± 5.0	34.4	72.2	14 TETRA	*ns*	ACE (R)	moderate
Goosey-Tolfrey et al. (2006) [54]‡,†	4	4	0	0.9 ± 0.1	12.1 ± 1.8	28.8 ± 3.2	75.0 ± 13.4	4 TETRA	*ns*	ACE (R)	moderate
Goosey-Tolfrey et al. (2014) [42]‡	9	9	0	1.5 ± 0.4	-	30 ± 5.0	70.6 ± 10.1	9 TETRA	13 ± 3 hrs/ week	Treadmill (I)	good
Griggs et al. (2015) [43]‡	8	7	1	1.6 ± 0.4	-	27 ± 4.2	65.2 ± 4.4	8 TETRA	11 ± 6.4 hrs/ week	Treadmill (S)	low
Leicht et al. (2012) [45]‡	8	8	0	1.7 ± 0.4	24.5 ± 4.9	29.2 ± 3.8	67.9 ± 6.7	8 TETRA	13.6 ± 5.6 hrs/week	Treadmill (I)	good
Morgulec-Adamowicz et al. (2011) [76]‡	30	7	0	1.6 ± 0.4	21.1 ± 6.3	31 ± 9.0	75.7 ± 8.2	7 TETRA	4–6 hrs/ week	Treadmill (*ns*)	moderate
	-	9	0	1.8 ± 0.5	26.4 ± 6.1	31 ± 8.0	69.8 ± 12.4	9 TETRA		Treadmill (*ns*)	
	-	6	0	1.8 ± 0.5	25.6 ± 5.6	30 ± 5.0	72.5 ± 14.6	6 TETRA		Treadmill (*ns*)	
	-	8	0	2.4 ± 0.5	30.2 ± 7.2	32 ± 5.0	81.4 ± 7.8	8 TETRA		Treadmill (*ns*)	
Taylor et al. (2010) [77]	7	6	1	1.2 ± 0.5	16.9 ± 4.9	30.9 ± 5.1	70.1 ± 13.8	7 TETRA	11 ± 3 hrs/ week	ACE (R)	moderate
West et al. (2013) [78]‡,†	7	7	0	1.3 ± 0.3	19.4 ± 3.8	31.7 ± 4.1	69.3 ± 13.5	7 TETRA	*ns*	ACE (R)	moderate
West et al. (2014a) [79]‡	8	7	1	1.3 ± 0.3	19 ± 2.1	29 ± 2.0	67.0 ± 15.0	8 TETRA	15 hrs/week	Treadmill (I)	moderate
West et al. (2014b) [80]‡	10	4	1	1.1 ± 0.2	-	27.9 ± 6.2	62.2 ± 9.2	5 TETRA	20 hrs/week	ACE (R)	moderate
	-	5	0	1.1 ± 0.2	-	30.5 ± 5.0	73.2 ± 13.3	5 TETRA		ACE (R)	

[E] These studies are excluded in the calculation of sports disciplines means of overview Table 2, since they only provide group data which includes data of athletes with TETRA.

[P] The data provided in these studies was only partially included of the athletes that had disabilities other than TETRA.

‡ Studies with data that is used in the meta-regression analyses

† Studies with individual data that is used in the pooled-data multiple regression analyses

Abbreviations: *ns* not specified, *TETRA* tetraplegia, *PARA* paraplegia, *AMP* amputation, *PM* poliomyelitis, *SB* spina bifida, *LA* Les Autres, *ACE* arm crank ergometer, *WERG* wheelchair ergometer, *(S)* speed increments, *(I)* incline increments, *(S-I)* combination of speed and incline increments, *(R)* resistance increments

https://doi.org/10.1371/journal.pone.0192903.t003

unstandardized.

$$\begin{aligned}&\text{Body}-\text{mass normalized VO}_{2\text{peak}}\\&\quad=40.85+\text{body mass}_i\cdot-0.12+\%\text{TETRA}_i\cdot-0.16+\text{WERG}_i\cdot4.54+\text{treadmill}_i\cdot4.22\end{aligned}\quad(2)$$

The factors included in Eq (2) all significantly contribute to the model (all $p < 0.001$) and explain 82% of the variance in body-mass normalized VO_{2peak}. The coefficients presented in the model are unstandardized

3.6 Pooled-data multiple regression analyses

The multiple regression analyses, based on 22 studies which provided individual data of 169 athletes in 4 different sports disciplines (wheelchair basketball, wheelchair racing, wheelchair tennis and wheelchair rugby), resulted in the following two equations as the best predictions of absolute (3) and body-mass normalized (4) VO_{2peak} values.

$$\begin{aligned}&\text{Absolute } VO_{2peak}\\&\quad = 1.22 + \text{body mass}_i \cdot 0.02[0.25] + \text{female}_i \cdot -0.62[-0.25] + \text{TETRA}_i \cdot -1.09[-0.63]\\&\qquad + \text{AMI}_i \cdot -0.29[0.10] + \text{WERG}_i \cdot 0.36[0.24] + \text{treadmill}_i \cdot 0.32[0.20]\end{aligned}$$

$$(F_{6,162} = 52.52, p = 0.00) \tag{3}$$

The factors included in Eq (3) all significantly contribute to the model (all $p < 0.01$) and explain 65% of the variance in absolute VO_{2peak}. The coefficients presented in the model are unstandardized [and standardized].

$$\begin{aligned}&\text{Body} - \text{mass normalized } VO_{2peak}\\&\quad = 49.11 + \text{bodymass}_i \cdot -0.24[-0.26] + \text{female}_i \cdot -9.79[-0.24] + \text{TETRA}_i\\&\qquad \cdot -16.52[-0.62] + \text{AMI}_i \cdot 5.38[0.12] + \text{WERG}_i \cdot 5.71[0.27] + \text{treadmill}_i \cdot 4.54[0.18]\end{aligned}$$

$$(F_{6,162} = 52.50, p = 0.00) \tag{4}$$

The factors included in Eq (4) all significantly contribute to the model (all $p < 0.01$) and the model explain 65% of the variance in body-mass normalized VO_{2peak}. The coefficients presented in the model are unstandardized [and standardized].

4. Discussion

This systematic literature review aimed to (i) identify VO_{2peak} for Paralympic sitting sports, (ii) examine between-sports differences and within-sport variations in VO_{2peak} and iii) determine the influence of other factors on VO_{2peak}. The main finding is that VO_{2peak} values in general reflect the sport-specific demands and the type of disability of the athletes who compete in the respective sitting sports disciplines. VO_{2peak} values range from 2.9 $L\cdot min^{-1}$ and 45.6 $ml\cdot kg^{-1}\cdot min^{-1}$ in Nordic sit skiing, an endurance sport with a continuously high physical effort over sustained periods, to 1.4 $L\cdot min^{-1}$ and 17.3 $ml\cdot kg^{-1}\cdot min^{-1}$ in shooting, a sport with low levels of displacement, to 1.3 $L\cdot min^{-1}$ and 18.9 $ml\cdot kg^{-1}\cdot min^{-1}$ in wheelchair rugby, a sport that includes athletes with TETRA. Within-sports variation in absolute and body-mass normalized VO_{2peak} was relatively small in the sports with high sample sizes and a strong level of evidence, i.e. wheelchair basketball, wheelchair racing and wheelchair rugby, but above 0.5 $L\cdot min^{-1}$ and 8 $mL\cdot kg^{-1}\cdot min^{-1}$ in all other sports. Since the VO_{2peak} values presented for each of the sports disciplines include data of athletes that differ in their sex, age, body mass, type of disability, training status and the mode they were tested in, we additionally conducted regression analyses. These analyses show that being a man, having an amputation, not being tetraplegic, testing in a wheelchair ergometer and treadmill mode, were favorable for high absolute and body-mass normalized VO_{2peak} values. Furthermore, high body mass was favourable for high absolute VO_{2peak} values and low body mass for high body-mass normalized VO_{2peak} values.

In line with our hypothesis, Nordic sit skiing, an endurance sport with continuously high physical efforts, was the Paralympic sitting sport with the highest observed absolute and body-mass normalized VO_{2peak} values. Although endurance disciplines by nature require high aerobic energy delivery, VO_{2peak} values may be particularly high in Nordic sit skiers since they compete in varying terrain, which requires both high absolute VO_{2peak} values to accompany the relatively large upper-body muscle mass required to produce sufficient power on flat terrain, as well as high body-mass normalized VO_{2peak} values to carry their body mass up inclines. The same applies to their able-bodied counterparts, standing cross-country skiers, who have shown some of the highest VO_{2max} values among Olympic athletes [9, 81, 82], although VO_{2peak} values in elite Nordic sit skiers are lower due to less active muscle mass while being tested in an upper-body mode and the adverse influence of having a disability. For example, athletes with a SCI display lower VO_{2peak} values, which is mainly related to loss of motor- and sympathetic nervous system control below the level of injury. Depending on the level and extent of injury, a SCI is associated with a range of autonomic dysregulations, which amongst other things attenuates exercise performance [83]. In fact, an inverse relationship between the level of SCI and VO_{2peak} has been found [14]. There is, however, lack of knowledge in terms of the difference in VO_{2peak} between the different disabilities represented in Paralympic sitting sports. Hutzler et al.[84] examined the aerobic power of fifty well-trained individuals with lower limb impairments including SCI, polio and amputations during arm-cranking tests in a standardized laboratory setting. It was found that individuals with high and low SCI (above and below T5, respectively) displayed lower aerobic power compared to individuals with lower limb amputations [84], which may also reflect a difference in VO_{2peak} between the SCI and other types of disability [85].

Even though not significantly different from some of the other sports disciplines, we also observed relatively high absolute VO_{2peak} values for Para ice hockey. Although this sport is characterized by short, repeated sprints requiring maximal power and speed production, aerobic capacity was shown to be highly correlated to the maintenance of sprint ability [86]. Furthermore, the high absolute VO_{2peak} values in Para ice hockey players may also be related to their large amount of upper-body muscle mass, which is required to produce power in sport-specific situations. In addition, the lack of a classification system in Para ice hockey allows athletes to perform on a high international level only if they possess good trunk control and the influence of disability is minimal, such as in athletes with a low level SCI or a lower limb amputation. Accordingly, we would have expected a low within-sports variation in the VO_{2peak} values of Para ice hockey players who are a homogenous group of Paralympic athletes with respect to gender and disabilities. However, the low number of studies and participants included in the present review in this sports discipline resulted in a limited level of evidence and wider confidence intervals for both absolute and body-mass normalized VO_{2peak} as compared to sports with a larger number of studies and participants. Therefore, the values presented might not be representative for the population of Para ice hockey players and we need to be cautious in our interpretation of VO_{2peak} values in this sport.

Wheelchair racing and handcycling are endurance sports where athletes need to sustain power over longer periods and, therefore, display relatively high body-mass normalized as well as absolute VO_{2peak} values. In fact, wheelchair racers in the present meta-analysis were found to display the second highest body-mass normalized VO_{2peak} values, which further highlights the high aerobic demands in this sports. Even though there is some variation within wheelchair racing for both absolute and body-mass normalized VO_{2peak} values, the relatively large amount of studies and participants in this sport resulted in a strong level of evidence and in narrower confidence intervals as compared to other sports disciplines with fewer studies and participants. The variance that remains can partially be explained by the classification system that allows athletes with a broad spectrum of disabilities to compete against each other in separate classes. For example, in athletes with a SCI, VO_{2peak} may be lower compared to athletes with other disabilities due to lack of sympathetic control to the paralyzed trunk and lower limbs [83] and even lower in individuals with a complete SCI above T6, who additionally lack innervation to the splanchnic area [87]. Moreover, individuals with TETRA and autonomic completeness of the injury lack sympathetic innervation to the heart and display considerably lower heart rates than athletes with other disabilities [88]. We, therefore, decided to exclude studies that included athletes with TETRA from the overview table to limit the variability in VO_{2peak}. Since all these factors may influence VO_{2peak}, the extent to which disability affects VO_{2peak} largely differs between Paralympic athletes, and may explain at least part of the variation in Paralympic sitting disciplines with different disability classes as compared to disciplines without. However, very small sample sizes in each of the disability classes, lack of detailed reporting on disability classes in most of the studies, and a change in the division of classes over time, prevented us from investigating the effect of disability classes on variation in VO_{2peak}. Concluding from the above, differences in VO_{2peak} values between sports are fairly well reflected by the sport-specific demands and, therefore, highest in sports with continuously high physical efforts. However, they are also influenced by the heterogeneity in disabilities between athletes and the number of studies and athletes within each sports discipline, which in turn lead to differences in the magnitude of within-sport variations.

Shooting is a sport with low levels of displacement and consequently low aerobic demands. This was also reflected by the low VO_{2peak} values revealed in this sport. However, caution is warranted in the interpretation of the VO_{2peak} values in shooting despite a moderate level of evidence due to wide confidence intervals, as a result of the few studies with small sample sizes in this sports discipline. In contrast to shooting, the within-sports variation is lower and the level of evidence strong in wheelchair rugby, which increases our ability to interpret the VO_{2peak} values in this sport with more accuracy. In wheelchair rugby, a sport only eligible to athletes with impairments in both the upper and the lower limbs, the low VO_{2peak} values can be explained by the extent of the impairment. A study by West et al. [88] found that in athletes with TETRA physiological responses, including VO_{2peak}, were lower as compared to other disabilities and varied based on autonomic completeness of the SCI. However, the competitiveness has increased in wheelchair rugby and athletes are training more today than previously. This is likely reflected by the increase of VO_{2peak} over time in this sport. Maybe the care (e.g. catheterization) and the access to better-adapted training facilities (e.g. endurance training equipment such as lying handbikes with supportive handles) and modalities (e.g. electrostimulation while exercising) have improved more over the last years in tetraplegic athletes than in athletes with other disabilities. Overall, VO_{2peak} values are lowest in sports with low levels of displacement or sports which include athletes with TETRA. However, the certainty in the interpretation of these values depends on the level of evidence and the within-sports variation, which are dependent on the amount of studies and sample sizes included in each sports discipline.

The VO_{2peak} values presented for each of the sports disciplines include data of athletes that differ in their sex, age, body mass, type of disability, training status and the mode they were tested in. Therefore, the effect of VO_{2peak} was considered in the meta-regression and pooled-data multiple regression analyses. These analyses indicate that being a man, having an amputation, not being tetraplegic, testing in a WERG or treadmill mode, as well as having higher or lower body mass, respectively, is favorable for high absolute and body-mass normalized VO_{2peak}. The finding that being a man is beneficial for VO_{2peak} is also in line with previous studies [89, 90]. Tetraplegia may negatively influence VO_{2peak} due to a small amount of innervated muscle mass and a lack of autonomic innervation as previously discussed. In addition, that a higher body mass is beneficial for high absolute VO_{2peak} and lower body mass for high body-mass normalized VO_{2peak} was shown previously [91]. Furthermore, the finding that the WERG mode resulted in higher VO_{2peak} values compared to ACE is in line with two previous studies [92, 93], although several studies also report no differences between modes [94–97]. The reason for the clear differences between employing WERG or wheelchair treadmill testing as compared to ACE might be related to the former two modes being more sports-specific for the athletes included in the regression analyses, who all participated in wheelchair sports. Smaller coefficients for the WERG and the wheelchair treadmill mode might be expected if sitting athletes of the non-wheelchair sports are tested. The extent to which these coefficients would decrease is speculative though, as most of the latter athletes are likely using a wheelchair in at least some parts of daily life. In this context the influence of the test protocol on VO_{2peak} should also be taken into consideration. During ACE, an increased crank rate led to increases in test time and VO_{2peak} [98], whereas similar values were found in stepwise as compared to ramp type protocols [99]. Even though caution is required when drawing conclusions from the meta-regression and pooled-data multiple regression analysis, our findings provide a point of departure for understanding the influence of the above-mentioned factors on VO_{2peak} in Paralympic sitting sports athletes.

Methodological considerations

VO_{2peak} values were provided for only 14 out of 21 Paralympic sitting sports disciplines. This is partly due to new sports disciplines being added to the Paralympic games. For example, athletes competed in para-triathlon and para-canoeing for the first time in the Paralympic games in Rio 2016, and VO_{2peak} values in these disciplines are hence missing. So far, the only sports where a considerable number of VO_{2peak} values was provided with a strong level of evidence and we can hence conclude with more certainty are wheelchair basketball, wheelchair racing and wheelchair rugby with 234, 205 and 114 included athletes, respectively. Thus, more studies and bigger data pools established through international collaborations are required. Alternatively, systematically combining the results of multiple studies in a literature review and meta-analysis can compensate for the small sample sizes in original studies in the Paralympic field. However, to allow for more valid analyses, future studies are encouraged to provide sufficient detail on outcome measures in their abstracts and to provide individual, more detailed anthropometric and training data of their athletes. Furthermore, possible changes in the demands of the sports and improvements in performance and physiological capacity over the years should further be elucidated in future studies.

The studies included in the current review vary widely in terms of test equipment, such as ergospirometers and weighing scales, as well as the test mode, warm-up procedure and test protocol (stepwise increments in resistance, speed, incline or a combination of speed and incline). The effect of variations in these factors on upper-body VO_{2peak} in athletes with a disability remains unclear. To enable a more valid comparison of findings between studies, future studies should aim at providing enough details on the above mentioned factors and on finding standardized criteria for determination of VO_{2peak} in upper-body exercise modes.

In case of the present literature review, the consequences of publication bias are not only related to being able to publish data with significant findings and/or positive findings. It may also be related to the nature of elite sports where many countries test their best athletes without publishing this information. The reason may be two-fold, since giving away interesting information may help competitors and/or simply because publishing would be too resource demanding. Furthermore, data of Paralympic athletes might not be published because of a too low number of participants included to run statistical analyses on the data or due to the tested athletes not being considered elite, which was especially the case a few decades ago. Therefore, the average VO_{2peak} values presented here might not fully reflect the VO_{2peak} of medal-winning elite athletes in many of the sports. However, we are confident that in the sports with a strong level of evidence (wheelchair basketball, wheelchair racing and wheelchair rugby), the ranges provided for the VO_{2peak} values reflect the aerobic capacity of athletes in the respective sports. Although we limited the effect of publication bias by excluding articles with a complete overlap of data, we cannot exclude that duplicate data of individual athletes is included in this review. The likelihood of publication bias is illustrated by the fact that 15 of the articles included are from the research-group of Goosey-Tolfrey et al.[36, 39–43, 45, 46, 53–55, 64, 65, 72, 79]. Additionally, VO_{2peak} was a secondary measure in many of the reviewed studies. We, therefore, screened a large amount of abstracts of studies that did not directly mention VO_{2peak} in their title in order to reduce the possibility to miss articles that could have fit our inclusion criteria.

A limitation in the present review is that information on training status, which is known to influence VO_{2peak}, was missing in a considerable amount of studies.

In the absence of a valid allometric scaling method for athletes with different disabilities and across various sports, we provided only absolute and body-mass normalized values in this review. However, we refer to the studies of Batterham et al. [100] for the challenges surrounding the units in which VO_{2peak} is presented and to Goosey-Tolfrey et al. [17] for the only study on scaling of VO_{2peak} values in athletes with a disability.

5. Conclusion

In the current review, VO_{2peak} values for Paralympic sitting sports were systematically reported in 14 out of 21 possible sitting sports disciplines. Of these, VO_{2peak} was highest in the typical endurance sports and lowest in sports with low levels of displacement and in those including athletes with TETRA. However, the only sports where a sufficient number of VO_{2peak} values are combined with a strong level of evidence, thereby allowing us to conclude with more certainty, are wheelchair basketball, wheelchair racing and wheelchair rugby. In contrast, VO_{2peak} values should be interpreted carefully in disciplines with limited level of evidence or with only one study mean and in disciplines with large within-sports variations. Large within-sports variation was found in sports with few included studies and corresponding low sample sizes. The VO_{2peak} values presented for each of the sports disciplines include data of athletes that differ in their sex, age, body mass, type of disability, training status and the mode they were tested in. The influence of these factors on VO_{2peak} was investigated in regression analyses, which indicated that–in wheelchair basketball, wheelchair racing, wheelchair tennis and wheelchair rugby athletes–being a man, having an amputation, not being tetraplegic, testing in a wheelchair ergometry or treadmill mode, was beneficial for attaining high absolute or body-mass normalized VO_{2peak} values. Furthermore, high body mass was favourable for high absolute VO_{2peak} values and low body mass for high body-mass normalized VO_{2peak} values. In general, the practical applications of this review are limited due to most sports disciplines having large within-sports variations in VO_{2peak}, a limited level of evidence or including only one study mean. Based on the findings of this study and as a take-home message for future studies, we encourage the use of standardized determination criteria for reaching VO_{2peak}, and the inclusion of more detailed information on training status, sex, age, body mass, type of disability and testing mode, as well as larger study samples from international collaborations.

Supporting information

S1 Fig. Boolean search string. Boolean search string with combined synonyms and MeSH terms (the latter only for the search in PubMed), which was entered into the data bases.
(PDF)

S1 Table. Adjusted black and downs checklist.
(PDF)

S2 Table. Quality scores of the 57 included studies. Scoring of quality items for each of the studies included in this systematic literature review on peak oxygen uptake in Paralympic sitting sports.
(PDF)

S3 Table. PRISMA 2009 checklist.
(PDF)

S1 Excel file. Extracted and analyzed data of this systematic literature review on peak oxygen uptake in Paralympic sitting sports.
(XLSX)

Acknowledgments

The support of the staff at the medical and health library at the St.Olav's University hospital in checking our search strategy and providing us with missing articles, as well as the help of Øyvind Salvesen with the statistical analysis, is highly appreciated.

Author Contributions

Conceptualization: Julia Kathrin Baumgart, Berit Brurok, Øyvind Sandbakk.

Data curation: Julia Kathrin Baumgart, Berit Brurok, Øyvind Sandbakk.

Formal analysis: Julia Kathrin Baumgart, Berit Brurok, Øyvind Sandbakk.

Funding acquisition: Øyvind Sandbakk.

Investigation: Julia Kathrin Baumgart, Berit Brurok.

Methodology: Julia Kathrin Baumgart, Berit Brurok, Øyvind Sandbakk.

Project administration: Julia Kathrin Baumgart, Berit Brurok, Øyvind Sandbakk.

Supervision: Julia Kathrin Baumgart, Berit Brurok, Øyvind Sandbakk.

Validation: Julia Kathrin Baumgart, Berit Brurok, Øyvind Sandbakk.

Visualization: Julia Kathrin Baumgart, Berit Brurok, Øyvind Sandbakk.

Writing – original draft: Julia Kathrin Baumgart.

Writing – review & editing: Julia Kathrin Baumgart, Berit Brurok, Øyvind Sandbakk.

References

1. Tweedy SM, Vanlandewijck YC. International Paralympic Committee position stand—background and scientific principles of classification in Paralympic sport. Br J Sports Med. 2011; 45(4):259–69. https://doi.org/10.1136/bjsm.2009.065060 PMID: 19850575
2. Bernardi M, Guerra E, Di Giacinto B, Di Cesare A, Castellano V, Bhambhani Y. Field evaluation of paralympic athletes in selected sports: implications for training. Med Sci Sports Exerc. 2010; 42 (6):1200–8. https://doi.org/10.1249/MSS.0b013e3181c67d82 PMID: 19997027
3. Arena R, Myers J, Williams MA, Gulati M, Kligfield P, Balady GJ, et al. Assessment of functional capacity in clinical and research settings: a scientific statement from the American Heart Association Committee on Exercise, Rehabilitation, and Prevention of the Council on Clinical Cardiology and the Council on Cardiovascular Nursing. Circulation. 2007; 116(3):329–43. https://doi.org/10.1161/CIRCULATIONAHA.106.184461 PMID: 17576872

4. Hegge AM, Bucher E, Ettema G, Faude O, Holmberg HC, Sandbakk O. Gender differences in power production, energetic capacity and efficiency of elite cross-country skiers during whole-body, upper-body, and arm poling. Eur J Appl Physiol. 2015.
5. Stenberg J, Astrand PO, Ekblom B, Royce J, Saltin B. Hemodynamic response to work with different muscle groups, sitting and supine. J Appl Physiol. 1967; 22(1):61–70. https://doi.org/10.1152/jappl.1967.22.1.61 PMID: 6017655
6. Bernardi M, Carucci S, Faiola F, Egidi F, Marini C, Castellano V, et al. Physical Fitness Evaluation of Paralympic Winter Sports Sitting Athletes. . . [corrected] [published erratum appears in Clin J Sport Med 2012; 22(2):209]. Clin J Sport Med. 2012; 22(1):26–30. https://doi.org/10.1097/JSM.0b013e31824237b5 PMID: 22222593
7. Joyner MJ, Coyle EF. Endurance exercise performance: the physiology of champions. J Physiol. 2008; 586(1):35–44. https://doi.org/10.1113/jphysiol.2007.143834 PMID: 17901124
8. Saltin B, Astrand PO. Maximal oxygen uptake in athletes. J Appl Physiol. 1967; 23(3):353–8. https://doi.org/10.1152/jappl.1967.23.3.353 PMID: 6047957
9. Tonnessen E, Haugen TA, Hem E, Leirstein S, Seiler S. Maximal Aerobic Capacity in the Winter Olympic Endurance Disciplines: Olympic Medal Benchmarks for the Time Period 1990–2013. Int J Sports Physiol Perform. 2015.
10. Wilmore JH. The assessment of and variation in aerobic power in world class athletes as related to specific sports. Am J Sports Med. 1984; 12(2):120–7. https://doi.org/10.1177/036354658401200206 PMID: 6377935
11. Bhambhani Y. Physiology of wheelchair racing in athletes with spinal cord injury. Sports Med. 2002; 32 (1):23–51. PMID: 11772160
12. Janssen TW, Dallmeijer AJ, Veeger DJ, van der Woude LH. Normative values and determinants of physical capacity in individuals with spinal cord injury. J Rehabil Res Dev. 2002; 39(1):29–39. PMID: 11930906
13. West CR, Mills P, Krassioukov AV. Influence of the neurological level of spinal cord injury on cardiovascular outcomes in humans: a meta-analysis. Spinal Cord. 2012; 50(7):484–92. https://doi.org/10.1038/sc.2012.17 PMID: 22391687
14. Haisma JA, Bussmann JB, Stam HJ, Sluis TA, Bergen MP, Dallmeijer AJ, et al. Changes in physical capacity during and after inpatient rehabilitation in subjects with a spinal cord injury. Arch Phys Med Rehabil. 2006; 87(6):741–8. https://doi.org/10.1016/j.apmr.2006.02.032 PMID: 16731207
15. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gotzsche PC, Ioannidis JP, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. Ann Intern Med. 2009; 151(4):W65–94. PMID: 19622512
16. Maron BJ, Mitchell JH. Revised eligibility recommendations for competitive athletes with cardiovascular abnormalities. Med Sci Sports Exerc. 1994; 26(10 Suppl):S223–6.
17. Goosey-Tolfrey VL, Batterham AM, Tolfrey K. Scaling behavior of VO2peak in trained wheelchair athletes. Med Sci Sports Exerc. 2003; 35(12):2106–11. https://doi.org/10.1249/01.MSS.0000099106.33943.8C PMID: 14652509
18. Downs SH, Black N. The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. J Epidemiol Community Health. 1998; 52(6):377–84. PMID: 9764259
19. Cummins C, Orr R, O'Connor H, West C. Global positioning systems (GPS) and microtechnology sensors in team sports: a systematic review. Sports Med. 2013; 43(10):1025–42. https://doi.org/10.1007/s40279-013-0069-2 PMID: 23812857
20. Hall JP, Barton C, Jones PR, Morrissey D. The biomechanical differences between barefoot and shod distance running: a systematic review and preliminary meta-analysis. Sports Med. 2013; 43 (12):1335–53. https://doi.org/10.1007/s40279-013-0084-3 PMID: 23996137
21. Hebert-Losier K, Supej M, Holmberg HC. Biomechanical factors influencing the performance of elite alpine ski racers. Sports Med. 2014; 44(4):519–33. https://doi.org/10.1007/s40279-013-0132-z PMID: 24374655
22. Leicht CA, Tolfrey K, Lenton JP, Bishop NC, Goosey-Tolfrey VL. The verification phase and reliability of physiological parameters in peak testing of elite wheelchair athletes. Eur J Appl Physiol. 2013; 113 (2):337–45. https://doi.org/10.1007/s00421-012-2441-6 PMID: 22718268
23. van Tulder M, Furlan A, Bombardier C, Bouter L. Updated method guidelines for systematic reviews in the cochrane collaboration back review group. Spine (Phila Pa 1976). 2003; 28(12):1290–9.
24. DerSimonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials. 1986; 7(3):177–88. PMID: 3802833

25. Campbell IG, Williams C, Lakomy HK. Physiological and metabolic responses of wheelchair athletes in different racing classes to prolonged exercise. J Sports Sci. 2004; 22(5):449–56. https://doi.org/10.1080/02640410410001675298 PMID: 15160598

26. van der Woude LH, Bouten C, Veeger HE, Gwinn T. Aerobic work capacity in elite wheelchair athletes: a cross-sectional analysis. Am J Phys Med Rehabil. 2002; 81(4):261–71. PMID: 11953543

27. Field A. Discovering statistics using IBM SPSS statistics: Sage; 2013.

28. Bhambhani Y, Forbes S, Forbes J, Craven B, Matsuura C, Rodgers C. Physiologic responses of competitive Canadian cross-country skiers with disabilities. Clin J Sport Med. 2012; 22(1):31–8. https://doi.org/10.1097/JSM.0b013e3182432f0c PMID: 22222595

29. Sandbakk O, Spencer M, Ettema G, Bucher Sandbakk S, Skovereng K, Welde B. The physiology and biomechanics of upper-body repeated sprints in ice sledge hockey. Int J Sports Physiol Perform. 2014; 9(1):77–84. https://doi.org/10.1123/ijspp.2012-0355 PMID: 23628782

30. Fischer G, Tarperi C, George K, Ardigo LP. An exploratory study of respiratory muscle endurance training in high lesion level paraplegic handbike athletes. Clin J Sport Med. 2014; 24(1):69–75. https://doi.org/10.1097/JSM.0000000000000003 PMID: 24326928

31. Knechtle B, Muller G, Knecht H. Optimal exercise intensities for fat metabolism in handbike cycling and cycling. Spinal Cord. 2004; 42(10):564–72. https://doi.org/10.1038/sj.sc.3101612 PMID: 15289799

32. 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):3431–7. https://doi.org/10.1007/s00421-012-2324-x PMID: 22278391

33. Gass GC, Camp EM. Physiological characteristics of trained Australian paraplegic and tetraplegic subjects. Med Sci Sports. 1979; 11(3):256–9. PMID: 522635

34. Bloxham LA, Bell GJ, Bhambhani Y, Steadward RD. Time motion analysis and physiological profile of Canadian world cup wheelchair basketball players. Sports Med Train Rehab. 2001; 10(3):183–98.

35. Coutts KD. Peak Oxygen Uptake of Elite Wheelchair Athletes. Adapt Phys Activ Q. 1990; 7(1):62–6.

36. Croft L, Dybrus S, Lenton J, Goosey-Tolfrey V. A comparison of the physiological demands of wheelchair basketball and wheelchair tennis. Int J Sports Physiol Perform. 2010; 5(3):301–15. PMID: 20861521

37. de Lira CAB, Vancini RL, Minozzo FC, Sousa BS, Dubas JP, Andrade MS, et al. Relationship between aerobic and anaerobic parameters and functional classification in wheelchair basketball players. Scand J Med Sci Sports. 2010; 20(4):638–43. https://doi.org/10.1111/j.1600-0838.2009.00934.x PMID: 19793219

38. Dwyer GB, Davis RW. The relationship between a twelve minute wheelchair push test and VO2 peak in women wheelchair athletes. Sports Med Train Rehab. 1998; 8(1):1–11.

39. Goosey-Tolfrey VL, Tolfrey K. The oxygen uptake-heart rate relationship in trained female wheelchair athletes. J Rehabil Res Dev. 2004; 41(3b):415–20. PMID: 15543459

40. Goosey-Tolfrey VL. Physiological profiles of elite wheelchair basketball players in preparation for the 2000 Paralympic Games. Adapt Phys Activ Q. 2005; 22(1):57–66.

41. Goosey-Tolfrey VL, Tolfrey K. The multi-stage fitness test as a predictor of endurance fitness in wheelchair athletes. J Sports Sci. 2008; 26(5):511–7. https://doi.org/10.1080/02640410701624531 PMID: 18274948

42. Goosey-Tolfrey VL, Paulson TA, Tolfrey K, Eston RG. Prediction of peak oxygen uptake from differentiated ratings of perceived exertion during wheelchair propulsion in trained wheelchair sportspersons. Eur J Appl Physiol. 2014; 114(6):1251–8. https://doi.org/10.1007/s00421-014-2850-9 PMID: 24610244

43. Griggs KE, Leicht CA, Price MJ, Goosey-Tolfrey VL. Thermoregulation During Intermittent Exercise in Athletes With a Spinal-Cord Injury. Int J Sports Physiol Perform. 2015; 10(4):469–75. https://doi.org/10.1123/ijspp.2014-0361 PMID: 25365654

44. Knechtle B, Kopfli W. Treadmill exercise testing with increasing inclination as exercise protocol for wheelchair athletes. Spinal Cord. 2001; 39(12):633–6. https://doi.org/10.1038/sj.sc.3101229 PMID: 11781859

45. Leicht CA, Bishop NC, Goosey-Tolfrey VL. Submaximal exercise responses in tetraplegic, paraplegic and non spinal cord injured elite wheelchair athletes. Scand J Med Sci Sports. 2012; 22(6):729–36. https://doi.org/10.1111/j.1600-0838.2011.01328.x PMID: 21599755

46. 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):1635–43. https://doi.org/10.1007/s00421-014-2886-x PMID: 24781928

47. Rotstein A, Sagiv M, Ben-Sira D, Werber G, Hutzler J, Annenburg H. Aerobic capacity and anaerobic threshold of wheelchair basketball players. Paraplegia. 1994; 32(3):196–201. https://doi.org/10.1038/sc.1994.36 PMID: 8008425

48. Schmid A, Huonker M, Stober P, Barturen J, Schmidt-Trucksass A, Durr H, et al. Physical performance and cardiovascular and metabolic adaptation on elite female wheelchair basketball players in wheelchair ergometry and in competition. Am J Phys Med Rehabil. 1998; 77(6):527–61 12p. PMID: 9862541

49. Vanlandewijck YC, Spaepen AJ, Lysens RJ. Wheelchair propulsion: functional ability dependent factors in wheelchair basketball players. Scand J Rehabil Med. 1994; 26(1):37–48. PMID: 8023084

50. Veeger HE, Hadj Yahmed M, van der Woude LH, Charpentier P. Peak oxygen uptake and maximal power output of Olympic wheelchair-dependent athletes. Med Sci Sports Exerc. 1991; 23(10):1201–9. PMID: 1836828

51. Zacharakis ED, Kounalakis SN, Nassis GP, Geladas ND. Cardiovascular drift in trained paraplegic and able-bodied individuals during prolonged wheelchair exercise: effect of fluid replacement. Appl Physiol Nutr Metab. 2013; 38(4):375–81. https://doi.org/10.1139/apnm-2012-0131 PMID: 23713529

52. Goll M, Wiedemann MSF, Spitzenpfeil P. Metabolic Demand of Paralympic Alpine Skiing in Sit-Skiing Athletes. J Sports Sci Med. 2015; 14(4):819–24. PMID: 26664279

53. Diaper NJ, Goosey-Tolfrey VL. A physiological case study of a paralympic wheelchair tennis player: reflective practise. J Sports Sci Med. 2009; 8(2):300–7. PMID: 24149542

54. Goosey-Tolfrey V, Castle P, Webborn N. Aerobic capacity and peak power output of elite quadriplegic games players. . .including commentary by Abel T. Br J Sports Med. 2006; 40(8):684–7 4p. https://doi.org/10.1136/bjsm.2006.026815 PMID: 16611721

55. Goosey-Tolfrey V, Swainson M, Boyd C, Atkinson G, Tolfrey K. The effectiveness of hand cooling at reducing exercise-induced hyperthermia and improving distance-race performance in wheelchair and able-bodied athletes. Journal of applied physiology (Bethesda, Md: 1985). 2008; 105(1):37–43.

56. Roy JL, Menear KS, Schmid MM, Hunter GR, Malone LA. Physiological responses of skilled players during a competitive wheelchair tennis match. J Strength Cond Res. 2006; 20(3):665–71. PMID: 16977715

57. Vinet A, Bernard PL, Poulain M, Varray A, Le Gallais D, Micallef JP. Validation of an incremental field test for the direct assessment of peak oxygen uptake in wheelchair-dependent athletes. Spinal Cord. 1996; 34(5):288–93. PMID: 8963977

58. Bhambhani YN, Burnham RS, Wheeler GD, Eriksson P, Holland LJ, Steadward RD. Physiological correlates of simulated wheelchair racing in trained quadriplegics. Can J Appl Physiol. 1995; 20(1):65–77. PMID: 7742771

59. Cooper RA, Horvath SM, Bedi JF, Drechsler-Parks DM, Williams RE. Maximal exercise response of paraplegic wheelchair road racers. Paraplegia. 1992; 30(8):573–81. https://doi.org/10.1038/sc.1992.117 PMID: 1522999

60. Cooper RA, O'Connor TJ, Robertson RN, Langbein WE, Baldini FD. An investigation of the exercise capacity of the Wheelchair Sports USA Team. Assist Technol. 1999; 11(1):34–42 9p.

61. Coutts KD, Stogryn JL. Aerobic and anaerobic power of Canadian wheelchair track athletes. Med Sci Sports Exerc. 1987; 19(1):62–5. PMID: 3821457

62. Crews D, Wells CL, Burkett L, McKeeman, Hopkins V. Physiological profile of four wheelchair marathon racers. Phys Sportsmed. 1982; 10(6):134–43. https://doi.org/10.1080/00913847.1982.11947252 PMID: 29261060

63. Gass EM, Gass GC, Pitetti K. Thermoregulatory responses to exercise and warm water immersion in physically trained men with tetraplegia. Spinal Cord. 2002; 40(9):474–80. https://doi.org/10.1038/sj.sc.3101341 PMID: 12185609

64. Goosey VL, Campbell IG. Pushing economy and propulsion technique of wheelchair racers at three speeds. Adapt Phys Activ Q. 1998; 15(1):36–50 15p.

65. Goosey VL, Campbell IG, Fowler NE. Effect of push frequency on the economy of wheelchair racers. Med Sci Sports Exerc. 2000; 32(1):174–81. PMID: 10647546

66. Hooker SP, Wells CL. Aerobic power of competitive paraplegic road racers. Paraplegia. 1992; 30 (6):428–36. https://doi.org/10.1038/sc.1992.94 PMID: 1635793

67. Knechtle B, Muller G, Willmann F, Eser P, Knecht H. Fat oxidation at different intensities in wheelchair racing. Spinal Cord. 2004; 42(1):24–8. https://doi.org/10.1038/sj.sc.3101548 PMID: 14713940

68. Morris AF. A case study of a female ultramarathon wheelchair road user. Paraplegia. 1986; 24 (4):260–4. https://doi.org/10.1038/sc.1986.36 PMID: 3763241

69. O'Connor TJ, Robertson RN, Cooper RA. Three-dimensional kinematic analysis and physiologic assessment of racing wheelchair propulsion. Adapt Phys Activ Q. 1998; 15(1):1–14 p.

70. Perret C, Labruyere R, Mueller G, Strupler M. Correlation of heart rate at lactate minimum and maximal lactate steady state in wheelchair-racing athletes. Spinal Cord. 2012; 50(1):33–6. https://doi.org/10.1038/sc.2011.97 PMID: 21894166

71. Shiba S, Okawa H, Uenishi H, Koike Y, Yamauchi K, Asayama K, et al. Longitudinal changes in physical capacity over 20 years in athletes with spinal cord injury. Arch Phys Med Rehabil. 2010; 91 (8):1262–6. https://doi.org/10.1016/j.apmr.2010.04.024 PMID: 20684908

72. Tolfrey K, Goosey-Tolfrey VL, Campbell IG. Oxygen uptake-heart rate relationship in elite wheelchair racers. Eur J Appl Physiol. 2001; 86(2):174–8. https://doi.org/10.1007/s004210100493 PMID: 11822477

73. Castle PC, Kularatne BP, Brewer J, Mauger AR, Austen RA, Tuttle JA, et al. Partial heat acclimation of athletes with spinal cord lesion. Eur J Appl Physiol. 2013; 113(1):109–15. https://doi.org/10.1007/s00421-012-2417-6 PMID: 22592455

74. Barfield JP, Malone LA, Arbo C, Jung AP. Exercise intensity during wheelchair rugby training. J Sports Sci. 2010; 28(4):389–98. https://doi.org/10.1080/02640410903508839 PMID: 20131143

75. Domaszewska K, Szymczak Ł, KryŚCiak J, Pospieszna B, Laurentowska M. Spirometric and ergospirometric evaluation of wheelchair rugby players. Trends Sport Sci. 2013; 20(2):89–94.

76. Morgulec-Adamowicz N, Kosmol A, Molik B, Yilla AB, Laskin JJ. Aerobic, anaerobic, and skill performance with regard to classification in wheelchair rugby athletes. Res Q Exerc Sport. 2011; 82(1):61–9. https://doi.org/10.1080/02701367.2011.10599722 PMID: 21462686

77. Taylor BJ, West CR, Romer LM. No effect of arm-crank exercise on diaphragmatic fatigue or ventilatory constraint in Paralympic athletes with cervical spinal cord injury. J Appl Physiol (1985). 2010; 109 (2):358–66.

78. West CR, Romer LM, Krassioukov A. Autonomic function and exercise performance in elite athletes with cervical spinal cord injury. Med Sci Sports Exerc. 2013; 45(2):261–7. https://doi.org/10.1249/MSS.0b013e31826f5099 PMID: 22914247

79. West CR, Goosey-Tolfrey VL, Campbell IG, Romer LM. Effect of abdominal binding on respiratory mechanics during exercise in athletes with cervical spinal cord injury. J Appl Physiol (1985). 2014; 117 (1):36–45.

80. West CR, Taylor BJ, Campbell IG, Romer LM. Effects of inspiratory muscle training on exercise responses in Paralympic athletes with cervical spinal cord injury. Scand J Med Sci Sports. 2014; 24 (5):764–72. https://doi.org/10.1111/sms.12070 PMID: 23530708

81. Sandbakk O, Holmberg HC. A reappraisal of success factors for Olympic cross-country skiing. Int J Sports Physiol Perform. 2014; 9(1):117–21. https://doi.org/10.1123/ijspp.2013-0373 PMID: 24088346

82. Sandbakk O, Holmberg HC. Physiological Capacity and Training Routines of Elite Cross-Country Skiers: Approaching the Upper Limits of Human Endurance. Int J Sports Physiol Perform. 2017:1–26.

83. Theisen D. Cardiovascular determinants of exercise capacity in the Paralympic athlete with spinal cord injury. Exp Physiol. 2012; 97(3):319–24. https://doi.org/10.1113/expphysiol.2011.063016 PMID: 22090064

84. 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):205–12. PMID: 9554023

85. Andersson EA, Lundahl G, Wecke L, Lindblom I, Nilsson J. Maximal aerobic power versus performance in two aerobic endurance tests among young and old adults. Gerontology. 2011; 57(6):502–12. https://doi.org/10.1159/000329174 PMID: 21860214

86. Baumgart JK, Sandbakk O. Laboratory Determinants of Repeated-Sprint and Sport-Specific-Technique Ability in World-Class Ice Sledge Hockey Players. Int J Sports Physiol Perform. 2016; 11 (2):182–90. https://doi.org/10.1123/ijspp.2014-0516 PMID: 26182436

87. Thijssen DH, Steendijk S, Hopman MT. Blood redistribution during exercise in subjects with spinal cord injury and controls. Med Sci Sports Exerc. 2009; 41(6):1249–54. https://doi.org/10.1249/MSS.0b013e318196c902 PMID: 19461541

88. West CR, Gee CM, Voss C, Hubli M, Currie KD, Schmid J, et al. Cardiovascular control, autonomic function, and elite endurance performance in spinal cord injury. Scand J Med Sci Sports. 2015; 25 (4):476–85. https://doi.org/10.1111/sms.12308 PMID: 25175825

89. Kline GM, Porcari JP, Hintermeister R, Freedson PS, Ward A, McCarron RF, et al. Estimation of VO2max from a one-mile track walk, gender, age, and body weight. Med Sci Sports Exerc. 1987; 19 (3):253–9. PMID: 3600239

90. Zinner C, Sperlich B, Wahl P, Mester J. Classification of selected cardiopulmonary variables of elite athletes of different age, gender, and disciplines during incremental exercise testing. Springerplus. 2015; 4:544.

91. Bergh U, Sjodin B, Forsberg A, Svedenhag J. The relationship between body mass and oxygen uptake during running in humans. Med Sci Sports Exerc. 1991; 23(2):205–11. PMID: 2017016

92. Bloemen MA, de Groot JF, Backx FJ, Westerveld RA, Takken T. Arm cranking versus wheelchair propulsion for testing aerobic fitness in children with spina bifida who are wheelchair dependent. J Rehabil Med. 2015; 47(5):432–7. https://doi.org/10.2340/16501977-1944 PMID: 25882374

93. Sawka MN, Glaser RM, Wilde SW, von Luhrte TC. Metabolic and circulatory responses to wheelchair and arm crank exercise. J Appl Physiol. 1980; 49(5):784–8. https://doi.org/10.1152/jappl.1980.49.5.784 PMID: 6776077

94. Gass EM, Harvey LA, Gass GC. Maximal physiological responses during arm cranking and treadmill wheelchair propulsion in T4-T6 paraplegic men. Paraplegia. 1995; 33(5):267–70. https://doi.org/10.1038/sc.1995.60 PMID: 7630652

95. Gayle GW, Pohlman RL, Glaser RM, Davis GM. Cardiorespiratory and perceptual responses to arm crank and wheelchair exercise using various handrims in male paraplegics. Res Q Exerc Sport. 1990; 61(3):224–32. https://doi.org/10.1080/02701367.1990.10608683 PMID: 2097677

96. Glaser RM, Sawka MN, Brune MF, Wilde SW. Physiological responses to maximal effort wheelchair and arm crank ergometry. J Appl Physiol Respir Environ Exerc Physiol. 1980; 48(6):1060–4. https://doi.org/10.1152/jappl.1980.48.6.1060 PMID: 7380703

97. Martel G, Noreau L, Jobin J. Physiological responses to maximal exercise on arm cranking and wheelchair ergometer with paraplegics. Spinal Cord. 1991; 29(7):447–56.

98. Smith PM, Price MJ, Doherty M. The influence of crank rate on peak oxygen consumption during arm crank ergometry. J Sports Sci. 2001; 19(12):955–60. https://doi.org/10.1080/026404101317108453 PMID: 11820689

99. 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):616–21. https://doi.org/10.1055/s-2004-817880 PMID: 15532006

100. Batterham AM, Tolfrey K, George KP. Nevill's explanation of Kleiber's 0.75 mass exponent: an artifact of collinearity problems in least squares models? J Appl Physiol (1985). 1997; 82(2):693–7.

Paper II

Comparison of peak oxygen uptake and exercise efficiency between upper-body poling and arm crank ergometry in trained paraplegic and able-bodied participants

Julia Kathrin Baumgart (MSc)[1,*], Laura Gürtler (MSc)[1], Gertjan Ettema (PhD)[1], Øyvind Sandbakk (PhD)[1]

[1] Centre for Elite Sports Research, Department of Neuromedicine and Movement Science, Norwegian University of Science and Technology, Trondheim, Norway

* *Corresponding Author:*
Address for Correspondence: Smistadgrenda 11, 7026 Trondheim, Norway
Telephone Number: 48436206
Email address: jk.baumgart@gmail.com

ORCHID:
Julia Kathrin Baumgart: 0000-0001-5628-6050
Gertjan Ettema: 0000-0002-3370-0957
Øyvind Sandbakk: 0000-0002-9014-5152

Running Title: VO_{2peak} and efficiency in upper-body poling

Abstract

Purpose To compare peak oxygen uptake (VO_{2peak}) and exercise efficiency between upper-body poling (UBP) and arm crank ergometry (ACE) in able-bodied (AB) and paraplegic participants (PARA).

Methods Seven PARA and eleven AB upper-body trained participants performed four 5-min submaximal stages, and an incremental test to exhaustion in UBP and ACE. VO_{2peak} was the highest 30-s average during the incremental test. Metabolic rate (joule/second=watt) at fixed power outputs of 40, 60, and 80 W was estimated using linear regression analysis on the original power-output-metabolic-rate data and used to compare exercise efficiency between exercise modes and groups.

Results VO_{2peak} did not significantly differ between UBP and ACE (p=0.101), although peak power output was 19% lower in UBP (p<0.001). Metabolic rate at fixed power outputs was 24% higher in UBP compared to ACE (p<0.001), i.e. exercise efficiency was lower in UBP. PARA had 24% lower VO_{2peak} compared to AB (p=0.010), although there were no significant differences in peak power output between PARA and AB (p=0.209).

Conclusions In upper-body trained PARA and AB participants, VO_{2peak} did not differ between UBP and ACE, indicating that these two test modes tax the cardiovascular system similarly when the upper-body is restricted. As such, the 18% lower peak power output in UBP compared to ACE may be explained by the coinciding lower efficiency.

Key words: aerobic power; endurance; Paralympic; spinal cord injury; VO_{2max}; VO_{2peak}

Abbreviations

AB	Able-bodied participants
ACE	Arm crank ergometry
BLa	Blood lactate
HR	Heart rate
MR	Metabolic rate
PARA	Participants with a paraplegia
PO	Power output
RER	Respiratory exchange ratio
RPE	Overall rating of perceived exertion
UBP	Upper-body poling
VE	Minute ventilation
VO_2	Oxygen uptake
VO_{2peak}	Peak oxygen uptake

Introduction

Peak oxygen uptake (VO_{2peak}) and exercise efficiency are key factors for endurance performance. In persons who are primarily able to use their upper-body during exercise, such as many Paralympic athletes, the mode most commonly used in assessing VO_{2peak} and efficiency is arm crank ergometry (ACE) (Drory et al. 1990; Glaser et al. 1980; Mossberg et al. 1999; Price et al. 2011; Smith et al. 2006; Smith et al. 2004; Smith et al. 2007; Smith et al. 2001; Tropp et al. 1997). However, sport-specificity of the test mode has been suggested to be of importance for achieving VO_{2peak} and efficiency that are reflective of the endurance capacity in the respective sport (McCafferty and Horvath 1977). For Para ice hockey players, sitting Para cross-country skiers and Para biathletes, the upper-body poling (UBP) movement is the most sport-specific. However, it has not yet been investigated whether VO_{2peak} and efficiency differ between ACE and UBP and if possible differences are caused by the respective movement of the arms and/or due to different use of the trunk.

In ACE, power is produced by continuous, asynchronous force application, whereas in poling – similarly to the wheelchair ergometry – power is generated discontinuously, yet in synchronous movements of both hands (Sawka 1986). During ACE, the involvement of the trunk is limited by the asynchronous movement of the hands, whereas during UBP and wheelchair ergometry the synchronous movement of the hands allows more involvement of the trunk. A higher VO_{2peak} may, therefore, be expected in UBP and wheelchair ergometry compared to ACE due to using more muscle mass. However, despite the differences in arm movement and the engagement of the trunk between ACE and wheelchair ergometry, some studies show no differences in VO_{2peak} values between these two modes (Arabi et al. 1997; Gass et al. 1995; Gayle et al. 1990; Glaser et al. 1980; Martel et al. 1991; Price and Campbell 1999b), whereas others report higher values in the wheelchair ergometry mode (Bloemen et al. 2015; Gass and Camp 1984; Sawka et al. 1980). Furthermore, previous studies have shown that wheelchair ergometry is less efficient than ACE (Glaser et al. 1980; Hintzy et al. 2002; Mukherjee and Samanta 2001). This is likely caused by higher coordinative demands of using the discontinuous movement and by production of power during a shorter portion of each cycle in the wheelchair ergometry mode (Mukherjee and Samanta 2001).

Irrespective of the upper-body mode used during exercise testing, VO_{2peak} values were found to be consistently lower in paraplegic (PARA) compared to able-bodied participants (AB) (Hopman et al. 2004; Leicht and Perret

2008; Price and Campbell 1999a). Although the evidence is currently limited, efficiency in both ACE and wheelchair ergometry does not seem to differ between PARA and AB (Glaser et al. 1980).

In the current study, we aimed to compare VO_{2peak} and exercise efficiency between ACE and UBP with the upper-body restricted in both modes in PARA and AB. We hypothesized that VO_{2peak} values would be similar in ACE and UBP, yet lower in PARA as compared to AB. In accordance with the lower efficiency previously found in wheelchair ergometry, exercise efficiency was expected to be lower in UBP compared to ACE in the current study.

Methods

Participants

The PARA group consisted of seven (6 men, 1 women) upper-body-trained individuals with a paraplegia and the AB group of eleven (9 men, 2 women) healthy able-bodied upper-body-trained controls (anthropometrics and training hours are presented in Table 1). PARA were significantly older and had significantly lower leg lean muscle mass (LLM) compared to AB (both comparisons, $p < 0.004$). PARA consisted of an ice sledge hockey player, two hand-cyclists, a wheelchair curler, a wheelchair judoist and two recreationally trained participants. AB were sub-elite cross-country skiers who trained 11.5 ± 3.2 hours/week, with approximately half of this training spent in modes including the upper-body. Whereas the total number of training sessions was significantly higher in AB (7.2 ± 2.9 sessions/week, $p = 0.009$), training sessions including upper-body exercise did not differ between AB and PARA (4.5 ± 2.4 vs 4.0 ± 1.9 sessions/week, $p = 0.687$). The participants were instructed to refrain from heavy training and alcohol consumption 24 hours before the start of the testing, caffeine intake the day of the testing and food intake 2 hours before testing. A questionnaire was filled in on each test day to monitor if the participants followed these instructions, as well as to exclude any prior illness or injury that might interfere with the testing. Participants provided written informed consent to voluntarily take part in the study and were informed about the possibility to withdraw from the study at any point in time without providing the reason for doing so. All procedures performed in studies involving human participants were in accordance with the ethical standards of the Regional Ethics Committee for Medical and Health Research in Mid-Norway and with the 1964 Helsinki declaration and its later amendments. The study was retrospectively registered in the Protocol Registration and Results System (NCT03284086).

Table 1 Sex, age, anthropometric and disability characteristics as well as weekly training hours of the participants with paraplegia and the able-bodied participants.

	Sex	Age	Height	Body mass	Lean muscle mass (kg)			Disability (level)	ASIA score	Training
		(years)	(cm)	(kg)	Arms (left + right)	Trunk	Legs (left + right)			hours/week
1	male	26	177	73.5	8.7	32.2	12.5	Paraplegia (L2)	C	5.0
2	male	27	178	59.4	7.2	26.2	13.7	Paraplegia (Th10)	D	6.0
3	male	19	165	79.2	9.5	31.8	7.7	Spina bifida (*N/A*)	*N/A*	7.0
4	male	48	195	95.0	10.6	37.6	14.1	Paraplegia (Th9)	A	8.0
6	female	40	168	65.0	5.1	21.2	12.0	Paraplegia (Th3)	D	6.0
6	male	43	178	83.5	9.7	32.6	11.5	Paraplegia (L1)	A	1.8
7	male	28	192	76.4	8.2	34.1	13.1	Paraplegia (L1)	A	3.5
Mean ± SD		33.8 ± 11.2	179 ± 11	74.4 ± 12.5	8.4 ± 1.8	30.8 ± 5.4	12.1 ± 2.1			5.7 ± 2.0
1	male	26	183	79.1	9.1	33.7	23.4			14
2	female	21	171	72.3	5.7	26.8	18.9			16
3	female	20	168	56.3	4.8	22.5	13.5			11
4	male	22	178	72.2	6.9	31.5	18.7			12.5
5	male	24	183	75.5	7.2	30.7	20.9			6
6	male	22	187	78.1	7.7	34.5	22.2			15
7	male	26	184	80.4	9.4	34.7	24.2			10
8	male	21	190	92	9.6	36.4	25.5			12.5
9	male	19	180	70	7.6	29.6	20.1			14
10	male	23	183	76.3	7.7	32.8	22.8			8
11	male	19	184	79.6	8.1	33.1	23.6			*ns*
Mean ± SD		22.4 ± 2.6	183 ± 3	78.1 ± 6.2	8.1 ± 1.0	33.0 ± 2.1	22.4 ± 2.2			11.5 ± 3.2

American Spinal Cord Injury Association (ASIA), not applicable (N/A), not specified (ns)

Overall design

The testing consisted of two test days, where participants performed four 5-min submaximal steady-state stages, an incremental test to exhaustion and a verification stage in UBP or ACE in a counterbalanced order. Tests were performed at the same time of day in order to minimize the bias of diurnal variation in performance (Atkinson and Reilly 1996). The time between tests was a minimum of 48 hours and a maximum of 4 days. On a separate day before or after the testing, body composition was assessed using dual-energy X-ray absorptiometry (DXA).

Test set-up

After being equipped with an oro-nasal mask (Hans Rudolph Inc, Kansas City, MO, USA) and a heart rate monitor (Polar Electro Inc., Port Washington, NY, USA), participants were tightly strapped into a seat construction, which consisted of a modified weight lifting bench placed in front of the UBP or ACE ergometer (see Figure 1a and 1b). The upper-body was fixed during all tests to limit differences in involvement of the trunk between UBP and ACE as well as AB and PARA. Furthermore, the legs were supported and fixed to minimize leg muscle activation. The spiroergometer (Oxycon Pro, Jaeger, Viasys BV, Bilthoven, the Netherlands) was calibrated against a known mixture of gases (15% O_2, 5% CO_2). The flow transducer was calibrated with a 3-L syringe (Calibration syringe D, SensorMedics, Yorba Linda, CA, USA). Respiratory parameters were assessed by open-circuit calorimetry, with expired gases passing through a mixing chamber and being measured continuously. The Concept2 ski-ergometer (Concept2, Morrisville, USA) was used during testing in the UBP mode. An ErgStick (Endurance Sports Research Limited, United Kingdom) was connected to the PM4 monitor of the Concept2 ski ergometer and the application Float (ErgStick Ltd, United Kingdom) continuously recorded power output (PO) and stroke rate. The ergometer's software has previously been validated with force and velocity measurements (Hegge et al. 2015). The ACE was custom-made from a road-bike (White, XXL Sport & Villmark AS, Norway) and equipped with an electronical brake system for indoor cycling (CompuTrainer^{TM}, RacerMate$^{®}$, Inc., Seattle, USA). The crank axis was aligned with the participant's shoulder height and the seat positioned so that the participant's elbows were slightly bent at maximal reach. The tire pressure was kept stable at six bars and the CompuTrainer^{TM} was calibrated prior to each test session. The in-built software (PerfPRO Studio$^{©}$, Dynastream Innovations Inc., Canada) continuously recorded PO and crank rate.

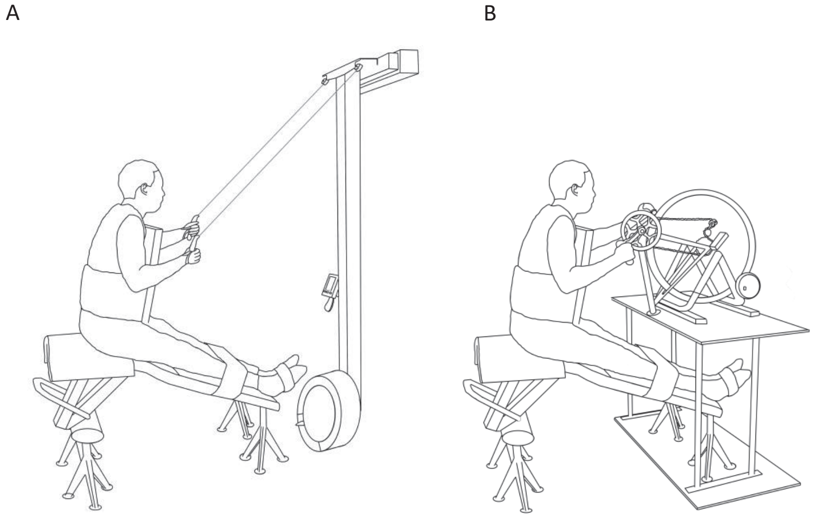


Fig. 1 Test set-up with the participant in a sitting position with the upper-body and the legs fixed in front of the Concept2 ski-ergometer (A) and the arm crank ergometer (B)

Test protocol

Submaximal stages

Prior to testing, participants familiarized themselves with the test set-up by 5 min of arm cranking or upper-body poling at low intensity (overall rating of perceived exertion (RPE) 8-9). The testing then commenced by performing four times 5-min submaximal stages at overall RPEs of 9 (very light), 11 (light), 13 (somewhat hard) and 15 (hard) on a 6-20 Borg scale (Borg 1982). Target RPE at increasing intensities from 9 to 15 (Hegge et al. 2015) was used instead of fixed workloads to ensure that the participants covered a similar range of exercise intensities relative to their maximal capacity. In the ACE mode, crank rate was self-chosen within 60-90 revolutions per minute, whereas stroke rate in the UBP mode was fully self-chosen.

Oxygen uptake (VO_2), respiratory exchange ratio (RER) and minute ventilation (VE) were recorded as 10-s averages. Heart rate (HR) was recorded every s. PO was recorded every second in ACE, and for every stroke in UBP and then interpolated at 1-s intervals. After each submaximal stage there was a 2- to 3-min break during which a 20 µL blood sample was taken from the fingertip and blood lactate (BLa) measured with the Biosen C-Line Sport lactate measurement system (EKF-diagnostic GmbH, Magdeburg, Germany). Furthermore, overall RPE was recorded. Steady-state PO, VO_2, RER, VE, and HR, were calculated by averaging the values during the last 2 min of each sub-maximal stage. There are three primary ways to describe mechanical efficiency during exercise: delta efficiency, net efficiency and gross efficiency. In brief, the challenges with net efficiency and delta efficiency, which are outlined more in detail by Ettema and Lorås (2009), concern the assumption that the processes related to the resting metabolism are independent of the processes associated with producing work. In comparison, gross efficiency, which is the ratio of PO and metabolic rate (MR), is a theoretically sound concept. However, it is affected by the diminishing effect of the resting metabolism with increasing PO. Therefore, we also consider the entire PO-MR relationship to interpret exercise efficiency in the current study.

MR in joule/second (watt) was calculated from VO_2 and RER by the use of a standard conversion table (Peronnet and Massicotte 1991). MR was then estimated from the original PO-MR relationship at a PO of 40, 60 and 80 W using linear regression analyses in Matlab R2016a (MathWorks Inc., Natic, USA). The resulting MR outcomes were used to investigate exercise efficiency between exercise modes and groups. In addition, gross efficiency was calculated as MR divided by PO at 40, 60 and 80 W.

Incremental test to exhaustion

After a 5-min passive break and a 3-min active recovery at the workload equivalent to the first stage (RPE 9), the participants performed an incremental test to exhaustion. The incremental test started at the individual PO of the second submaximal stage (RPE 11) (rounded to the nearest 10-watt value) of the respective mode. PO was then increased by 10 W every 1 min. Termination criteria were a drop in PO and a plateau (3 values with < 2 $mL\cdot kg^{-1}\cdot min^{-1}$ difference) (Taylor et al. 1955) or drop in VO_2 (> 2 $mL\cdot kg^{-1}\cdot min^{-1}$). BLa was measured 1 and 3 min after the incremental test. Furthermore, overall RPE was recorded directly after the incremental test. After a 5-min passive break and a 3-min active recovery, participants performed a verification stage where they directly increased the workload to the peak PO of the incremental test to verify that no higher VO_{2peak} can be obtained despite a longer duration spent at high workload (Leicht et al. 2013).

30-s moving averages were calculated for the PO and the respiratory parameters and the highest values were defined as peak values. 3-s moving averages were calculated for the HR data and the highest value defined as peak HR. The higher of the two blood lactate values was defined as peak BLa.

Statistics

A linear mixed model with fixed coefficients and random intercept was employed to investigate the effect of exercise mode, group and exercise intensity on PO, physiological and perceptual parameters during the submaximal stages and the incremental test. This model investigates the effect of one factor (exercise mode or group or exercise intensity) while adjusting for the two other factors. The same model was used to investigate the effect of exercise mode and group on exercise efficiency. Paired-samples T-tests were used to compare gross efficiency between exercise modes and groups at each of the three POs. An alpha level of 0.05 was used to indicate statistical significance. IBM SPSS Statistics 24.0 (SPSS Inc., Chicago, USA) was used for all statistical analyses.

Results

Peak values from incremental test

An overview of the peak values reached during the incremental test is provided in Table 2. During UBP, when group was adjusted for, participants produced 19% lower peak PO compared to ACE ($p < 0.001$) but displayed 0.08 higher RER ($p < 0.001$). PARA had a 24% lower VO_{2peak} ($p = 0.010$) and 1.2 higher RPE ($p = 0.018$) compared to AB. However, peak PO despite being 14% lower in PARA compared to AB, was not significantly different between PARA and AB ($p = 0.209$). A significant interaction in peak VE existed between exercise mode and group ($p = 0.049$). When investigating each group separately, AB displayed a trend towards a 22% higher peak VE in UBP compared to ACE ($p = 0.069$), whereas in PARA there was no significant difference between modes ($p = 0.804$).

Submaximal values

All outcome parameters significantly increased from the first stage (RPE 9) to the fourth stage (RPE 15) (all comparisons, $p < 0.001$) (Figure 2a and 2b). During UBP, at a given RPE, participants produced 16% lower PO ($p < 0.001$) and displayed 7% higher VO_2 ($p = 0.005$), 9%-point higher % of VO_{2peak} ($p < 0.001$), 8% higher MR ($p = 0.001$), 0.04 higher RER ($p < 0.001$), 19% higher VE ($p < 0.001$), 6% higher HR ($p = 0.001$), 7%-point higher % of peak HR ($p < 0.001$) and 0.50 $mmol \cdot L^{-1}$ higher BLa ($p = 0.002$) compared to ACE. PARA had a trend towards 18% lower PO ($p = 0.081$), and-displayed 20% lower VO_2 ($p = 0.016$) and 22% lower MR ($p = 0.046$). No significant differences between neither modes nor groups were found in % of peak PO and % of peak VE (all comparisons, $p > 0.689$). No significant differences between groups were found in % of VO_{2peak}, VE, RER, HR, % of peak HR and BLa (all comparisons, $p < 0.283$). Furthermore, no significant differences in RPE at 30, 40, 50 and 60% of VO_{2peak} were found ($p = 0.993$).

Exercise efficiency

MR was 24% higher in UBP compared to ACE ($p < 0.001$), i.e. exercise efficiency was lower in UBP (Figure 3A). In line with this, gross efficiency calculated at 40, 60 and 80 W was significantly lower in UBP (10.4 ± 0.9, 11.4 ± 0.8 and 12.0 ± 0.9) compared to ACE (12.9 ± 1.8, 14.0 ± 1.8 and 14.7 ± 1.9) (all comparisons $p < 0.001$). MR was not significantly different between PARA and AB ($p = 0.323$) (Figure 3B and 3C).

Table 2 Peak power output, peak physiological and perceptual responses (Mean ± SD) during the incremental test to exhaustion in the upper-body poling and arm crank ergometry mode in the seven paraplegic and eleven able-bodied participants

		Upper-body poling			**Arm crank ergometry**	
	Paraplegic		Able-bodied	Paraplegic		Able-bodied
Peak PO (Watt)		118 ± 34			146 ± 33*	
	104 ± 35		127 ± 31	136 ± 38		152 ± 29
VO_{2peak} ($mL \cdot kg^{-1} \cdot min^{-1}$)		35.9± 7.8			37.3 ± 8.0	
	30.3 ± 6.1		39.5 ± 6.6†	32.7 ± 7.0		40.3 ± 7.3†
Peak VE ($L \cdot min^{-1}$)		145 ± 37*			126 ± 40	
	131 ± 47		154 ± 28	124 ± 44		126 ± 39
Peak RER		1.19 ± 0.05*			1.11 ± 0.06	
	1.20 ± 0.05		1.19 ± 0.06	1.15 ± 0.06†		1.09 ± 0.04
Peak HR ($beats \cdot min^{-1}$)		176 ± 16			178 ± 17	
	172 ± 20		178 ± 14	182 ± 14		176 ± 18
Peak BLa ($mmol \cdot L^{-1}$)		10.8 ± 1.9			9.9 ± 3.0	
	10.2 ± 2.3		11.3 ± 1.5	10.3 ± 3.7		9.6 ± 2.6
RPE_O (6-20)		18.8 ± 1.2			18.4 ± 1.2	
	19.3 ± 0.5		18.5 ± 1.4	19.2 ± 0.6†		17.8 ± 1.2

*significantly higher in either upper-body poling or arm crank ergometry at an alpha level of 0.05

†significantly higher in either able-bodied or paraplegic participants at an alpha level of 0.05

Power output (PO), peak oxygen uptake (VO_{2peak}), minute ventilation (VE), respiratory exchange ratio (RER), heart rate (HR), blood lactate (BLa), overall rating of perceived exertion (RPE)

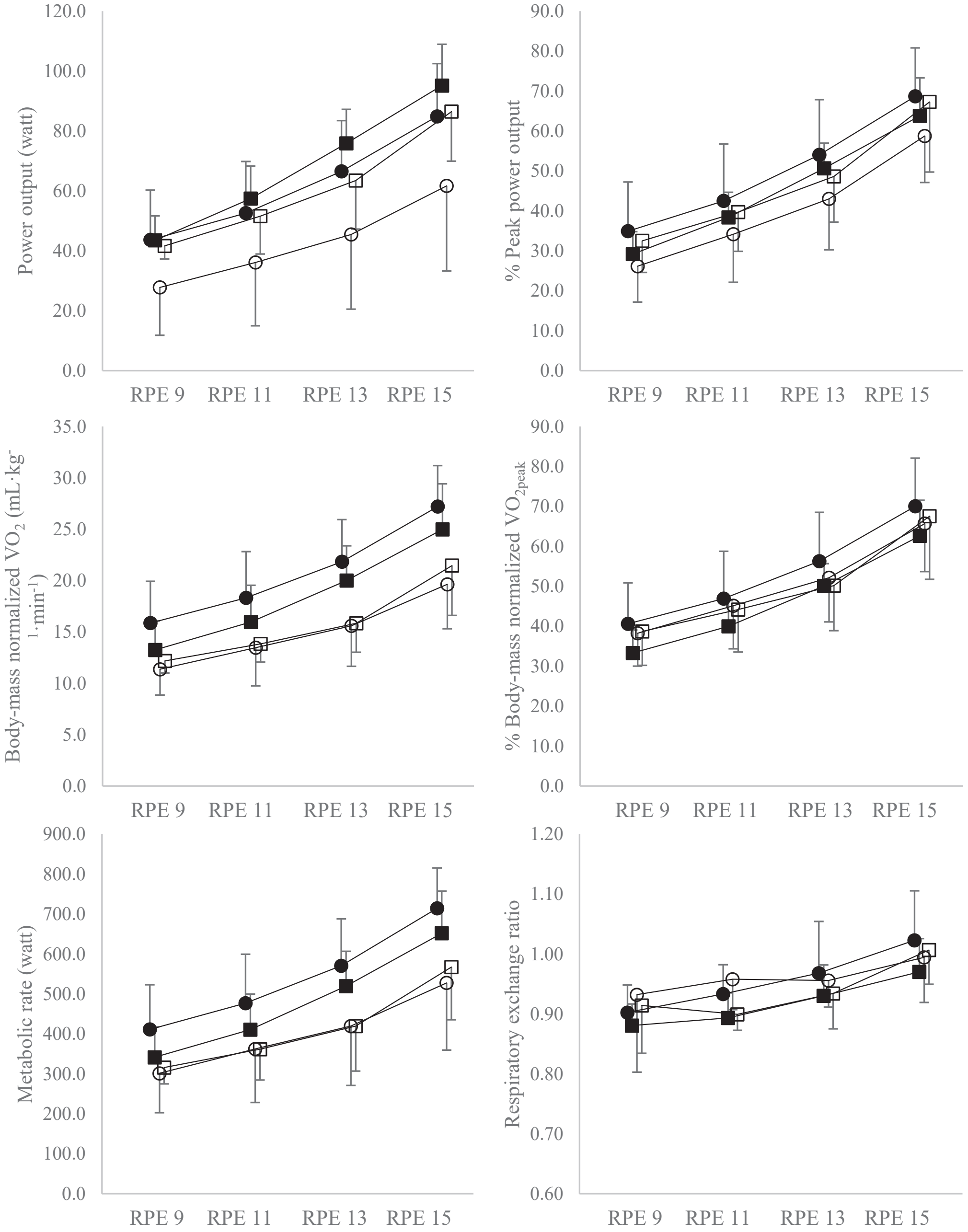


Fig. 2a Power output and physiological parameters at a rating of perceived exertion (RPE) of 9, 11, 13 and 15 presented both as absolute values and as percentage of peak. Furthermore, RPE is presented at 30, 40, 50 and 60% of peak oxyen uptake (VO_{2peak}). (Circles represent the UBP mode, squares the ACE mode. Open symbols represent the PARA participants, closed symbols the AB participants)

Abbreviation: Oxygen uptake (VO_2)

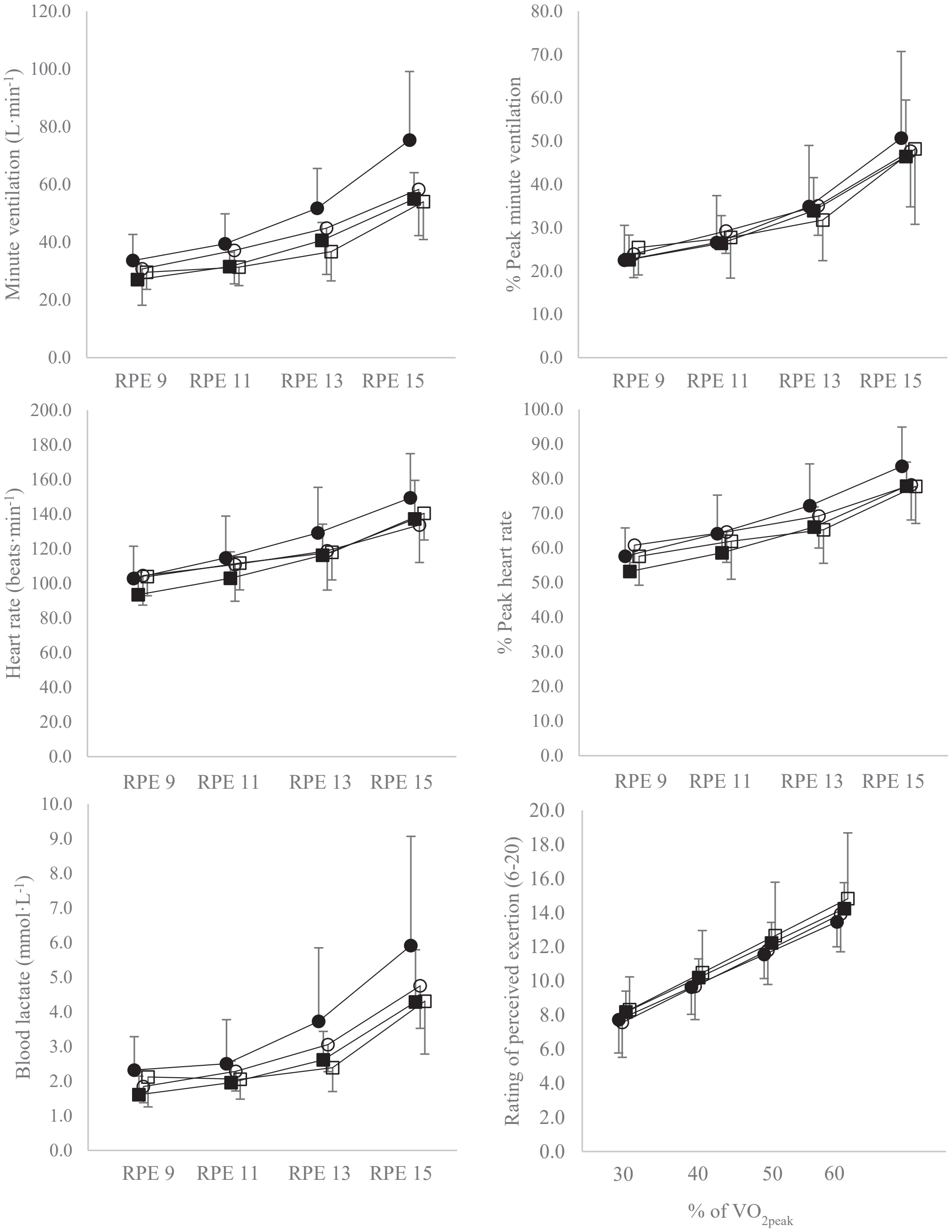


Fig. 2b See Fig. 2a.

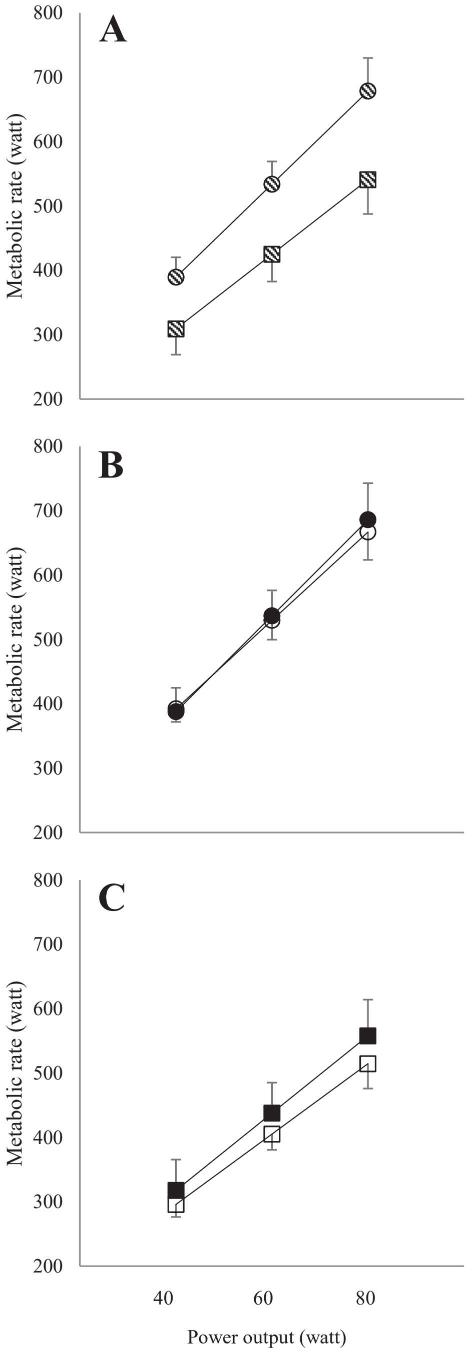


Fig. 3 Metabolic-rate-work-rate relationship in the comparisons of A) upper-body poling (circles) and arm crank ergometry (squares) with paraplegic and able-bodied participants pooled, B) paraplegic (open circles) and able-bodied participants (closed circles) in the upper-body poling mode, and C) paraplegic (open squares) and able-bodied participants (closed squares) in the arm crank ergometry mode

Discussion

The aim of this study was to compare VO_{2peak} and exercise efficiency between upper-body poling (UBP) and arm crank ergometry (ACE) in paraplegic (PARA) and able-bodied (AB) participants. As expected, VO_{2peak} did not differ between UBP and ACE, indicating that both modes tax the cardiovascular system similarly. However, there was an 19~~8~~% lower peak PO produced in UBP that coincided with the 24% higher metabolic energy at a given power output (i.e., lower gross efficiency). PARA did not differ from AB in exercise efficiency, but PARA had 24% lower VO_{2peak} compared to AB.

Differences between UBP and ACE

This is the first study to investigate differences in VO_{2peak} and exercise efficiency between UBP and ACE. We found no difference in VO_{2peak} between UBP and ACE, which indicates that – with the upper-body restricted – the cardiorespiratory system is taxed equally in both exercise modes when working until voluntary exhaustion. In addition, no differences in peak HR, peak BLa and RPE were found between UBP and ACE, indicating that a similar level of exhaustion was reached at the end of the tests. However, although the peak aerobic energy delivery capacity and the ability to reach exhaustion did not differ between exercise modes, peak PO was clearly lower in UBP compared to ACE. The difference in peak PO is likely explained by UBP being a less efficient test mode, which is also supported lower efficiency in UBP compared to ACE at submaximal workloads.

The higher MR at a given power (i.e. lower gross efficiency) in UBP may be related to that power is produced in a discontinuous movement, which includes larger fluctuations in instantaneous power, compared to in ACE where the movement is more continuous. In line with this, studies comparing wheelchair propulsion to ACE have found that the discontinuous movement during wheelchair propulsion is less efficient (Glaser et al. 1980; Hintzy et al. 2002; Mukherjee and Samanta 2001). Discontinuous force application has been shown to increase power fluctuations within strokes, a strategy that costs more for the production of the same PO (Glaser et al. 1980). Furthermore, in UBP the participants move their arms up against gravity before pulling down on the ropes. This movement fundamentally differs from ACE where the arms are supported by the cranks throughout the whole cycle, a movement pattern that has previously been associated with higher exercise efficiency (Mukherjee and Samanta 2001), arguably due to reutilization of kinetic energy. Altogether, the lower exercise efficiency in UBP compared to ACE may be explained by the different movement characteristics.

The percentage of peak PO employed at the various RPE-matched submaximal stages was almost identical between UBP and ACE. However, UBP showed a higher MR and a trend towards lower absolute PO during each of these stages, which is associated with the lower exercise efficiency in UBP. In addition, all related physiological outcome parameters (e.g., VO_2, MR, RER, VE, HR, BLa) were significantly higher at a given RPE during UBP compared to ACE. Therefore, the greater physiological stress during UBP might be related to differences in local metabolic responses in the working upper-body muscles, such as a higher local oxygen desaturation and a lower local muscle blood flow (this is indicated by unpublished data from our research group), in response to the higher instantaneous power production during each stroke in the UBP compared to the ACE mode. However, further studies measuring local muscle blood flow and desaturation are needed to investigate this hypothesis.

Differences between AB and PARA

As expected, PARA displayed significantly lower VO_{2peak} compared to AB, which might partially be due to a more limited ability to recruit muscle mass during testing. Even though we tried to minimize differences in trunk and leg stabilization between PARA and AB, we still observed leg muscle contractions in AB especially towards the end of the incremental test. In addition, VO_{2peak} might be lower in PARA due the inability to redistribute blood from the paralyzed trunk and lower limbs, which is related to a reduced stroke volume and, at higher exercise intensities, to a reduced cardiac output (Hopman et al. 1993; Thijssen et al. 2009). In PARA with an injury level above Th6, VO_{2peak} may be further restricted due to reduced blood redistribution from the splanchnic vascular bed (Thijssen et al. 2009). Furthermore, the lower VO_{2peak} in PARA may be related to the fact that AB perform more overall training hours. Whereas the amount of upper-body training did not differ between PARA and AB and the two groups had a similar amount of muscle mass in the upper-body, AB had twice the amount of overall training hours due to additional exercise with lower-body and whole-body exercise modes. Additionally, PARA consisted of a group of athletes from different sports, whereas AB were all cross-country skiers. As such, PARA might be less specifically trained for the upper-body poling movement, and one might expect the difference in VO_{2peak} to be bigger in UBP as compared to ACE. However, differences in VO_{2peak} were similar between PARA and AB both in UBP and ACE. This indicates that, when the upper-body is restricted, sports-specificity does not seem to have a major effect on VO_{2peak}.

There was no difference in MR at a given PO between PARA and AB, indicating that PARA were equally efficient as AB. In addition, AB had a higher VO_2 but also a comparably higher absolute PO at all RPE-matched submaximal stages. Hence, as a percentage of VO_{2peak} and of peak PO, participants exercised at the same relative intensity in both groups. Furthermore, none of the other physiological parameters significantly differed between PARA and AB when expressed as a percentage of peak values. Concluding from the above, differences in the submaximal responses between AB compared to PARA are due to AB working at higher PO and not due to differences in exercise efficiency.

Methodological considerations

While the fixed position of the upper-body reduced potential differences in the use of the muscles of the trunk and pelvic region between UBP and ACE as well as AB and PARA, it likely influenced VO_{2peak} and other related outcome parameters as well. Not restricting upper-body movement (as is more commonly seen when UBP is used in during training and competition), would have led to a different use of the trunk in UBP compared to ACE, in particular in the comparison of PARA versus AB. In UBP, trunk movement can easily contribute to increased power production and thereby elevated MR (Hegge et al. 2015). In comparison, due to the asynchronous arm movements in ACE, there is a lower contribution of the trunk movement to power production. Further studies are needed to compare the effect of fixed trunk versus allowing the trunk to move freely on VO_{2peak} during UBP incremental exercise to exhaustion.

Conclusion

In upper-body trained PARA and AB participants, VO_{2peak} did not differ between UBP and ACE, indicating that the movement patterns of these two test modes tax the cardiovascular system to a similar extent when the trunk is fixed. The 18% lower peak PO in UBP compared to ACE may be explained by the coinciding lower efficiency in the same mode. Furthermore, the lower VO_{2peak} in PARA compared to AB is likely related to their disability, i.e. less active muscle mass during testing and a limited blood redistribution below the level of injury. However, there was no difference in exercise efficiency between PARA and AB in the two modes. Overall, the findings of this study provide a good starting point for understanding the differences in outcome parameters between two commonly used test modes and between PARA and AB athletes. However, to allow coaches and researchers to implement our findings into practice, future research should complement our results by investigating whether differences in trunk involvement between UBP and ACE lead to differences in VO_{2peak} and efficiency.

Acknowledgements

We appreciate the time and effort of the participants who took part in this study. Furthermore, we thank Katrin Daetwyler and Knut Skovereng for their assistance in the laboratory.

Compliance with ethical standards

Funding This study was funded by the Centre for Elite Sports Research, Department of Neuromedicine and Movement Science, Norwegian University of Science and Technology. The laboratory facilities and equipment were provided by NeXt Move, Norwegian University of Science and Technology (NTNU). *NeXt Move* is funded by the Faculty of Medicine at NTNU and Central Norway Regional Health Authority. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Conflicts of interest The authors declare no financial or non-financial conflicts of interest.

References

Arabi H, Vandewalle H, Pitor P, de Lattre J, Monod H (1997) Relationship between maximal oxygen uptake on different ergometers, lean arm volume and strength in paraplegic subjects. Eur J Appl Physiol Occup Physiol 76:122-127 doi:10.1007/s004210050223

Atkinson G, Reilly T (1996) Circadian variation in sports performance. Sports Med 21:292-312

Bloemen MA, de Groot JF, Backx FJ, Westerveld RA, Takken T (2015) Arm cranking versus wheelchair propulsion for testing aerobic fitness in children with spina bifida who are wheelchair dependent. J Rehabil Med 47:432-437 doi:10.2340/16501977-1944

Borg GA (1982) Psychophysical bases of perceived exertion. Med Sci Sports Exerc 14:377-381

Drory Y, Ohry A, Brooks ME, Dolphin D, Kellermann JJ (1990) Arm crank ergometry in chronic spinal cord injured patients. Arch Phys Med Rehabil 71:389-392

Ettema G, Loras HW (2009) Efficiency in cycling: a review. Eur J Appl Physiol 106:1-14 doi:10.1007/s00421-009-1008-7

Gass EM, Harvey LA, Gass GC (1995) Maximal physiological responses during arm cranking and treadmill wheelchair propulsion in T4-T6 paraplegic men. Paraplegia 33:267-270 doi:10.1038/sc.1995.60

Gass GC, Camp EM (1984) The maximum physiological responses during incremental wheelchair and arm cranking exercise in male paraplegics. Med Sci Sports Exerc 16:355-359

Gayle GW, Pohlman RL, Glaser RM, Davis GM (1990) Cardiorespiratory and perceptual responses to arm crank and wheelchair exercise using various handrims in male paraplegics. Res Q Exerc Sport 61:224-232 doi:10.1080/02701367.1990.10608683

Glaser RM, Sawka MN, Brune MF, Wilde SW (1980) Physiological responses to maximal effort wheelchair and arm crank ergometry. J Appl Physiol Respir Environ Exerc Physiol 48:1060-1064

Hegge AM, Bucher E, Ettema G, Faude O, Holmberg HC, Sandbakk O (2015) Gender differences in power production, energetic capacity and efficiency of elite cross-country skiers during whole-body, upper-body, and arm poling. Eur J Appl Physiol doi:10.1007/s00421-015-3281-y

Hintzy F, Tordi N, Perrey S (2002) Muscular efficiency during arm cranking and wheelchair exercise: a comparison. Int J Sports Med 23:408-414 doi:10.1055/s-2002-33734

Hopman MT, Houtman S, Groothuis JT, Folgering HT (2004) The effect of varied fractional inspired oxygen on arm exercise performance in spinal cord injury and able-bodied persons. Arch Phys Med Rehabil 85:319-323

Hopman MT, Verheijen PH, Binkhorst RA (1993) Volume changes in the legs of paraplegic subjects during arm exercise. J Appl Physiol (1985) 75:2079-2083 doi:10.1152/jappl.1993.75.5.2079

Leicht C, Perret C (2008) Comparison of Blood Lactate Elimination in Individuals With Paraplegia and Able-Bodied Individuals During Active Recovery From Exhaustive Exercise. The Journal of Spinal Cord Medicine 31:60-64 doi:10.1080/10790268.2008.11753982

Leicht CA, Tolfrey K, Lenton JP, Bishop NC, Goosey-Tolfrey VL (2013) The verification phase and reliability of physiological parameters in peak testing of elite wheelchair athletes. Eur J Appl Physiol 113:337-345 doi:10.1007/s00421-012-2441-6

Martel G, Noreau L, Jobin J (1991) Physiological responses to maximal exercise on arm cranking and wheelchair ergometer with paraplegics. Spinal Cord 29:447-456

McCafferty WB, Horvath SM (1977) Specificity of exercise and specificity of training: a subcellular review. Res Q 48:358-371

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

Mukherjee G, Samanta A (2001) Physiological response to the ambulatory performance of hand-rim and arm-crank propulsion systems. J Rehabil Res Dev 38:391-399

Peronnet F, Massicotte D (1991) Table of nonprotein respiratory quotient: an update. Can J Sport Sci 16:23-29

Price M, Campbell I (1999a) Thermoregulatory responses of spinal cord injured and able-bodied athletes to prolonged upper body exercise and recovery. Spinal Cord 37

Price MJ, Bottoms L, Smith PM, Nicholettos A (2011) The effects of an increasing versus constant crank rate on peak physiological responses during incremental arm crank ergometry. J Sports Sci 29:263-269 doi:10.1080/02640414.2010.525520

Price MJ, Campbell IG (1999b) Thermoregulatory and physiological responses of wheelchair athletes to prolonged arm crank and wheelchair exercise. Int J Sports Med 20:457-463 doi:10.1055/s-1999-8831

Sawka MN (1986) Physiology of upper body exercise. Exerc Sport Sci Rev 14:175-211

Sawka MN, Glaser RM, Wilde SW, von Luhrte TC (1980) Metabolic and circulatory responses to wheelchair and arm crank exercise. J Appl Physiol 49:784-788

Smith PM, Amaral I, Doherty M, Price MJ, Jones AM (2006) The influence of ramp rate on VO2peak and "excess" VO2 during arm crank ergometry. Int J Sports Med 27:610-616 doi:10.1055/s-2005-865857

Smith PM, Doherty M, Drake D, Price MJ (2004) The influence of step and ramp type protocols on the attainment of peak physiological responses during arm crank ergometry. Int J Sports Med 25:616-621 doi:10.1055/s-2004-817880

Smith PM, Doherty M, Price MJ (2007) The effect of crank rate strategy on peak aerobic power and peak physiological responses during arm crank ergometry. J Sports Sci 25:711-718 doi:10.1080/02640410600831955

Smith PM, Price MJ, Doherty M (2001) The influence of crank rate on peak oxygen consumption during arm crank ergometry. J Sports Sci 19:955-960 doi:10.1080/026404101317108453

Taylor HL, Buskirk E, Henschel A (1955) Maximal oxygen intake as an objective measure of cardio-respiratory performance. J Appl Physiol 8:73-80

Thijssen DH, Steendijk S, Hopman MT (2009) Blood redistribution during exercise in subjects with spinal cord injury and controls. Med Sci Sports Exerc 41:1249-1254 doi:10.1249/MSS.0b013e318196c902

Tropp H, Samuelsson K, Jorfeldt L (1997) Power output for wheelchair driving on a treadmill compared with arm crank ergometry. Br J Sports Med 31:41-44

Paper III

frontiers in Physiology

ORIGINAL RESEARCH
published: 30 October 2017
doi: 10.3389/fphys.2017.00857

Comparison of Peak Oxygen Uptake and Test-Retest Reliability of Physiological Parameters between Closed-End and Incremental Upper-Body Poling Tests

Julia K. Baumgart, Knut Skovereng and Øyvind Sandbakk*

Department of Neuromedicine and Movement Science, Centre for Elite Sports Research, Norwegian University of Science and Technology, Trondheim, Norway

OPEN ACCESS

Edited by:
Billy Sperlich,
Integrative and Experimentelle Trainingswissenschaft, Universität Würzburg, Germany

Reviewed by:
Hannes Gatterer,
University of Innsbruck, Austria
Stefanos Volianitis,
Aalborg University, Denmark

****Correspondence:***
Julia K. Baumgart
jk.baumgart@gmail.com

Specialty section:
This article was submitted to Exercise Physiology, a section of the journal Frontiers in Physiology

Received: *18 July 2017*
Accepted: *13 October 2017*
Published: *30 October 2017*

Citation:
Baumgart JK, Skovereng K and Sandbakk Ø (2017) Comparison of Peak Oxygen Uptake and Test-Retest Reliability of Physiological Parameters between Closed-End and Incremental Upper-Body Poling Tests. Front. Physiol. 8:857. doi: 10.3389/fphys.2017.00857

Objective: To compare peak oxygen uptake (VO_{2peak}) and the test-retest reliability of physiological parameters between a 1-min and a 3-min closed-end and an incremental open-end upper-body poling test.

Methods: On two separate test days, 24 healthy, upper-body trained men (age: 28.3 ± 9.3 years, body mass: 77.4 ± 8.9 kg, height: 182 ± 7 cm) performed a 1-min, a 3-min and an incremental test to volitional exhaustion in the same random order. Respiratory parameters, heart rate (HR), blood lactate concentration (BLa), rating of perceived exertion (RPE), and power output were measured. VO_{2peak} was determined as the single highest 30-s average. Relative reliability was assessed with the intra-class correlation coefficient ($ICC_{2,1}$) and absolute reliability with the standard error of measurement (SEM) and smallest detectable change (SDC).

Results: The incremental (3.50 ± 0.46 $L \cdot min^{-1}$ and 45.4 ± 5.5 $mL \cdot kg^{-1} \cdot min^{-1}$) and the 3-min test (3.42 ± 0.47 $L \cdot min^{-1}$ and 44.5 ± 5.5 $mL \cdot kg^{-1} \cdot min^{-1}$) resulted in significantly higher absolute and body-mass normalized VO_{2peak} compared to the 1-min test (3.13 ± 0.40 $L \cdot min^{-1}$ and 40.4 ± 5.0 $mL \cdot kg^{-1} \cdot min^{-1}$) (all comparisons, $p < 0.001$). Furthermore, the incremental test resulted in a significantly higher VO_{2peak} as compared to the 3-min test ($p < 0.001$). VO_{2peak} was significantly higher on day 1 than day 2 for the 1-min test ($p < 0.05$) and displayed a trend toward higher values on day 2 for the incremental test ($p = 0.07$). High and very high ICCs across all physiological parameters were found for the 1-min (0.827–0.956), the 3-min (0.916–0.949), and the incremental test (0.728–0.956). The SDC was consistently small for HR (1-min: 4%, 3-min: 4%, incremental: 3%), moderate for absolute and body-mass normalized VO_{2peak} (1-min: 5%, 3-min: 6%, incremental: 7%) and large for BLa (1-min: 20%, 3-min: 12%, incremental: 22%).

Conclusions: Whereas both the 3-min and the incremental test display high relative reliability, the incremental test induces slightly higher VO_{2peak}. However, the 3-min test seems to be more stable with respect to day-to-day differences in VO_{2peak}. The 1-min test would provide a reliable alternative when short test-duration is desirable, but is not recommended for testing VO_{2peak} due to the clearly lower values.

Keywords: peak aerobic capacity, endurance performance, all-out, 3-min, exhaustion

INTRODUCTION

Exercise testing in a sitting position is relevant for determining upper-body physiological capacities and monitoring training progression in both Paralympic sitting athletes as well as able-bodied athletes involved in an upper-body sport. Various test protocols have been used to determine peak oxygen uptake (VO_{2peak}) in upper-body modes, with the most common test procedure comprising incremental increases in workload until voluntary exhaustion (Bar-Or and Zwiren, 1975; Bhambhani et al., 1991; Leicht et al., 2009, 2013; Hutchinson et al., 2017). In addition, a 3-min self-paced closed-end test is a common procedure to assess VO_{2peak} in upper-body modes (Skovereng et al., 2013; Flueck et al., 2015; Hegge et al., 2015a,b; Baumgart and Sandbakk, 2016).

In cycling, the 3-min and incremental tests resulted in equally high VO_{2peak} values (Sperlich et al., 2011). The 3-min test additionally includes indices of performance and anaerobic capacity (i.e., accumulated oxygen deficit) (Losnegard et al., 2012), and therefore covers a more complementary set of measurements in a single test as compared to incremental workloads. In addition, the ability to increase the utilization of VO_2 rapidly plays an important role in sports where high power outputs are produced over a relatively short time period. Examples are middle distance sports or intermittent activities such as cross-country skiing or Para cross-country skiing where hard work is performed in steep uphills followed by recovery in the subsequent downhill sections. Therefore, it would be of interest to explore the maximal rate of VO_2 uptake during a test of shorter duration than traditionally employed. However, VO_{2peak} and corresponding physiological responses during closed-end tests of different duration and an incremental protocol have not yet been compared in upper-body exercise modes.

In an athletic context, sport-specificity of the testing mode is important in eliciting performance related peak responses (Roels et al., 2005). Upper-body poling is the most sports-specific mode for ice sledge hockey players and cross-country sit skiers as well as for testing upper-body capacity in cross-country skiers, biathletes and Nordic combined athletes. Thirty-nine Paralympic and 27 Olympic gold medals are contended for in these events, highlighting the importance of reliable test concepts for such sports. The 3-min and the incremental test are regarded reliable for the determination of VO_{2peak} as well as other physiological and perceptual parameters during arm-crank (Bar-Or and Zwiren, 1975; Leicht et al., 2009; Flueck et al., 2015) and wheelchair ergometry (Bhambhani et al., 1991; Leicht et al., 2013). However, the test-retest reliability of physiological parameters in upper-body poling needs to be established before meaningful differences between athletes and repeated tests within athletes can be interpreted.

The determination of test-retest reliability requires a relatively large group of homogeneous participants since most statistical measures of absolute and relative reliability are sensitive to population heterogeneity (Atkinson and Nevill, 1998; Hopkins, 2000; Weir, 2005). High test-retest reliability can solely be the result of a large spread of data points as compared to small intra-participant day-to-day variation (Atkinson and Nevill, 1998; Hopkins, 2000; Weir, 2005). Paralympic athletes represent a small group of participants with a large heterogeneity of physical capacities and are not preferable in this context.

Therefore, the aim of this study was to compare VO_{2peak} and the test-retest reliability between a 1-min and a 3-min closed-end and an incremental open-end upper-body poling test in able-bodied upper-body-trained participants. We hypothesized that the 3-min and the incremental upper-body poling tests would display high test-retest reliability as these protocols were previously found reliable for arm-crank and wheelchair ergometry. In line with previous research in cycling, we expected that the 3-min and incremental protocol would not differ in VO_{2peak}.

MATERIALS AND METHODS

Participants

Twenty-four able-bodied upper-body-trained male individuals (age 28.3 ± 9.3, body mass 77.4 ± 8.9 kg, and height 1.8 ± 0.1 m) participated in this study. Participants were mainly cross-country skiers ($N = 23$) and additionally one rower who regularly trained cross-country skiing, all of whom participated in recreational or national level cross-country skiing and rowing races, respectively. All were highly trained with a running VO_{2max} of 66 ± 7 $mL\cdot kg^{-1}\cdot min^{-1}$ (range 53.0–75.9) and an average of 39 ± 11 (range 22.5–75) training hours per month (based on self-reported training hours from their training diary logs; www.olt-dagbok.net), most of which was endurance training and a considerable part employing the upper-body. The participants were instructed to refrain from heavy training

---

Abbreviations: BLa_{peak}, Peak blood lactate; HR_{peak}, Peak heart rate; ICC, Interclass correlation coefficient; PO_{peak}, Peak power output; RPE_M, Muscular rate of perceived exertion; RPE_O, Overall rate of perceived exertion; RPE_R, Respiratory rate of perceived exertion; SDC, Smallest detectable change; SEM, Standard error of the measurement; VE, Minute ventilation; VO_{2max}, Maximal oxygen uptake; VO_{2peak}, Peak oxygen uptake; VCO_{2peak}, Peak carbon dioxide production.

and alcohol consumption 24 h before the start of the testing, caffeine intake the day of the testing and food intake 2 h before. A questionnaire was filled out on each day to monitor if the participants followed these instructions, as well as to exclude any prior illness or injury that might have interfered with the testing. The study was pre-approved by the Regional Committee for Medical and Health Research Ethics of Mid-Norway and conducted in accordance with the Declaration of Helsinki. All participants signed an informed consent form prior to participation in the experiment and were made aware that they could withdraw from the study at any point without providing an explanation.

Overall Design

The testing consisted of two test days, where participants performed a 1-min and a 3-min all out and an incremental test to exhaustion in an upper-body poling mode on a Concept2 ski-ergometer (Concept2, Inc., Morrisville, USA). Each participant performed the tests in the same order and at the same time of the day (to minimize the bias of diurnal variation in performance; Atkinson and Reilly, 1996). The test order was randomized between participants. The time between test days was a minimum of 48 h and an average of 4 ± 3 days (range 2–11 days). Before the start of the testing on the first day, the participants' body mass was assessed by the built-in weighing scale of a bioelectrical impedance analyzer (Inbody Co., Ltd., Seoul, Korea).

Test Set-Up and Familiarization

After being equipped with an oro-nasal mask (Hans Rudolph Inc, Kansas City, MO, USA) and a heart rate (HR) monitor (Polar Electro Inc., Port Washington, NY, USA), the participants tightly strapped themselves around the hips and thighs into a seat construction (see **Figure 1**). They were then familiarized with the test setup and mode by performing four times 5-min submaximal stages mode at an overall rate of perceived exertion of 9 (very light), 11 (light), 13 (somewhat hard), and 15 (hard) on a 6–20 Borg scale (Borg, 1982; Shephard et al., 1992).

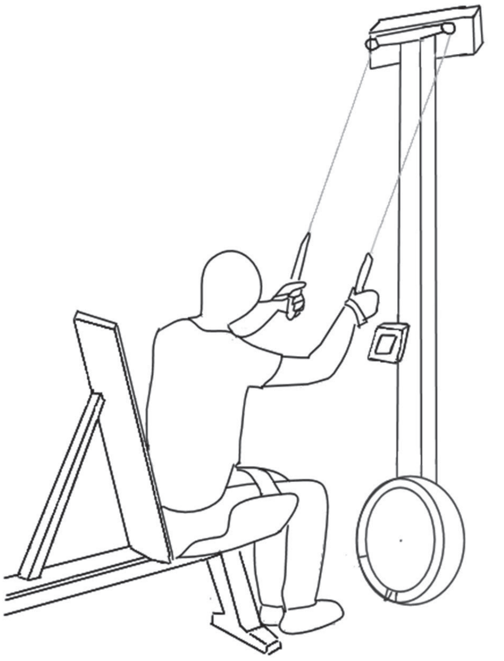


FIGURE 1 | Test set-up with the participant in a sitting position, strapped around the hips and thighs, in front of the Concept2 ski-ergometer.

Test Protocols

A 15-min break followed the submaximal familiarizing stages, before a standardized 5-min warm-up preceded each of the peak tests, consisting of three and two min at the power output of the third (RPE 13) and fourth submaximal stage (RPE 15), respectively. The third minute of warm-up was inter-dispersed by two 5-s sprints at 90–95% of maximal sprint power. Both, the 1-min and the 3-min test were self-paced closed-end tests with the instruction to find a power output that the participants thought they could maintain throughout the test. Three pacing strategies were possible: (1) a higher power output at the start with a drop toward the end (positive pacing), (2) a stable power output throughout the test (even pacing), and (3) a lower power output at the start with an increase toward the end (negative pacing) (Atkinson et al., 2007). Positive and negative pacing were defined as more than a 10% increase or decrease of the last 30-s average as compared to the initial 30-s average power output. The incremental test started at the individual power output of the third submaximal stage (rounded to the nearest 5-W value) and participants were instructed to continuously increase power output by 10 W every 30 s. Between each of the maximal tests, a rest period of 26 ± 3 min was given and the participants were optionally allowed to drink water or sports drink. It has previously been shown that a recovery period of 20 min between maximal tests allows participants to maintain performance (Weltman et al., 1979; Vesterinen et al., 2009; Moxnes and Moxnes, 2014).

VO_2, VCO_2, and VE were measured breath-by-breath using a spiroergometer (Oxycon Pro, Jaeger, Viasys BV, CA, USA) which was calibrated against a known mixture of gases (5% CO_2, 15% O_2) and a known air flow (from a 3 L syringe) prior to each test. HR was continuously recorded during the tests. A blood sample was taken 1 and 3 min after each test and blood lactate analyzed by a Biosen C-Line Sport lactate measurement system (EKF-diagnostic GmbH, Magdeburg, Germany). Overall (RPE_O), respiratory (RPE_R) and muscular rate of perceived exertion (RPE_M), and were recorded after each test as described more in detail by Shephard et al. (1992). Power output per stroke was recorded by the skiergometer's internal software (Concept2, Morrisville, USA) and recorded by a Sony Alpha 58 video camera (Sony Corporation, NY, USA).

Data Processing and Statistical Analyses

A minimum number 21 participants was determined by a-priori analyses in G*Power 3.1, with an effect size of 0.65 (calculated as a Cohen's d based on VO_{2peak} values in a similar sample from Baumgart and Sandbakk, 2016, an alpha level of 0.05 and a power of 0.80. Two of the 24 participants were not able to complete the 1-min and the 3-min test on the second day. In these two tests, the data from 22 participants were analyzed. Breath-by-breath respiratory data was interpolated at individually fitted sample frequencies, resampled at 1-s intervals and 10, 30, and 60-s averages were calculated in MATLAB 8.1.0. (R2016a; Mathworks Inc., Natick, MA) The single highest 30-s average value was then identified as VO_{2peak} as recommended by Robergs et al. (2010) and used in the further analyses. The highest 10 and 60-s averages were used to investigate if changes in averaging procedure affected the results. Moving 3-s averages were calculated for the HR data and the highest value defined as peak heart rate (HR_{peak}). The higher of the two blood lactate values was defined as peak blood lactate (BLa_{peak}). Thirty seconds averages were calculated for the PO data and the highest value defined as peak power output (PO_{peak}).

Data are presented as mean ± *SD* unless specified differently and an α level of 0.05 was employed to indicate statistical significance. All calculations and statistical tests were executed in Microsoft Excel (Version 2010, Microsoft Cooperation, The Microsoft Network, LLC, Richmond, USA) or in SPSS 22.0 (Software for Windows, SPSS Inc., Chicago, IL, USA).

Assumptions

The assumption of homoscedasticity was examined by plotting the individual test-retest differences against the individual means and by calculating the Pearson's *r* correlation coefficient between the two. A correlation of $r > 0.25$ was used to define heteroscedasticity (O'Donoghue, 2013). Heteroscedastic variables (VCO_{2peak}, $VE_{peak,}$ and PO_{peak} of the 1-min test, and RPE_O and RPE_M of the incremental test) were transformed using the natural logarithm. However, this procedure did not improve the heteroscedasticity and we hence used the non-transformed data. The assumption of normally distributed test-retest differences was assessed by the Shapiro-Wilk test of normality and Normal Q-Q plots. Paired-samples *T*-Tests were used to assess systematic bias in physiological variables, RPE and PO_{peak} between the two test days. Independent-samples *T*-Tests were used to investigate whether using the same or unequal pacing strategies led to differences in the VO_{2peak} delta values from day 1 to day 2. A general linear mixed model was used to investigate the interaction effect of test order and the type of upper-body poling peak test on VO_{2peak}.

Comparison of Tests

Paired-samples *T*-tests were used to compare physiological variables, RPE_M, RPE_R, $RPE_{O,}$ and PO_{peak} between the three tests. To investigate the influence of a different averaging procedure on VO_{2peak} we compared 10 and 60-s average to the 30-s average described above with paired-samples *T*-tests. The average over day 1 and 2 for each variable was used for these comparisons.

Absolute Reliability

Absolute reliability was assessed by the standard error of measurement (SEM) and the smallest detectable change (SDC). The SEM was calculated as $SD_{\text{diff}}/\sqrt{2}$ (Hopkins, 2000), and the 80% SDC as $\text{SEM}\cdot 1.28\cdot\sqrt{2}$ (Bland and Altman, 1986).

Relative Reliability

Intraclass correlation coefficients ($ICC_{2,1}$) with 95% CI were calculated as a measure of relative reliability (Weir, 2005). Ranges of 0.26–0.49, 0.50–0.69, 0.70–0.89, and 0.90–1.0 were classified as low, moderate, high, and very high ICC according to Munro's criteria (Plichta et al., 2013).

RESULTS

Comparison of Tests

Individual differences and mean values of day 1 to day 2 and corresponding limits of agreement are visualized in Bland-Altman plots in **Figure 2** and displayed in **Table 1**. All data used in the analyses of this study are found in Datasheet 1. Based on the average values of test day 1 and 2, the incremental (45.4 ± 5.5 $mL\cdot kg^{-1}\cdot min^{-1}$, 196 ± 28 Watt) and the 3-min test (44.5 ± 5.5 $mL\cdot kg^{-1}\cdot min^{-1}$, 202 ± 37 W) resulted in significantly higher VO_{2peak} and lower PO_{peak} as compared to the 1-min test (40.4 ± 5.0 $mL\cdot kg^{-1}\cdot min^{-1}$, 253 ± 46 W) (all $p < 0.001$). Additionally, the incremental test resulted in significantly higher VO_{2peak} as well as lower PO_{peak} as compared to the 3-min test (both $p < 0.001$) (see **Supplementary Figure 1**). A plateau in VO_{2peak} (2 consecutive 30-s values within 2 $mL\cdot kg^{-1}\cdot min^{-1}$) was observed in ~80% of tests both during the 3-min and the incremental protocols of day 1 and 2, without any difference between test protocol or order of test day.

As compared to the 30-s average used in the above, employing a 10 or 60-s average would have resulted in significantly higher or lower VO_{2peak}, respectively, in all three tests (all comparisons $p < 0.001$). When using 10-s averages, the difference in VO_{2peak} between the tests would have remained unchanged (1-min: 41.3 ± 5.3, 3-min: 45.4 ± 5.5, incremental: 46.3 ± 5.6 $mL\cdot kg^{-1}\cdot min^{-1}$) compared to using 30-s averages. However, using 60-s averages, the VO_{2peak} difference between the 1-min as compared to the 3-min and the incremental test would have increased, and the differences between the 3-min and the incremental test decreased (1-min: 32.6 ± 4.2, 3-min: 43.8 ± 5.7, incremental: 44.2 ± 5.4).

Relative and Absolute Reliability

High and very high ICCs across all physiological outcome parameters and PO_{peak} were found for the 1-min, the 3-min and the incremental test (**Table 2**). In all three tests, the SDC was consistently small for HR_{peak} (1-min: 4%, 3-min: 4%, incremental: 3%), moderate for absolute and body-mass normalized VO_{2peak} (1-min: 5%, 3-min: 6%, incremental: 7%) as well as PO_{peak} (1-min: 10%, 3-min: 9%, incremental: 6%) and large for BLa_{peak} (1-min: 20%, 3-min: 12%, incremental: 22%).

Fourteen and 9 participants changed their pacing strategy from day 1 to day 2 for the 1-min and the 3-min test, respectively. However, there were no differences in VO_{2peak} between day 1 and day 2 for neither the 1-min (1.1 ± 1.7 vs. 0.8 ± 1.7 $mL\cdot kg^{-1}\cdot min^{-1}$, $p = 0.67$) nor the 3-min test (0.7 ± 2.0 vs. 0.2 ± 2.5 $mL\cdot kg^{-1}\cdot min^{-1}$, $p = 0.53$) when comparing those who

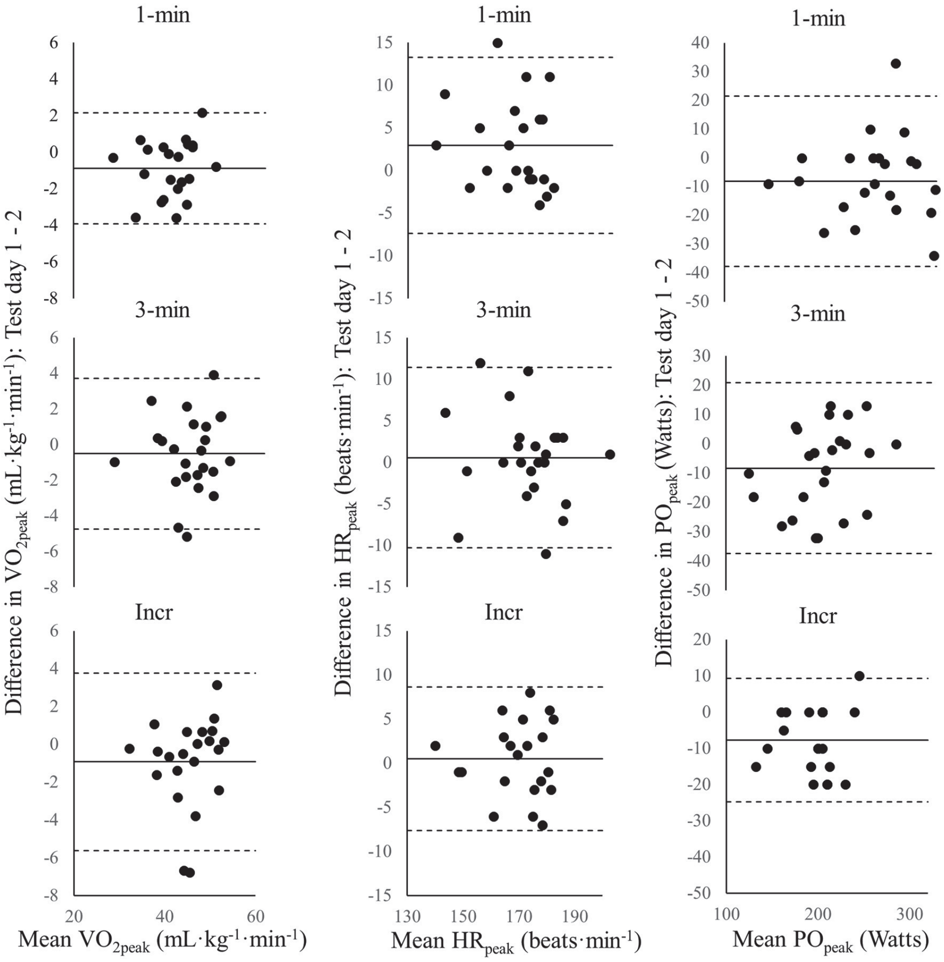


FIGURE 2 | Bland-Altman plots for the individual mean body-mass normalized peak oxygen uptake (VO_{2peak}), peak heart rate (HR_{peak}), and peak power output (PO_{peak}) of test day 1 and 2 vs. the difference between day 1 and 2 in VO_{2peak}, HR_{peak}, and PO_{peak} for the 1-min, the 3-min, and the incremental (Incr) test. The solid line is the group mean and the dotted lines indicate ±1.96·*SD*.

changed and those who maintained a stable pacing strategy across test days (**Supplementary Figures 2**, **3**).

There was no significant interaction between type of test and test order on VO_{2peak} ($p = 0.779$). Furthermore, VO_{2peak} did not differ between test day 1 and 2 for the 3-min test (below 1% change, $p > 0.05$) or the incremental test (~2%, $p = 0.068$ and 0.085), but increased significantly for the 1-min test (~2%, $p = 0.014$ and 0.007). PO_{peak} was significantly higher on test day 1 as compared to test day 2 for the 1-, the 3-min and the incremental test (2, 1, and 4%, all $p < 0.015$). In line with the increased PO_{peak}, time to exhaustion significantly increased in the incremental test on day 2 (from 326 ± 63 to 346 ± 70 s, $p = 0.003$).

DISCUSSION

The aim of this study was to compare VO_{2peak} and test-retest reliability of physiological parameters between a 1-min and a 3-min closed-end and an incremental open-end upper-body poling test. The incremental and the 3-min test resulted in significantly higher VO_{2peak} as compared to the 1-min test, with the incremental test inducing slightly higher VO_{2peak} than the 3-min test. High and very high ICCs across all physiological parameters (0.728–0.956) and PO_{peak} (0.923–0.955) were found for all three tests. The SDC, as a measure of absolute reliability, was consistently small for HR_{peak}, moderate for VO_{2peak} and PO_{peak}, but large for BLa_{peak} for all three tests. Furthermore, the

TABLE 1 | Power output, physiological and perceptual parameters of test day 1 and 2 for a 1-min, a 3-min, and an incremental upper-body poling test in able-bodied, upper-body trained participants (means ± *SD*).

	1-min			3-min			Incremental		
	Day 1	**Day 2**	***p*-value**	**Day 1**	**Day 2**	***p*-value**	**Day 1**	**Day 2**	***p*-value**
PO_{peak} (Watts)	234 ± 50	246 ± 45*	<0.001	179 ± 34	187 ± $33^{*,\dagger}$	<0.001	192 ± 29	200 ± $28^{*,\dagger,\ddagger}$	<0.001
VO_{2peak} ($mL \cdot kg^{-1} \cdot min^{-1}$)	40.0 ± 5.2	40.9 ± 5.0*	0.014	44.2 ± 5.7	44.7 ± $5.5^{\dagger}$	0.262	45.0 ± 5.8	45.9 ± $5.5^{\dagger,\ddagger}$	0.085
VO_{2peak} ($L \cdot min^{-1}$)	3.09 ± 0.42	3.17 ± 0.39*	0.007	3.40 ± 0.48	3.44 ± $0.46^{\dagger}$	0.270	3.46 ± 0.45	3.54 ± $0.49^{\dagger,\ddagger}$	0.068
VCO_{2peak} ($L \cdot min^{-1}$)	3.46 ± 0.60	3.56 ± 0.49	0.152	4.03 ± 0.63	4.12 ± $0.62^{\dagger}$	0.147	3.97 ± 0.52	4.19 ± $0.55^{*,\dagger}$	0.001
VE ($L \cdot min^{-1}$)	145 ± 32	144 ± 27	0.677	161 ± 29	162 ± $30^{\dagger}$	0.806	161 ± 23	165 ± $22^{*,\dagger}$	0.044
HR_{peak} ($beats \cdot min^{-1}$)	168 ± 11	165 ± 12*	0.016	172 ± 13	171 ± $14^{\dagger}$	0.611	171 ± 14	171 ± $14^{\dagger}$	0.578
BLa_{peak} ($mmol \cdot L^{-1}$)	11.0 ± 2.1	10.9 ± 2.5	0.868	11.6 ± 2.4	11.8 ± $2.2^{\dagger}$	0.489	11.4 ± 2.3	12.0 ± $2.2^{\dagger}$	0.166
RPE_O (6–20)	18.1 ± 1.3	17.8 ± 1.4	0.318	18.1 ± 1.0	18.3 ± $1.2^{\dagger}$	0.465	18.3 ± 0.9	18.3 ± 1.2	0.935
RPE_R (6–20)	17.6 ± 1.5	17.4 ± 1.6	0.554	17.7 ± 1.7	17.9 ± 1.8	0.484	17.7 ± 1.8	17.8 ± 1.7	0.544
RPE_M (6–20)	18.3 ± 1.1	18.2 ± 1.3	0.618	18.3 ± 1.1	18.4 ± 1.2	0.656	18.6 ± 0.9	18.4 ± 1.2	0.432

Calculations are based on data from 22 participants for the 1-min and the incremental test and 24 participants for the 3-min test.

Peak oxygen uptake (VO_{2peak}), peak carbon dioxide production (VCO_{2peak}), minute ventilation (VE), peak heart rate (HR_{peak}), peak blood lactate (BLa_{peak}), overall rate of perceived exertion (RPE_O), respiratory rate of perceived exertion (RPE_R), muscular rate of perceived exertion (RPE_M), peak power output (PO_{peak}).

**Significant differences from day 1 to day 2 at an alpha level of 0.05.*

† Mean value of day 1 and day 2 significantly different from 1-min test mean value at an alpha level of 0.05.

‡ Mean value of day 1 and day 2 significantly different from 3-min test mean value at an alpha level of 0.05

TABLE 2 | Interclass correlation coefficients (ICC) and [95% confidence interval (CI)], standard error of the measurement (SEM), smallest detectable change (SDC) of power output, physiological, and perceptual parameters for a 1-min, a 3-min and an incremental upper-body poling test in able-bodied, upper-body trained participants.

	1-min					3-min					Incremental				
	$ICC_{2,1}$	**95% CI**	**SEM**	**SDC**	**SDC%**	**$ICC_{2,1}$**	**95% CI**	**SEM**	**SDC**	**SDC%**	**$ICC_{2,1}$**	**95% CI**	**SEM**	**SDC**	**SDC%**
PO_{peak} (Watts)	0.946	[0.879–0.976]	10.7	19.4	7.6	0.923	[0.827–0.967]	10.5	19.0	9.4	0.955	[0.895-0.981]	6.1	11.1	5.7
VO_{2peak} ($mL \cdot kg^{-1} \cdot min^{-1}$)	0.956	[0.897–0.981]	1.1	2.0	4.9	0.942	[0.871–0.974]	1.5	2.8	6.2	0.933	[0.846-0.972]	1.7	3.1	6.7
VO_{2peak} ($L \cdot min^{-1}$)	0.952	[0.888–0.980]	0.1	0.2	5.4	0.949	[0.886–0.978]	0.1	0.2	6.2	0.938	[0.857–0.974]	0.1	0.2	6.9
VCO_{2peak} ($L \cdot min^{-1}$)	0.903	[0.781–0.959]	0.2	0.4	10.8	0.929	[0.843–0.969]	0.2	0.4	8.7	0.922	[0.822–0.967]	0.2	0.3	7.8
VE ($L \cdot min^{-1}$)	0.905	[0.786–0.959]	11.2	20.2	14.0	0.919	[0.822–0.964]	10.1	18.2	11.3	0.922	[0.822–0.967]	7.5	13.6	8.3
HR_{peak} ($beats \cdot min^{-1}$)	0.926	[0.831–0.969]	3.7	6.8	4.1	0.935	[0.856–0.971]	3.9	7.1	4.1	0.956	[0.897–0.981]	2.9	5.3	3.1
BLa_{peak} ($mmol \cdot L^{-1}$)	0.827	[0.629–0.924]	1.2	2.1	19.9	0.916	[0.816–0.963]	0.8	1.5	12.4	0.728	[0.450–0.877]	1.4	2.6	22.2
RPE_O (6–20)	0.707	[0.415–0.867]	0.9	1.6	8.9	0.833	[0.652–0.924]	0.6	1.1	5.8	0.429	[0.019–0.715]	0.9	1.7	9.1
RPE_R (6–20)	0.726	[0.447–0.876]	1.0	1.8	10.4	0.892	[0.767–0.952]	0.7	1.3	7.2	0.885	[0.744–0.951]	0.7	1.3	7.5
RPE_M (6–20)	0.580	[0.219–0.801]	0.9	1.6	8.9	0.791	[0.575–0.904]	0.6	1.2	6.3	0.679	[0.369–0.853]	0.8	1.4	7.4

Calculations are based on data from 22 participants for the 1-min and the incremental test and 24 participants for the 3-min test.

Peak oxygen uptake (VO_{2peak}), peak carbon dioxide production (VCO_{2peak}), minute ventilation (VE), peak heart rate (HR_{peak}), peak blood lactate (BLa_{peak}), overall rate of perceived exertion (RPE_O), respiratory rate of perceived exertion (RPE_R), muscular rate of perceived exertion (RPE_M), peak power output (PO_{peak}).

3-min closed-end test was more stable with respect to day-to-day differences in VO_{2peak} as compared to the incremental and 1-min test.

We found that the 3-min and the incremental test resulted in higher VO_{2peak} values than the 1-min test, demonstrating that 1-min duration is too short for the kinetics of the cardio-respiratory system to respond to the increased work demand during upper-body work. This is supported by the absence of a plateau in VO_{2peak} during the 1-min test in all participants. In contrast, a plateau or drop in VO_{2peak} at the end of the 3-min and the incremental test was observed in the majority of our participants' tests. Even though no study had previously compared a 1-min test to a 3-min or incremental protocol, Price et al. (2014) found significantly lower VO_{2peak} during a 30-s Wingate test as compared to an incremental protocol. Furthermore, the incremental protocol led to slightly higher VO_{2peak} values than the 3-min test, which is in line with a comparable study in cross-country skiing (McGawley, 2017). However, the meaningfulness of the 1 $mL \cdot kg^{-1} \cdot min^{-1}$ higher VO_{2peak} during the incremental test in the current study can be questioned, since both tests reach a plateau and the difference was in part influenced by the averaging procedure. In the current study the highest 30-s average was chosen to indicate VO_{2peak} as recommended by Robergs et al. (2010). If we shortened the duration to the single highest 10-s VO_{2peak} value, the difference between tests would have stayed stable but the peak values were consistently higher. In contrast, if VO_{2peak} would have been defined over two consecutive 30-s periods, VO_{2peak} differences between the 3-min test and the incremental test become negligible, yet the difference in VO_{2peak} between both these tests and the 1-min test would have increased. The latter is logical since there is a lag in the VO_2 kinetics response to the increased work demands included in the 1-min average. Thus, a 30-s average was deemed most appropriate in the current study to be able to compare tests, without taking the initial part of the test into consideration. Concluding from the above, the 1-min is not recommended as a VO_{2peak} test due to the clearly lower responses, whereas the 3-min test might slightly underestimate VO_{2peak}, with the magnitude depending on the averaging procedure.

Our finding of high relative reliability of physiological parameters of the three upper-body poling tests, reflected by high ICCs, are in line with several previous studies. Three minutes closed-end and incremental arm crank ergometry tests displayed similar ICCs in not specifically upper-body trained able-bodied participants (Leicht et al., 2009; Flueck et al., 2015; Hutchinson et al., 2017) as well as incremental wheelchair ergometry or treadmill tests in athletes with different disabilities (Bhambhani et al., 1991; Leicht et al., 2013). The current data shows that the ranks of the participants remain stable from test day 1 to test day 2 also during upper-body poling. However, caution is needed in the ICC's interpretation as it is a measure of the between-subjects variation in relation to the within-subjects variation and can be inflated merely by sample heterogeneity (Atkinson and Nevill, 1998; Hopkins, 2000). In a previous study on the reliability of VO_{2peak} during an incremental wheelchair treadmill test, Leicht et al. (2013) tried to circumvent a too large spread between participants by grouping together athletes with similar disabilities and training status, which consequently lead to small group sample sizes. To achieve a sufficient sample size, yet at the same time have a homogeneous sample, we chose to recruit upper-body trained male participants for our study. Given that the participants in our study were highly and relatively similarly upper-body trained, we expected them to be more homogeneous than athletes with a disability. However, the coefficient of variation of the body-mass normalized VO_{2peak} of 12% during the incremental test was higher than the 8% variation found in a group of participants with lower-limb disabilities (Leicht et al., 2013). As such, the interpretability of the ICC as a measure of test-retest reliability in upper-body testing remains limited, since even homogeneous able-bodied participants show heterogeneous responses.

In comparison to relative reliability outcomes, absolute reliability measures provide the possibility to investigate the degree to which repeated measurements vary for individuals. In this study, the small SDC for HR_{peak} and moderate SDC for VO_{2peak} and PO_{peak} indicate acceptable absolute reliability of all three peak tests. However, the rather large SDCs for BLa_{peak}, which are in line with previous studies (Leicht et al., 2009, 2013; Flueck et al., 2015), suggest that BLa_{peak} cannot be used as a reliable outcome measure in upper-body testing to exhaustion. That the SDC for absolute and body-mass normalized VO_{2peak} was only moderate can be explained by the higher values on day 2 for the 1-min and the incremental test, and for PO_{peak} for all three tests. The higher PO_{peak} and consequently higher VO_{2peak} values during the 1-min test on day 2 may be attributed to motivation to beat their previous score, although we have no data supporting this speculation. The higher VO_{2peak} during the incremental test on test day 2 can in part be explained by two participants having 0.5–0.6 $L{\cdot}min^{-1}$ and 7 $mL{\cdot}kg^{-1}{\cdot}min^{-1}$ higher absolute and body-mass normalized VO_{2peak}, respectively. If the data of the two participants were excluded, VO_{2peak} differences between test day 1 vs. 2 would have become non-significant. During the incremental test, the higher PO_{peak} on day 2 is related to half of the participants being able to sustain at least one extra 30-s stage with a higher PO_{peak} on day 2. Overall, the 3-min test is the most stable with respect to day-to-day differences and, therefore, the most reliable of the three upper-body poling tests.

Methodological Considerations

Our participants were highly trained for the poling movement and the exercise intensities used in this study, and they familiarized themselves with four times 5-min submaximal warm-up stages. We, therefore, chose to not perform a separate familiarization session for the peak tests in advance. However, in hindsight and as a recommendation for further studies, a separate familiarization session should be performed for all tests if the main outcome measure is PO_{peak} and for the 1-min and the incremental test if the main outcome measure is VO_{2peak}.

Furthermore, it remains to be investigated if other durations of the incremental test would result in different VO_{2peak} responses. As the participants in our study performed a thorough warm-up before starting the incremental test, we do not expect higher VO_{2peak} values with increases in duration of the test, but a follow-up study is needed to confirm this.

To be able to identify meaningful differences in body-mass normalized VO_{2peak} with paired comparison tests and similar participants in future studies, we estimated a sample size of 26 participants by $n = 8{\cdot}SDC^2{\cdot}(SEM^2)^{-1}$ as proposed by Hopkins (2000). Relatively similar numbers apply for most of the variables used in our approach. It is often challenging to recruit so many similarly upper-body trained participants, and in particular when aiming to test Paralympic athletes which are homogeneous with respect to their disability. Therefore, as large sample sizes as possible should be aimed at, if necessary through international collaborations. In addition, detailed description of the testing procedure and individual data should be made available so high-quality meta-analyses can be performed in the future.

CONCLUSION

In conclusion, we find acceptable absolute and relative reliability of a 1-min and a 3-min closed-end, and an incremental upper-body poling VO_{2peak} test in able-bodied, upper-body trained individuals. However, the 1-min test is not recommended as a VO_{2peak} test due to the clearly lower values than the 3-min and the incremental test. Whereas the 3-min test is more stable with respect to day-to-day differences in VO_{2peak}, the incremental test leads to slightly higher VO_{2peak}.

AUTHOR CONTRIBUTIONS

JKB, KS, and ØS substantially contributed to the conception and design of the study. JKB and KS acquired the data. JKB analyzed the data and all three authors were involved in the interpretation of data. JKB drafted the study with KS and ØS critically revising it for important intellectual content. The final version sent in for publication was approved by all three authors. Agreement to be accountable for all aspects of the work was reached between JKB, KS, and ØS.

FUNDING

This study was funded by the Centre for Elite Sports Research, Department of Neuromedicine and Movement Science, Norwegian University of Science and Technology, Trondheim, Norway.

ACKNOWLEDGMENTS

The eager participation of the participations is deeply appreciated. The authors acknowledge the financial support of the Centre for Elite Sports Research in Trondheim.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fphys.2017.00857/full#supplementary-material

Supplementary Figure 1 | Bland-Altman plots for the individual mean body-mass normalized peak oxygen uptake (VO_{2peak}) and peak power output (PO_{peak}) of test day 1 and 2 vs. the difference in VO_{2peak} and PO_{peak} between the 3-min and the incremental test. The solid line is the group mean and the dotted lines indicate ±1.96·*SD*.

Supplementary Figure 2 | Development of power output and VO_2 (presented as 30-s averages) of the 1-min test plotted individually for each participant over time. Solid lines demark the power output, dotted lines the VO_2. Black lines are for test day 1 and red lines for test day 2.

Supplementary Figure 3 | Development of power output and VO_2 (presented as 30-s averages) of the 3-min test plotted individually for each participant over time. Solid lines demark the power output, dotted lines the VO_2. Black lines are for test day 1 and red lines for test day 2.

REFERENCES

Atkinson, G., and Nevill, A. M. (1998). Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. *Sports Med.* 26, 217–238. doi: 10.2165/00007256-199826040-00002

Atkinson, G., Peacock, O., St. Clair Gibson, A., and Tucker, R. (2007). Distribution of power output during cycling: impact and mechanisms. *Sports Med.* 37, 647–667. doi: 10.2165/00007256-200737080-00001

Atkinson, G., and Reilly, T. (1996). Circadian variation in sports performance. *Sports Med.* 21, 292–312. doi: 10.2165/00007256-199621040-00005

Bar-Or, O., and Zwiren, L. D. (1975). Maximal oxygen consumption test during arm exercise–reliability and validity. *J. Appl. Physiol.* 38, 424–426.

Baumgart, J. K., and Sandbakk, Ø. (2016). Laboratory determinants of repeated-sprint and sport-specific-technique ability in World-class ice sledge hockey players. *Int. J. Sports Physiol. Perform.* 11, 182–190. doi: 10.1123/ijspp.2014-0516

Bhambhani, Y. N., Eriksson, P., and Steadward, R. D. (1991). Reliability of peak physiological responses during wheelchair ergometry in persons with spinal cord injury. *Arch. Phys. Med. Rehabil.* 72, 559–562.

Bland, J. M., and Altman, D. G. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1, 307–310. doi: 10.1016/S0140-6736(86)90837-8

Borg, G. A. (1982). Psychophysical bases of perceived exertion. *Med. Sci. Sports Exerc.* 14, 377–381. doi: 10.1249/00005768-198205000-00012

Flueck, J. L., Lienert, M., Schaufelberger, F., and Perret, C. (2015). Reliability of a 3-min all-out arm crank ergometer exercise test. *Int. J. Sports Med.* 36, 809–813. doi: 10.1055/s-0035-1548811

Hegge, A. M., Bucher, E., Ettema, G., Faude, O., Holmberg, H. C., and Sandbakk, O. (2015a). Gender differences in power production, energetic capacity and efficiency of elite cross-country skiers during whole-body, upper-body, and arm poling. *Eur. J. Appl. Physiol.* 116, 291–300. doi: 10.1007/s00421-015-3281-y

Hegge, A. M., Myhre, K., Welde, B., Holmberg, H. C., and Sandbakk, Ø. (2015b). Are gender differences in upper-body power generated by elite cross-country skiers augmented by increasing the intensity of exercise? *PLoS ONE* 10:e0127509. doi: 10.1371/journal.pone.0127509

Hopkins, W. G. (2000). Measures of reliability in sports medicine and science. *Sports Med.* 30, 1–15. doi: 10.2165/00007256-200030010-00001

Hutchinson, M. J., Paulson, T. A. W., Eston, R., and Goosey-Tolfrey, V. L. (2017). Assessment of peak oxygen uptake during handcycling: test-retest reliability and comparison of a ramp-incremented and perceptually-regulated exercise test. *PLoS ONE* 12:e0181008. doi: 10.1371/journal.pone.0181008

Leicht, A. S., Sealey, R. M., and Sinclair, W. H. (2009). The reliability of VO_{2peak} determination in healthy females during an incremental arm ergometry test. *Int. J. Sports Med.* 30, 509–515. doi: 10.1055/s-0029-1202351

Leicht, C. A., Tolfrey, K., Lenton, J. P., Bishop, N. C., and Goosey-Tolfrey, V. L. (2013). The verification phase and reliability of physiological parameters in peak testing of elite wheelchair athletes. *Eur. J. Appl. Physiol.* 113, 337–345. doi: 10.1007/s00421-012-2441-6

Losnegard, T., Myklebust, H., and Hallén, J. (2012). Anaerobic capacity as a determinant of performance in sprint skiing. *Med. Sci. Sports Exerc.* 44, 673–681. doi: 10.1249/MSS.0b013e3182388684

McGawley, K. (2017). The reliability and validity of a four-minute running time-trial in assessing [Formula: see text]max and performance. *Front. Physiol.* 8:270. doi: 10.3389/fphys.2017.00270

Moxnes, J. F., and Moxnes, E. D. (2014). Mathematical simulation of energy expenditure and recovery during sprint cross-country skiing. *Open Access J. Sports Med.* 5, 115–121. doi: 10.2147/OAJSM.S62020

O'Donoghue, P. (2013). *Statistics for Sport and Exercise Studies: An Introduction.* Abingdon: Routledge.

Plichta, S. B., Kelvin, E. A., and Munro, B. H. (2013). *Munro's Statistical Methods for Health Care Research.* Philadelphia, PA: Wolters Kluwer Health/Lippincott Williams & Wilkins.

Price, M., Beckford, C., Dorricott, A., Hill, C., Kershaw, M., Singh, M., et al. (2014). Oxygen uptake during upper body and lower body Wingate anaerobic tests. *Appl. Physiol. Nutr. Metab.* 39, 1345–1351. doi: 10.1139/apnm-2013-0405

Robergs, R. A., Dwyer, D., and Astorino, T. (2010). Recommendations for improved data processing from expired gas analysis indirect calorimetry. *Sports Med.* 40, 95–111. doi: 10.2165/11319670-000000000-00000

Roels, B., Schmitt, L., Libicz, S., Bentley, D., Richalet, J. P., and Millet, G. (2005). Specificity of V? o2max and the ventilatory threshold in free swimming and cycle ergometry: comparison between triathletes and swimmers. *Br. J. Sports Med.* 39, 965–968. doi: 10.1136/bjsm.2005.020404

Shephard, R. J., Vandewalle, H., Gil, V., Bouhlel, E., and Monod, H. (1992). Respiratory, muscular, and overall perceptions of effort: the influence of hypoxia and muscle mass. *Med. Sci. Sports Exerc.* 24, 556–567. doi: 10.1249/00005768-199205000-00010

Skovereng, K., Ettema, G., Welde, B., and Sandbakk, Ø (2013). On the relationship between upper-body strength, power, and sprint performance in ice sledge hockey. *J. Strength Cond .Res.* 27, 3461–3466. doi: 10.1519/JSC.0b013e31828f2799

Sperlich, B., Haegele, M., Thissen, A., Mester, J., and Holmberg, H. C. (2011). Are peak oxygen uptake and power output at maximal lactate steady state obtained from a 3-min all-out cycle test? *Int. J. Sports Med.* 32, 433–437. doi: 10.1055/s-0031-1271770

Vesterinen, V., Mikkola, J., Nummela, A., Hynynen, E., and Hakkinen, K. (2009). Fatigue in a simulated cross-country skiing sprint competition. *J. Sports Sci.* 27, 1069–1077. doi: 10.1080/02640410903081860

Weir, J. P. (2005). Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. *J. Strength Cond. Res.* 19, 231–240. doi: 10.1519/15184.1

Weltman, A., Stamford, B. A., and Fulco, C. (1979). Recovery from maximal effort exercise: lactate disappearance and subsequent performance. *J. Appl. Physiol. Respir. Environ. Exerc. Physiol.* 47, 677–682.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Corrigendum: Comparison of Peak Oxygen Uptake and Test-Retest Reliability of Physiological Parameters between Closed-End and Incremental Upper-Body Poling Tests

Julia K. Baumgart, Knut Skovereng, Øyvind Sandbakk

Department of Neuromedicine and Movement Science, Centre for Elite Sports Research, Norwegian University of Science and Technology, Trondheim, Norway

*** Correspondence**: jk.baumgart@gmail.com

Keywords: peak aerobic capacity, endurance performance, all-out, 3-min, exhaustion

Corrigendum on: Baumgart, J.K., Skovereng, K., and Sandbakk, Ø. (2017). Comparison of Peak Oxygen Uptake and Test-Retest Reliability of Physiological Parameters between Closed-End and Incremental Upper-Body Poling Tests. *Frontiers in Physiology* 8(857). doi: 10.3389/fphys.2017.00857.

Error in Table

In the original article, there was a mistake in **Table 1** as published. The mistake concerns the peak power output values provided for the 1-min and the 3-min test. We initially based the calculations on the mean peak power output of the 1-min and the 3-min test on the values provided by the internal software of the Concept2 ski ergometer, which are cumulative averages (i.e. the first average is an average over the first 30 s, the second average is an average over the first minute, the third over one and a half minutes and so forth). However, when submitting this manuscript we recalculated the mean peak power output to reflect 30-s averages that are not cumulative and hence independent of the power output produced in the previous 30-s period.

The corrected **Table 1** appears below. The authors apologize for this error and state that this does not change the scientific conclusions of the article in any way.

Text Correction

In the original article, there was an error. The mistake is in line with what is described in the above.

A correction has been made to the section **Results**, subsection Comparison of Tests, paragraph 1:

Based on the average values of test day 1 and 2, the incremental (45.4 ± 5.5 $mL·kg^{-1}·min^{-1}$, 196 ± 28 Watt) and the 3-min test (44.5 ± 5.5 $mL·kg^{-1}·min^{-1}$, 201 ± 36 Watt) resulted in significantly higher VO_{2peak} and lower PO_{peak} as compared to the 1-min test (40.4 ±5.0 $mL·kg^{-1}·min^{-1}$, 256 ± 47 Watt) (all $p < 0.001$). Additionally, the incremental test resulted in significantly higher VO_{2peak} ($p = 0.03$).

The authors apologize for this error and state that this does not change the scientific conclusions of the article in any way.

Table 1. Power output, physiological and perceptual parameters of test day 1 and 2 for a 1-min, a 3-min and an incremental upper-body poling test in able-bodied, upper-body trained participants (means ± SD). Calculations are based on data from 22 participants for the 1-min and the incremental test and 24 participants for the 3-min test.

	1-min			3-min			Incremental		
	Day 1	**Day 2**	**p-value**	**Day 1**	**Day 2**	**p-value**	**Day 1**	**Day 2**	**p-value**
Power output (Watt)	254±46	259±47*	<0.001	198±40	203±33*,†	<0.001	192±29	200±28*,†	<0.001
VO_{2peak} ($mL \cdot kg^{-1} \cdot min^{-1}$)	40.0±5.2	40.9±5.0*	0.014	44.2±5.7	44.7±5.5†	0.262	45.0±5.8	45.9±5.5†,‡	0.085
VO_{2peak} ($L \cdot min^{-1}$)	3.09±0.42	3.17±0.39*	0.007	3.40±0.48	3.44±0.46†	0.270	3.46±0.45	3.54±0.49†,‡	0.068
VCO_{2peak} ($L \cdot min^{-1}$)	3.46±0.60	3.56±0.49	0.152	4.03±0.63	4.12±0.62†	0.147	3.97±0.52	4.19±0.55*,†	0.001
VE ($L \cdot min^{-1}$)	145±32	144±27	0.677	161±29	162±30†	0.806	161±23	165±22*,†	0.044
HR_{peak} ($beats \cdot min^{-1}$)	168±11	165±12*	0.016	172±13	171±14†	0.611	171±14	171±14†	0.578
BLa_{peak} ($mmol \cdot L^{-1}$)	11.0±2.1	10.9±2.5	0.868	11.6±2.4	11.8±2.2†	0.489	11.4±2.3	12.0±2.2†	0.166
RPE_O (6-20)	18.1±1.3	17.8±1.4	0.318	18.1±1.0	18.3±1.2†	0.465	18.3±0.9	18.3±1.2	0.935
RPE_R (6-20)	17.6±1.5	17.4±1.6	0.554	17.7±1.7	17.9±1.8	0.484	17.7±1.8	17.8±1.7	0.544
RPE_M (6-20)	18.3±1.1	18.2±1.3	0.618	18.3±1.1	18.4±1.2	0.656	18.6±0.9	18.4±1.2	0.432

Peak oxygen uptake (VO_{2peak}), peak carbon dioxide production (VCO_{2peak}), minute ventilation (VE), peak heart rate (HR_{peak}), peak blood lactate (BLa_{peak}), overall rate of perceived exertion (RPE_O), respiratory rate of perceived exertion (RPE_R), muscular rate of perceived exertion (RPE_M)

* Significant differences from day 1 to day 2 at an alpha level of 0.05
† mean value of day 1 and day 2 significantly different from 1-min test mean value at an alpha level of 0.05
‡ mean value of day 1 and day 2 significantly different from 3-min test mean value at an alpha level of 0.05

Paper IV

Examination of gas exchange and blood lactate thresholds in Paralympic athletes during upper-body poling

Authors

Julia Kathrin Baumgart[1,*], Maaike Moes[2], Knut Skovereng[1], Gertjan Ettema[1], Øyvind Sandbakk[1]

Institutions

[1] Centre for Elite Sports Research, Department of Neuroscience and Movement Science, Faculty of Medicine and Health Sciences, Norwegian University of Science and Technology, Trondheim, Norway

[2] Department of Human Movement Sciences, Faculty of Health, Medicine and Life Sciences, Maastricht University, Maastricht, The Netherlands

***Corresponding author**

Email: jk.baumgart@gmail.com

Abstract

Objectives: The primary aim was to compare physiological and perceptual outcome parameters identified at common gas exchange and blood lactate (BLa) thresholds in Paralympic athletes while upper-body poling. The secondary aim was to compare the fit of the breakpoint models used to identify threshols in the gas exchange thresholds data versus continuous linear and curvilinear (no-breakpoint) models.

Methods: Fifteen elite ice sledge hockey players performed seven to eight 5-min stages at increasing workload until exhaustion during upper-body poling. Two regression lines were fitted to the oxygen uptake (VO_2)-carbon dioxide (VCO_2) and minute ventilation (VE)/VO_2 data to determine the ventilatory threshold (VT), and to the VCO_2-VE and VE/VCO_2 data to determine the respiratory compensation threshold (RCT). The first lactate threshold (LT1) was determined by the first rise in BLa (+0.4 $mmol·L^{-1}$ and +1.0 $mmol·L^{-1}$) and a breakpoint in the log-log transformed VO_2-BLa data, and the second lactate threshold (LT2) by a fixed rise in BLa above 4 $mmol·L^{-1}$ and by employing the modified D_{max} method. Paired-samples t-tests were used to compare the outcome parameters within and between the different threshold methods. The fit of the two regression lines (breakpoint model) used to identify thresholds in the gas exchange data was compared to that of a single regression line, an exponential and a 3rd order polynomial curve (no-breakpoint models).

Results: All outcome parameters identified with the VT (i.e., breakpoints in the VO_2-VCO_2 or VE/VO_2 data) were significantly higher than the ones identified with a fixed rise in BLa (+0.4 or +1.0 $mmol·L^{-1}$) at the LT1 (e.g. BLa: 5.1±2.2 or 4.9±1.8 vs 1.9±0.6 or 2.3±0.5 $mmol·L^{-1}$, $p<0.001$), but were not significantly different from the log-log transformed VO_2-BLa data (4.3±1.6 $mmol·L^{-1}$, $p<0.13$). The outcome parameters identified with breakpoints in the VCO_2-VE data to determine the RCT (e.g. BLa: 5.5±1.4 $mmol·L^{-1}$) and with the modified D_{max} method at the LT2 (5.5±1.1 $mmol·L^{-1}$) were higher compared to parameters identified with VE/VCO_2 method (4.9±1.5 $mmol·L^{-1}$) and a fixed BLa value of 4 $mmol·L^{-1}$ (all $p<0.03$).

Although we were able to determine the VT and RCT via different gas exchange threshold methods with good fit in all 15 participants (mean $R^2 > 0.931$), the continuous no-breakpoint models had the highest probability of being the best models for the VO_2-VCO_2 and the VCO_2-VE data (>68%).

Conclusions: In Paralympic athletes who exercise in the upper-body poling mode, the physiological and perceptual outcome parameters identified at the VT and the ones identified with fixed methods at the LT1 showed large differences, demonstrating that these cannot be used interchangeably to estimate the aerobic threshold. In addition, the close location of the VT, RCT and LT2 does not allow us to distinguish the aerobic and anaerobic threshold, indicating the presence of only one threshold in athletes with a disability exercising in an upper-body mode. Furthermore, the better fit of continuous no-breakpoint models indicates no presence of clear breakpoints in the gas exchange data for most participants. This makes us question if breakpoints used to determine thresholds in the gas exchange data really exist in an upper-body exercise mode in athletes with disabilities.

Key words: aerobic threshold, anaerobic threshold, ventilatory threshold, respiratory compensation threshold, lactate threshold, disability

Introduction

In able-bodied endurance athletes performing lower-body or whole-body exercise, gas exchange and blood lactate (BLa) threshold concepts are well-established in the diagnosis of endurance performance as well as in the prescription of systematic training with different exercise intensity zones [1]. Two thresholds are commonly described in the literature: 1) The aerobic threshold (AT) – determined by the ventilatory threshold (VT) or the first lactate threshold (LT1) – separates low- from moderate-intensity exercise [2, 3]. 2) The anaerobic threshold (ANT) – determined by the respiratory compensation threshold (RCT) or the second lactate threshold (LT2) – separates moderate- from high-intensity exercise [2, 3]. However, to what extent the outcome parameters identified at the VT and LT1 as well as the RCT and LT2 coincide in Paralympic sitting sport athletes who exercise in an upper-body mode remains to be investigated.

Various methods have been employed to determine the VT and the RCT, as well as the LT1 and the LT2 [3-6]. The VT is based on a disproportionate increase (i.e. a breakpoint) in carbon dioxide production (VCO_2) and minute ventilation (VE) in relation to oxygen uptake (VO_2) [3, 7], and the LT1 on an onset in BLa concentration above resting levels that marks the beginning of exercise [5] or on a breakpoint in the log-log transformed VO_2-BLa data [4]. Even though these physiological changes occur above the VT and LT1, the body is still able to maintain equilibrium at intensities up to the ANT, and aerobic metabolism (indicated by measurements of oxygen uptake and the corresponding energy equivalent) reflects overall energy expenditure [2]. The ANT marks the point beyond which any attempt of the body to maintain metabolic equilibrium at a constant rate of work fails [6]. The RCT is based on a disproportionate increase (i.e. a breakpoint) of VE in relation to VCO_2 [3], a mechanism that has been suggested to correspond with the point where BLa starts to accumulate with constant workload [6]. In contrast, it has been argued that the changes in gas exchange with increasing work rate are continuous transitions where fatigue gradually accumulates rather than clear breakpoints [8].

The assumption that the VT corresponds with the LT1, and the RCT with the LT2, are based on the initial studies by Beaver et al. [3] and Wassermann et al. [6, 9, 10] from the 1980's. However, there has been a continuous debate around the existence of and the physiological link between these different thresholds [2, 11-14]. Although physiological parameters identified at the VT and LT1, and at the RCT and LT2 have shown high correlations in able-bodied participants during cycling and running in some studies [15, 16], others find low correlations [17]. In wheelchair basketball and wheelchair rugby athletes with a spinal cord injury, the % of VO_{2peak} was lower at the LT1 compared to the VT, whereas it did not significantly differ at the LT2 and RCT [18]. In contrast, in able-bodied swimmers, there were no significant differences in physiological outcome parameters at the LT1 and the VT [19].

Whereas a range of studies have investigated the VT during upper-body exercise in able-bodied participants and participants with a disability [20-29], knowledge is limited on whether gas exchange and BLa threshold concepts can be used interchangeably in athletes with disabilities who exercise in an upper-body mode, or whether breakpoints exist in the gas exchange data of these athletes. Therefore, the primary aim of this study was to compare physiological and perceptual outcome parameters at the gas exchange and BLa thresholds in the data obtained from Paralympic athletes while upper-body poling. The secondary aim was to compare the fit of breakpoint models used to identify gas exchange thresholds with continuous linear or curvilinear (no-breakpoint) models.

Methods

Participants

Fourteen male and one female endurance-trained Norwegian Para ice hockey players participated in this study. Anthropometrics and training hours per month of the participants are depicted in Table 1. All participants were healthy and free of injuries at the time of testing. The study was approved by the Norwegian Data Protection Authority and conducted in accordance with the Declaration of Helsinki. All participants signed an informed consent form prior to voluntarily taking part in the study, and were made aware that they could withdraw from the study at any point without providing an explanation.

Table 1. Sex, age, anthropometric and disability characteristics as well as monthly training hours of 15 Norwegian national team Para ice hockey players participating in this study.

	Sex	Age (years)	Body mass (kg)	Height (cm)	Disability (level of injury)	Training hrs/month
1	Male	53	83.3	186	Paraplegia (Th12-L1)	25
2	Male	18	75.7	160	Spina bifida (L5)	49
3	Male	27	61.0	160	Athrogryposis multiplex congenita	63
4	Male	31	69.4	184	Hereditary spastic paraplegia	45
5	Male	28	90.0	173	Paraplegia (Th10)	26
6	Male	21	70.4	164	Spina bifida (ns)	59
7	Male	33	70.5	160	Spina bifida (Th12)	67
8	Male	34	75.3	173	Paraplegia (Th11-12)	48
9	Female	22	70.0	167	Spina bifida (L3-S1)	33
10*	Male	22	63.4	164	Paraplegia (Th11-12)	28
11*	Male	18	64.2	154	Spina bifida (ns)	54
12	Male	20	68.0	186	Paraplegia (Th12)	40
13	Male	20	77.0	163	Cerebral Palsy (motor only)	23
14	Male	28	66.5	173	Amputation (single leg above the knee)	80
15	Male	32	63.2	165	Paraplegia (ns)	56
Mean ± SD		27.1±8.9	71.2±8.0	170±10	-	47±18

* Players are from the Norwegian national B-team

Thoracic (Th), lumbar (L), sacral (S), not specified (ns)

Experimental design

The testing consisted of two consecutive test days at similar test times, during which participants performed an incremental test to exhaustion on day one, followed by seven to eight 5-min stages at gradually increasing effort for each stage until exhaustion on day two. All tests were performed in upper-body poling on a Concept2 ski ergometer 1 (Concept2, Inc., Morrisville, USA, http://www.concept2.com/service/skierg/skierg-1), while sitting in an ice sledge hockey seat.

Test set-up

After being equipped with an oro-nasal mask (Hans Rudolph Inc, Kansas City, MO, USA) and a heart rate monitor (Polar Electro Inc., Port Washington, NY, USA), the participants were tightly strapped around the thighs and hips into an ice sledge hockey seat that was mounted on a wooden platform (Fig 1). The distance of the seat to the Concept2 ski ergometer and the position of the feet depended on personal preference but was the same for test day one and two. The ski ergometer uses wind resistance, which is generated by the spinning flywheel. The ski ergometer has a spiral damper with settings from one to ten, which works like a gearing system. We had this damper set at "eight" for all participants. Power output was measured with the ergometer's software, which was previously validated with force and velocity measurements using a force cell (Noraxon USA inc., Scottsdal, AZ, USA) and the Oqus cameras of the Qualisys motion capture system (Qualisys AB, Gothenburg, Sweden) as described by Hegge et al. [30]. The Metamax II ergospirometer CORTEX Biophysik GmbH, Leipzig, Germany) was calibrated against a known mixture of gases (16% O_2 and 4% CO_2) and ambient air prior to the testing procedure of every second participant. Before each athlete was tested, the flow transducer was calibrated with a 3 L syringe and then connected to the oro-nasal mask, which allowed for the measurement of breath-by-breath respiratory parameters.

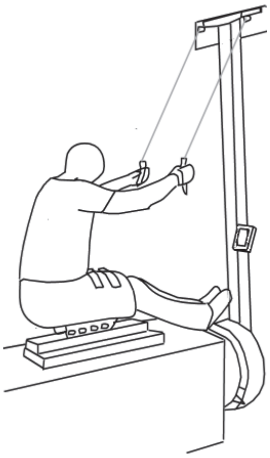


Fig 1. **Test set-up.** The participants were strapped in around the hips and thighs in an ice sledge hockey seat mounted on a platform in front of the Concept2 ski-ergometer.

Test protocol

The participants were instructed to refrain from heavy training and alcohol consumption 24 hours before, caffeine intake the day of, and food intake two hours before testing. Additionally, the participants were instructed to void their bladder directly before arriving at the laboratory. A questionnaire was filled out on each of the two test days to monitor if the participants followed these instructions, as well as to exclude any prior illness or injury that might have interfered with the testing.

Test day one

A standardized warm-up of five 5-min submaximal stages with a 2- to 3-min break between stages was performed in the upper-body poling mode at an overall rating of perceived exertion (RPE) of 7 (very light), 9 (very light), 11 (light), 13 (somewhat hard) and 15 (hard). Next to serving as a warm-up, the submaximal stages were used to familiarize the participants with the use of the Borg scale [31] to indicate RPE after the incremental test and each of the 5-min stages on day two. After a 5-min break, the incremental test started at the individual power output of the third submaximal stage (rounded to the nearest 10-point value), and participants were instructed to continuously increase power output by 10 W every 30 s. The test was terminated when the participant, despite strong verbal encouragement, could no longer maintain the required power output of the 30-s stage and the VO_2 values either plateaued or decreased (a drop of more than 2 $mL \cdot kg^{-1} \cdot min^{-1}$). After the incremental test, participants recovered passively for five min and actively for three min (at the power output of the first submaximal stage). They then performed a verification stage at a 10% higher power output than the peak power output of the incremental test (rounded to the nearest 10-point value) to verify the attainment of a true VO_{2peak} [32]. The verification stage was terminated when the participant dropped more than 10% of target power output for more than five s.

Test day 2

Seven to eight 5-min stages were performed with a 2- to 3-min break between stages and in the same upper-body poling mode. The first stage started at 20% of the individual peak power output obtained during the incremental test on day one, with increases of 10% (of the individual peak power) for each consecutive stage. The last stage was terminated when the participant, despite strong verbal encouragement, could no longer maintain the power output of that stage and dropped more than 10% in the target power output for longer than five s. The intermittent exercise protocol was chosen to take a BLa sample from the fingertip in between stages. The duration of five min per stage was chosen, since in an upper-body mode two to three min are needed to achieve steady-state of physiological outcome parameters[33].

Outcome measurements

Heart rate was measured every second with a heart rate monitor (Polar Electro Inc., Port Washington, NY, USA), and respiratory parameters (i.e., VO_2, VCO_2, VE, and respiratory exchange ratio (RER)) were measured breath-by-breath and averaged over 10 s by the in-built software of a Metamax II. A blood sample was taken from the fingertip and BLa analysed with a Lactate Pro device (Arkray Inc., Japan) at rest and directly after each of the submaximal stages on day one and day two, and one and three min after the incremental test and the verification stage as well as the last stage of day two. Overall RPE was recorded after each of the submaximal stages on day one and two, as well as after the incremental test on day one and the last stage on day two. Power output was displayed per stroke and saved as 20-s averages during the submaximal stages on day one and day two by the in-built Concept2 software (Concept2, Morrisville, VT, USA). Peak power output during the incremental test and during the verification stage was registered as the highest 30-s average.

Data analysis

Data processing

Peak power output and gas exchange outcome parameters were calculated as the highest 30-s moving average and HR_{peak} as the highest 3-s moving average of the incremental test performed on test day one. The gas exchange, heart rate and power output data of the last two min (12 x 10-s averages) of each complete 5-min stage conducted on test day two was included for data analysis in MATLAB (R2016a; Mathworks Inc., Natick, MA). The analyses in the following were based on the concatenated 2-min gas exchange data for the VT and RCT and on the BLa values after each 5-min stages for the LT1 and LT2.

Different methods were used to determine both the VT and the RCT, as well as the LT1 and the LT2. For the determination of the VT, VO_2 was plotted against VCO_2 (V-slope method) [3] as well as the stages against VE/VO_2 and VE/VCO_2 (ventilatory equivalent method) [7] and two regression lines fit to the data.

For a valid detection of the VT with the ventilatory equivalent method, the VE/VO$_2$ had to increase before an increase in VE/VCO$_2$ [15, 34]. For the detection of the RCT, VCO$_2$ was plotted against VE [3] and two regression lines fit to the data. The LT1 was determined as the first fixed rise in BLa concentration by 0.4 and 1 mmol·L^{-1} above the lowest individual BLa value [5, 35]. Additionally, the LT1 was determined by breakpoints in the log-log transformed VO$_2$-BLa relationship [4]. The LT2 was determined by a fixed BLa concentration of 4 mmol·L^{-1} [36]. Additionally, the LT2 was determined by the modified D$_{max}$ method, which identifies the point on the 3rd order polynomial curve fitted to the BLa values that yields the maximal perpendicular distance to the straight line formed by the first stage with an increase of 0.4 mmol·L^{-1} and the BLa measured after the last stage [5]. Outcome parameters (% of peak power output, % of VO$_{2peak}$, % of HR$_{peak}$, as well as BLa and RPE) were interpolated at the thresholds identified with each of the above described methods used to determine the VT, LT1, RCT and LT2.

Statistical analyses

Paired-samples t-tests were used to compare the physiological and perceptual outcome parameters within the VT, LT1, RCT and LT2, and between all four different thresholds. Pearson's r was used to investigate relationships between the outcome parameters identified with the different methods used to determine VT, LT1, RCT and LT2. Ranges of 0.26-0.49, 0.50-0.69, 0.70-0.89 and 0.90-1.0 were used to indicate low, moderate, high and very high correlations according to Munro's criteria [37]. An α level of 0.05 was used to indicate statistical significance.

To compare the fit of breakpoint models versus continuous linear or curvilinear (no-breakpoint) models to the gas exchange data, two regression lines (equation 1) versus a single linear regression line (equation 2), an exponential curve (equation 3), and a 3rd order polynomial curve (equation 4) were fitted to the VO$_2$-VCO$_2$, VE/VO$_2$, VE/VCO$_2$, and VCO$_2$-VE data by linear least squares fitting.

$$y = \begin{cases} a_1 + b_1\, x, & t < k \\ a_2 + b_2\, x, & t \geq k \end{cases} \tag{1}$$

$$y = a \ + b\, x \tag{2}$$

$$y = a + c \cdot exp\left(\frac{x + g}{d}\right) \tag{3}$$

$$y = a \ + b_1\, x + b_2\, x^2 + b_3\, x^3 \tag{4}$$

y is the variable of interest, a the y-axis offset, b the slope coefficients, c and d spreading coefficients, g the x-axis offset and k the point where the first and the second regression line of the piecewise function cross. To compare the fit of the four models, the Akaike information criterion (AIC) (equation 5) [38] and the Akaike weights (w_i) (equation 7) for each model i relative to the set of R candidate models were calculated based on the delta AIC (Δ_i) (equation 6) [39, 40].

$$AIC = n \cdot log\left(\frac{SS_{er}}{n}\right) + 2 \cdot K \tag{5}$$

$$Delta\ AIC = \ \Delta_i = AIC_i - AIC_{2reg} \tag{6}$$

$$AIC\ weight = w_i = \frac{exp^{(-\frac{\Delta i}{2})}}{\sum_{r=1}^{R} exp^{(-\frac{\Delta r}{2})}} \tag{7}$$

n is the number of data points, SS_{er} the error sums of squares, and K the number of parameters +1 of each model. Our rationale was that a better fit of the two regression lines (breakpoint model) as compared to the linear/curvilinear models (continuous no-breakpoint models), would suggest the presence of a breakpoint.

Results

The outcome parameters identified with different methods used to determine the VT, LT1, RCT and LT2 (Fig 2) are presented as percentage of the respective peak power output, VO_{2peak} and peak HR obtained during the incremental test (Table 2).

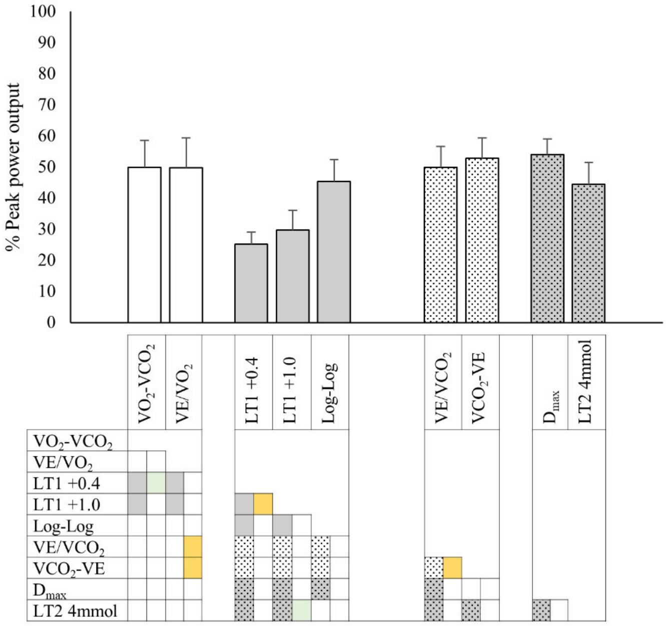


Fig 2. Outcome parameters at the VT, LT1, RCT and LT2. Outcome parameters at the aerobic threshold are determined by the V-slope (VO_2-VCO_2 data) and the ventilatory equivalent method (VE/VO_2 data) to identify the VT, and a fixed rise in BLa of 0.4 and 1.0 mmol·L^{-1} and the log-log transformed VO_2-BLa data to identify the LT1. Outcome parameters at the anaerobic threshold are determined by the ventilatory equivalent method (VE/VCO_2 data) and the respiratory compensation point (VCO_2-VE data) to identify the RCT, and the modified D_{max} method and a fixed BLa value of 4 mmol·L^{-1} to identify the LT2. The data used is from 15 elite ice sledge hockey players following a protocol with stepwise increases in workload every 5 min while upper-body poling.

The data presented is the mean of all 15 participants and error bars denote +1 SD.

Filled squares in the left columns denote significant differences between methods at an alpha level of 0.05.

Filled squares in the right columns denote significant correlations of methods (green = moderate, yellow = high, orange = very high).

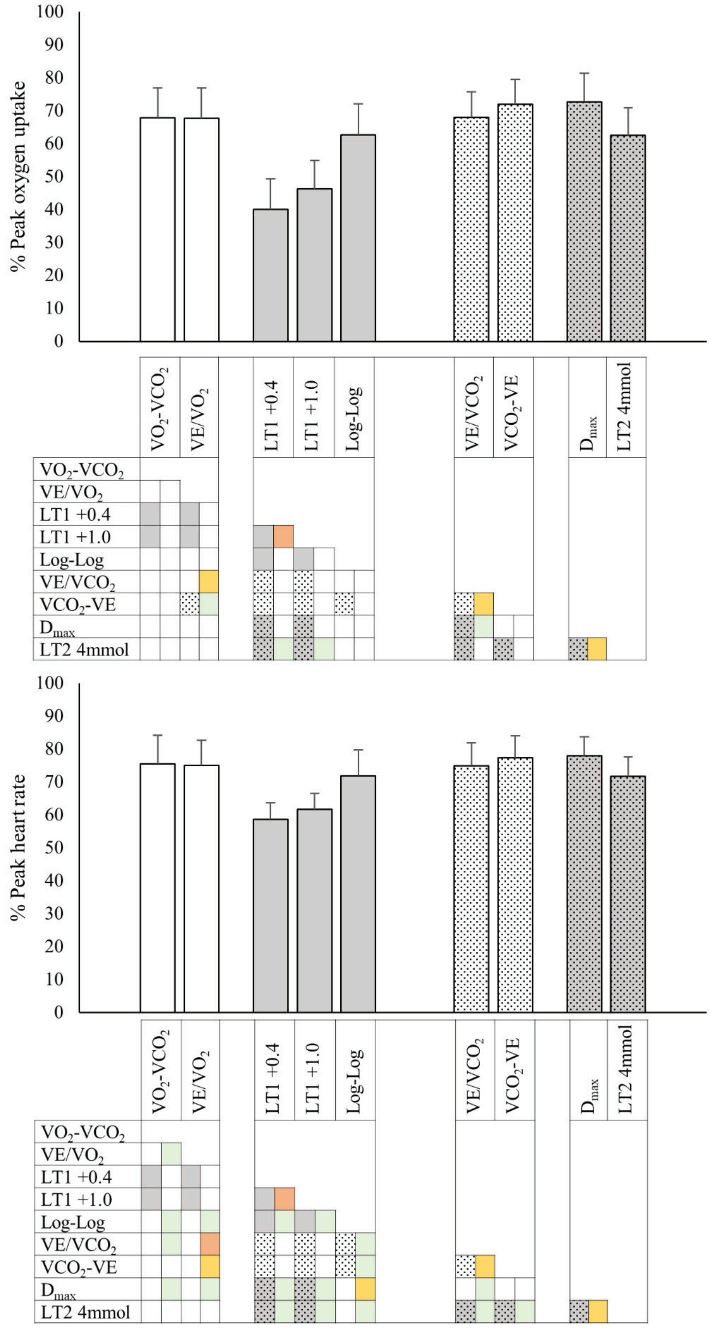


Fig 2 (continued).

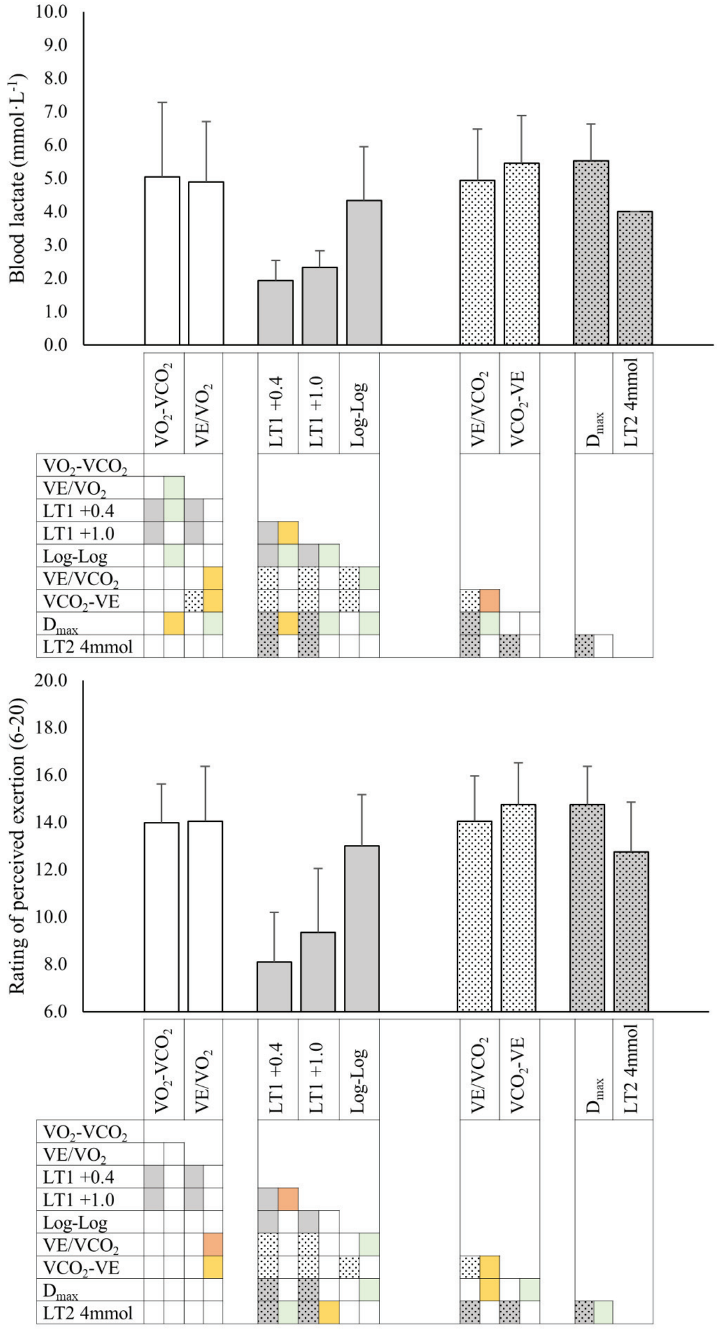


Fig 2 (continued).

Table 2. Mean ± SD (95% CI) peak power output and peak physiological and perceptual outcome parameters.

	Peak values
Peak power output (W)	144 ± 37 (125-163)
VO_{2peak} ($mL \cdot kg^{-1} \cdot min^{-1}$)	36 ± 7 (32-39)
HR_{peak} ($beats \cdot min^{-1}$)	188 ± 12 (182-194)
Blood lactate ($mmol \cdot L^{-1}$)	14.4 ± 1.5 (13.7-15.2)
RPE (6-20)	19.7 ± 0.5 (19.4-19.9)

The data was collected during an incremental test to exhaustion while upper-body poling of 15 Norwegian ice sledge hockey players.

Peak oxygen uptake (VO_{2peak}), peak heart rate (HR_{peak}), rating of perceived exertion (RPE)

All outcome parameters identified at VT with either the V-slope or the ventilatory equivalent method were significantly higher than the ones at both the LT1 (+0.4) and LT1 (+1.0) (all $p<0.001$), but not significantly different from the ones identified with the log-log transformed VO_2-BLa method (all $p>0.06$) (Fig 2). Additionally, most of the outcome parameters identified at the VT did not significantly correlate with the corresponding ones at LT1 (+0.4) or LT1 (+1.0) (exception: BLa at LT1 (+0.4): $r=0.55$, $p=0.03$; all other outcome parameters: $r<0.38$, $p>0.17$) (S1 Excel file, sheet "correlations"). All outcome parameters at LT1 (+0.4) and LT1 (+1.0) were highly or very highly correlated (all $r>0.83$, $p<0.001$). In addition, some of the outcome parameters identified with breakpoints in the log-log transformed VO_2-BLa moderately correlated with the outcome parameters identified by the V-slope method (HR: $r=0.64$, $p=0.01$; BLa: $r=0.54$, $p=0.04$) and the breakpoints in the VE/VO_2 data of the ventilatory equivalent method (HR: $r=0.54$, $p=0.04$).

The outcome parameters identified with breakpoints in the VCO_2-VE data at the RCT (e.g. BLa: 5.5±1.4 $mmol \cdot L^{-1}$) and with the modified D_{max} method at the LT2 (5.5±1.1 $mmol \cdot L^{-1}$) were higher compared to parameters identified with VE/VCO_2 method (4.9±1.5 $mmol \cdot L^{-1}$) and a fixed BLa value of 4 $mmol \cdot L^{-1}$ (all $p<0.03$). Furthermore, there was no significant difference between the outcome parameters identified with V-slope method used to determine the VT and the ones identified with breakpoints in the VE/VCO_2 data used to determine the RCT ($p>0.22$). However, most outcome parameters identified at the breakpoints in the VE/VO_2 and VE/VCO_2 data (ventilatory equivalent method) were highly or very highly correlated with those identified at the breakpoints in the VCO_2-VE data (RCT) (exception: % of VO_{2peak} $r=0.67$, $p=0.006$; all other outcome parameters: $r>0.73$, $p<0.01$) (Fig 2). In addition, most outcome parameters identified at the thresholds in the VE/VCO_2 data were moderately to highly correlated with the same outcome parameters at the thresholds identified with the modified D_{max} method (exception: % of peak power output: $r=0.44$, $p=0.10$; all other outcome parameters: $r>0.57$, $p<0.03$). Furthermore, there was no significant difference between the outcome parameters identified with the log-log transformed VO_2-BLa method used to determine the LT1 and at a fixed BLa concentration of 4 mmol·L^{-1} used to determine the LT2 and ($p>0.43$).

For the gas exchange data, all fitting procedures for the VO_2-VCO_2 and the VCO_2-VE plots, including the single linear regression line, showed very good fit on the data for all 15 participants (mean $r^2>0.97$) (Table 3). However, the fit of the breakpoint model compared to the continuous no-breakpoint models on the VO_2-VCO_2 and the VCO_2-VE data was only better among five participants. Accordingly, the continuous no-breakpoint models had 71% and 68% probability of being the best models for the VO_2-VCO_2 and the VCO_2-VE data, respectively (Table 4). Exemplary VO_2-VCO_2 and VCO_2-VE plots are illustrated in Fig 3 and 4, respectively.

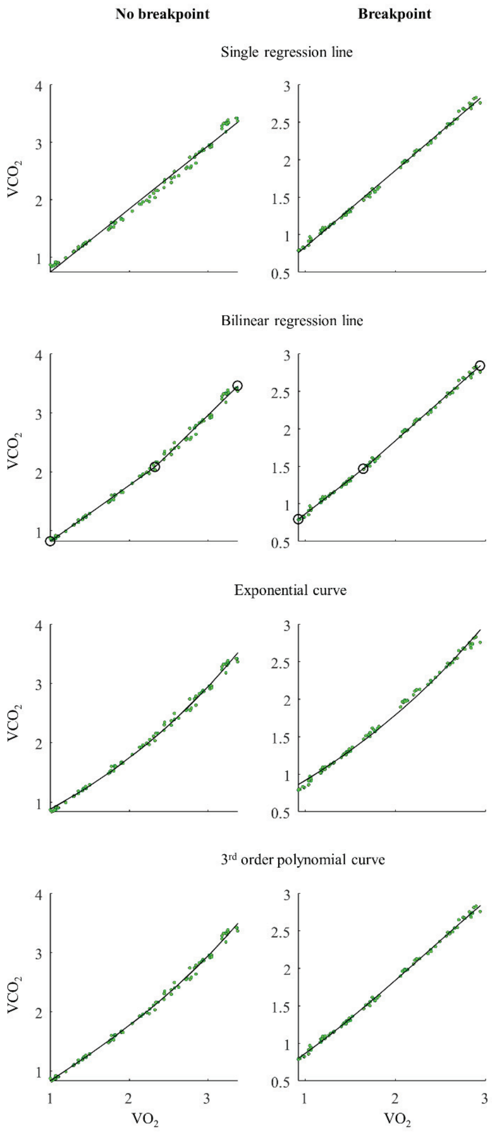


Fig 3. Exemplary VO_2-VCO_2 plots. The VO_2-VCO_2 data was fitted with a single regression line, a bilinear regression line, an exponential curve, and a 3rd order polynomial curve for an athlete without breakpoint (the four plots to the left) and with suggested breakpoint presence (the four plots to the right). (Note that the plots of the five athletes with a suggested breakpoint also show a rather linear increase in the VO_2-VCO_2 relationship.)

Oxygen uptake (VO_2), carbon dioxide production (VCO_2)

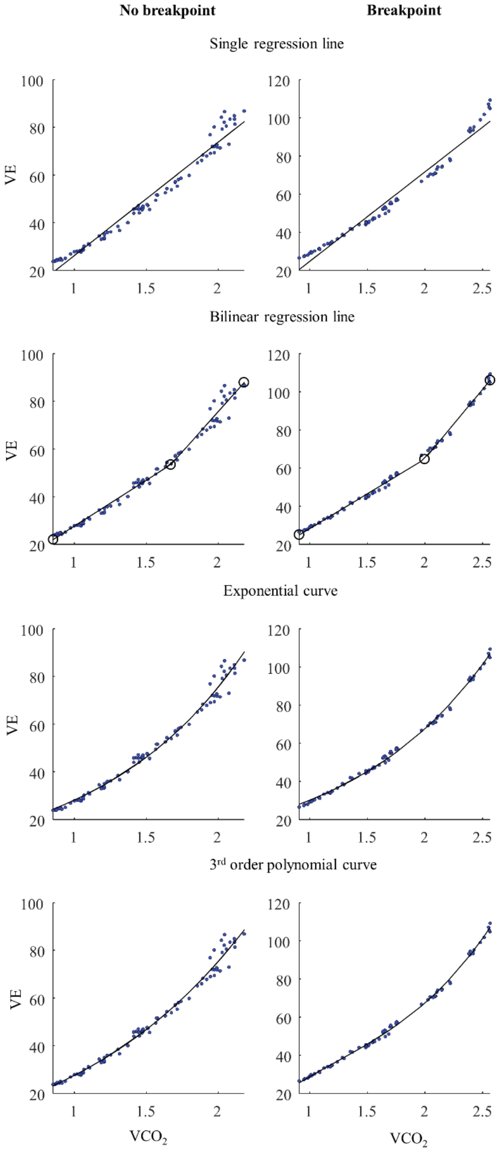


Fig 4. Exemplary VCO_2-VE plots. The VCO_2-VE data was fitted with a single regression line, a bilinear regression line, an exponential curve, and a third order polynomial curve for an athlete without breakpoint (the four plots to the left) and with suggested breakpoint presence (the four plots to the right). (Note that the plots of the five athletes with a suggested breakpoint show a rather curvilinear increase in the VCO_2-VE relationship.)

Carbon dioxide production (VCO_2), minute ventilation (VE)

Table 3. The coefficient of determination (mean r^2 ± SD (range) for the two regression lines (breakpoint model) and the single regression line, exponential and 3rd order polynomial curve (continuous no-breakpoint models) fitted to the gas exchange data of 15 elite ice sledge hockey players following a protocol with stepwise increases in workload every 5 min while upper-body poling.

	Two regression lines	**Single regression line**	**Exponential curve**	**3rd order polynomial curve**
VO_2-VCO_2 plots	0.995 ± 0.005 (0.993-0.998)	0.994 ± 0.005 (0.991-0.996)	0.993 ± 0.005 (0.991-0.996)	0.996 ± 0.005 (0.993-0.998)
VE/VO_2 plots	0.931 ± 0.069 (0.896-0.966)	0.764 ± 0.094 (0.716-0.811)	0.919 ± 0.064 (0.886-0.951)	0.932 ± 0.064 (0.900-0.964)
VE/VCO_2 plots	0.940 ± 0.044 (0.918-0.962)	0.700 ± 0.142 (0.628-0.772)	0.920 ± 0.044 (0.898-0.942)	0.940 ± 0.041 (0.919-0.961)
VCO_2-VE plots	0.995 ± 0.003 (0.994-0.997)	0.968 ± 0.015 (0.960-0.976)	0.992 ± 0.006 (0.989-0.994)	0.996 ± 0.003 (0.994-0.998)

Oxygen uptake (VO_2), carbon dioxide production (VCO_2), minute ventilation (VE)

Table 4. Akaike weights (w_i) representing a measure of strength of evidence for probability of best fit of the two regression lines (breakpoint model) and the single regression line, exponential and 3rd order polynomial curve (continuous no-breakpoint models) (mean w_i ± SD (95 % CI) [# of participants with better fit of the respective model compared to the two regression lines]) fitted to the gas exchange data of 15 elite ice sledge hockey players following a protocol with stepwise increases in workload every 5 min while upper-body poling.

	Two regression lines	**Single regression line**	**Exponential curve**	**3rd order polynomial curve**
VO_2-VCO_2 plots	0.29 ± 0.35 (0.11-0.46)	0.07 ± 0.18 (-0.03-0.16) [#0]	0.03 ± 0.10 (-0.01-0.08) [#0]	0.61 ± 0.37 (0.42-0.79) [#10]
VE/VO_2 plots	0.41 ± 0.45 (0.18-0.64)	0.00 ± 0.00 (0.00-0.00) [#0]	0.13 ± 0.24 (0.01-0.26) [#2]	0.46 ± 0.40 (0.25-0.66) [#7]
VE/VCO_2 plots	0.47 ± 0.49 (0.22-0.72)	0.00 ± 0.00 (0.00-0.00) [#0]	0.09 ± 0.19 (-0.01-0.19) [#2]	0.44 ± 0.43 (0.22-0.66) [#6]
VCO_2-VE plots	0.31 ± 0.44 (0.09-0.54)	0.00 ± 0.00 (0.00-0.00) [#0]	0.00 ± 0.01 (0.00-0.01) [#0]	0.68 ± 0.44 (0.46-0.91) [#10]

Oxygen uptake (VO_2), carbon dioxide production (VCO_2), minute ventilation (VE)

In the gas exchange data displayed in the VE/VO_2 plots and the VE/VCO_2 plots, the breakpoint model fitted better than the continuous no-breakpoint models in six and seven of the athletes, respectively (Fig 5). Accordingly, it is unclear if in general the breakpoint or continuous no-breakpoint models fit the VE/VO_2 and the VE/VCO_2 data best (Table 4). The rise in VE/VO_2 occurred earlier than the VE/VCO_2 only in four athletes (Fig S1). The VT detection by the VE/VO_2 relationship was, therefore, only valid in these four athletes. In none of these four athletes, did the breakpoint model fit the VE/VO_2 data better than the continuous no-breakpoint models.

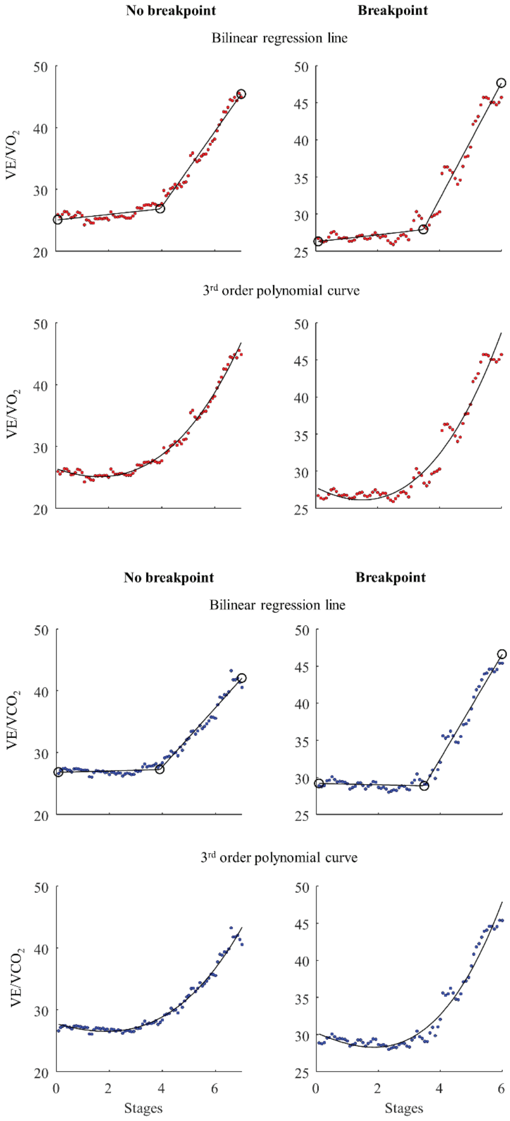


Fig 5. Exemplary VE/VO$_2$ and VE/VCO$_2$ plots. Exemplary VE/VO$_2$ data fitted with a bilinear regression line and a 3rd order polynomial curve for an athlete without breakpoint (upper two plots to the left) and with suggested breakpoint presence (upper two plots to the right). Exemplary VE/VCO$_2$ data fitted with a bilinear regression line and a 3rd order polynomial curve for an athlete without breakpoint (lower two plots to the left) and with suggested breakpoint presence (upper two plots to the right).

Oxygen uptake (VO$_2$), carbon dioxide production (VCO$_2$), minute ventilation (VE)

Discussion

The main aim of this study was to compare physiological and perceptual outcome parameters identified with common gas exchange and BLa thresholds methods used to determine the VT, LT1, RCT and LT2 in Paralympic athletes while upper-body poling. Furthermore, we compared the fit of breakpoint models used to determine gas exchange thresholds to the fit of continuous linear or curvilinear (i.e., no-breakpoint) models. The LT1 occurred at much lower exercise intensity than the VT although both are used as indicators of AT, whereas there were no or minor differences between the methods used to identify the RCT and LT2 that determine the ANT. Furthermore, the RCT and LT2 did not differ from the VT. In addition, the outcome parameters corresponding to the LT1 and LT2 using the log-log transformed VO_2-BLa data and the modified D_{max} method, respectively, were significantly higher than ones identified with fixed BLa values at the LT1 and LT2 (i.e., rise in BLa of +0.4/1.0 at LT1 or BLa concentration of 4 $mmol \cdot L^{-1}$ at LT2). We were able to determine breakpoints at the VT and RCT with different gas exchange methods with good fit in all 15 participants, although continuous no-breakpoint models showed even better fit for most participants.

The physiological and perceptual outcome parameters identified with a fixed rise in BLa at the LT1 were significantly lower than the ones at the VT, and the outcome parameters using these methods only low or moderately correlated with each other. Overall, this indicates that these two thresholds cannot be used interchangeably to determine the AT. In addition, thresholds identified by a fixed BLa increase at the LT1 were significantly lower compared with the breakpoints identified in the log-log transformed VO_2-BLa data, showing that individually adjustable BLa methods did not correspond with fixed methods in determining the LT1. The early occurrence of a rise in BLa in upper-body exercise is in accordance with Beneke et al. [41], who found BLa to be higher at a given workload in activities involving smaller muscle mass, where power output per kg of active muscle mass and, thus, local metabolic stress is increased compared to lower body exercise. In addition, BLa accumulation after cessation of exercise was shown to be faster in individuals with a spinal cord injury as compared to able-bodied individuals [42]. However, although outcome parameters identified with breakpoints in the log-log transformed VO_2-BLa data are not significantly lower than the ones identified at the VT, outcome parameters identified with methods using fixed BLa values to identify the LT1 are much lower than the VT.

As estimates of the ANT, the outcome parameters identified with the D_{max} method to determine LT2 did not significantly differ from the ones identified with breakpoints in the VCO_2-VE data at the RCT, whereas the outcome parameters identified with breakpoints in the VE/VCO_2 data were significantly lower than these. However, the outcome parameters identified by the latter method differ only marginally from the two other ANT methods, indicating that the exercise intensity where a disproportionate increase in BLa and in VE occurs is located relatively similar. Note that we decided to not correct for multiple comparisons and rather present the uncorrected p-values from paired samples t-tests instead. Although we are aware of the subsequent increased chances of making a type 1 errors, the decreased chances of making a type a type 2 errors were regarded more important, which is in accordance with Rothman (1990)[43]. However, if Bonferonni corrections would have been used in this specific case, there would have been no significant differences between the outcome measures identified at with these three methods.

Furthermore, most of the outcome parameters identified with the different methods at the LT2 and RCT are low to moderately correlated, coinciding with high individual variation in the outcome parameters within each of the methods used to identify the LT2 and RCT. This indicates that an individual with a high LT2 does not necessarily display a high RCT. The high individual variation may be explained by disability-related differences in the cardio-respiratory system that might affect physiological responses to upper-body exercise. For example, athletes with a spinal cord injury exercising in an upper-body mode were shown to vary considerably in their VO_{2peak} depending on their level of injury [44], which might also reflect differences in the % of VO_{2peak} that can be sustained during exercise. In addition, the inclusion of one participant that was much older than the rest and one female participant may have contributed to the high variation. Furthermore, individual variation in physiological responses may be higher in upper-body exercise compared to lower-body exercise. Altogether, it is questionable whether the similar outcome parameters identified at the LT2 and the RCT on a group basis, result in similar outcome parameters at the LT2 and RCT for the individual sitting athlete when training in an upper-body mode.

The thresholds identified by the breakpoints in the VE/VO_2 at the VT and in VE/VCO_2 at the RCT did not significantly differ and were highly correlated. This, together with the rather linear increase in the VO_2-VCO_2 relationship suggests that it is solely the disproportionate rise in VE that leads to a rather rapid increase in the data of the VE/VO_2, and the VE/VCO_2 plots, and to discernible breakpoints in approximately half of the participants. Together with the close location of the breakpoints identified in the VCO_2-VO_2 data at the VT and the VCO_2-VE data at the RCT, this indicates that a two-phase (low-high) rather than a three-phase (low-moderate-high) intensity zone model could be applicable in athletes with a disability who exercise in an upper-body mode. This is in contrast to significant differences between the VT and the RCT in Dekerle et al. [45], who test able-bodied participants in the arm crank ergometry mode, and Leicht et al. [18], who tested wheelchair athletes in the wheelchair treadmill mode. However, our findings are in line with a study of Pires et al. [46] who also found one rather than two thresholds in the gas exchange data in upper-body trained able-bodied participants during exercise in the arm crank ergometry mode. Whether the discrepancies between studies are related to employment of e.g. different populations, protocols or exercise modes needs to be examined further in other experimental designs.

All gas exchange threshold methods have in common that there is an a priori assumption of the presence of a breakpoint, defined as “a place where an interruption or change occurs” [47]. However, the presence or absence of breakpoints in the gas exchange data is a debated topic [8, 12]. Thus, in addition to the breakpoint models used to identify the VT and the RCT in the present study, we fitted continuous no-breakpoint models to the data to investigate if there are clear breakpoints in our data. Here, we found good fit for the breakpoint model used to identify the gas exchange thresholds, but better fit for the curvilinear no-breakpoint models in most cases. We, hence, question if clear breakpoints really exist in the gas exchange data of athletes with disabilities in an upper-body exercise mode. However, since there were only minor differences in the fit of the different models, even in the participants with suggested breakpoint presence, the practical consequences of these differences can be questioned.

Conclusion

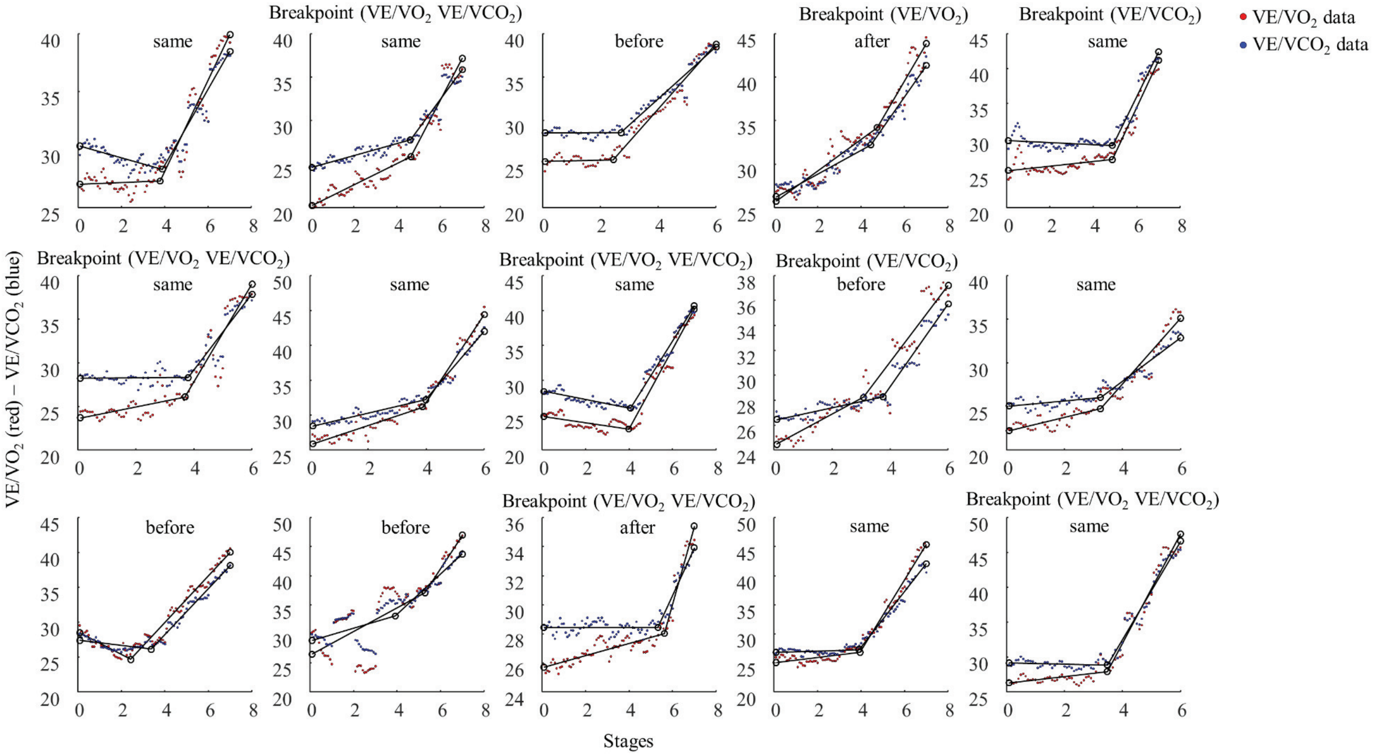
In Paralympic athletes who exercise in upper-body poling, the physiological and perceptual outcome values identified at the VT and the LT1 showed large differences, which demonstrates that these cannot be used interchangeably to identify the AT. In addition, the close location of the VT, RCT and LT2 does not allow us to distinguish the AT and ANT, indicating that there might only be one threshold in athletes with a disability exercising in an upper-body mode. Furthermore, continuous no-breakpoint models fit the gas exchange data better than breakpoint models in most participants. We, hence, question if clear breakpoints in the gas exchange data really exist in an upper-body exercise mode in athletes with disabilities.

Acknowledgements

The laboratory equipment was provided by NeXt Move, Norwegian University of Science and Technology (NTNU). *NeXt Move* is funded by the Faculty of Medicine at NTNU and Central Norway Regional Health Authority. The funder~~s~~ had no role in study design, how the data collection and analysis was performed, decision to publish, or preparation of the manuscript. The authors acknowledge the support of the Olympic and Paralympic Centre in Oslo and the Centre for Elite Sports Research in Trondheim in conducting this research. The eager participation of the athletes is deeply appreciated. None of the authors have any conflicts of interest to declare.

Supporting information

S1 Excel file. Data. Data and analyses conducted in this study of gas exchange and blood lactate threshold in Paralympic sitting athletes.



S1 Fig. VE/VO$_2$ and VE/VCO$_2$ plots fitted with two regression lines. The data is of the six or seven completed stages of each of the 15 athletes. Breakpoint presence is indicated above each individual plot. Furthermore, it is indicated in the second row above the figures whether the two thresholds occur at the same time, or the VE/VO$_2$ occurs before or after the VE/VCO$_2$ threshold.

Oxygen uptake (VO$_2$), carbon dioxide production (VCO$_2$), minute ventilation (VE)

References

1. Seiler S, Toennessen E. Intervals, Thresholds, and Long Slow Distance: the Role of Intensity and Duration in Endurance Training Sportscience. 2009;13:32-53.
2. Binder RK, Wonisch M, Corra U, Cohen-Solal A, Vanhees L, Saner H, et al. Methodological approach to the first and second lactate threshold in incremental cardiopulmonary exercise testing. Eur J Cardiovasc Prev Rehabil. 2008;15(6):726-34.
3. Beaver WL, Wasserman K, Whipp BJ. A new method for detecting anaerobic threshold by gas exchange. J Appl Physiol (1985). 1986;60(6):2020-7.
4. Beaver WL, Wasserman K, Whipp BJ. Improved detection of lactate threshold during exercise using a log-log transformation. J Appl Physiol (1985). 1985;59(6):1936-40.
5. Bishop D, Jenkins DG, Mackinnon LT. The relationship between plasma lactate parameters, Wpeak and 1-h cycling performance in women. Med Sci Sports Exerc. 1998;30(8):1270-5.
6. Wasserman K. The anaerobic threshold: definition, physiological significance and identification. Adv Cardiol. 1986;35:1-23.
7. Reinhard U, Muller PH, Schmulling RM. Determination of anaerobic threshold by the ventilation equivalent in normal individuals. Respiration. 1979;38(1):36-42.
8. Myers J, Ashley E. Dangerous curves. A perspective on exercise, lactate, and the anaerobic threshold. Chest. 1997;111(3):787-95.
9. Wasserman K. The anaerobic threshold measurement to evaluate exercise performance. Am Rev Respir Dis. 1984;129(2 Pt 2):S35-40.
10. Wasserman K. Determinants and detection of anaerobic threshold and consequences of exercise above it. Circulation. 1987;76(6 Pt 2):Vi29-39.
11. Faude O, Kindermann W, Meyer T. Lactate threshold concepts: how valid are they? Sports Med. 2009;39(6):469-90.
12. Hopker JG, Jobson SA, Pandit J. Controversies in the physiological basis of the 'anaerobic threshold'and their implications for clinical cardiopulmonary exercise testing. Anaesthesia. 2011;66(2):111-23.
13. Peronnet F, Aguilaniu B. Lactic acid buffering, nonmetabolic CO2 and exercise hyperventilation: a critical reappraisal. Respir Physiol Neurobiol. 2006;150(1):4-18.
14. Brooks GA. Anaerobic threshold: review of the concept and directions for future research. Med Sci Sports Exerc. 1985;17(1):22-34.
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 Sports Exerc. 2001;33(11):1841-8.
16. Maffulli N, Testa V, Capasso G. Anaerobic threshold determination in master endurance runners. J Sports Med Phys Fitness. 1994;34(3):242-9.
17. Chicharro JL, Perez M, Vaquero AF, Lucia A, Legido JC. Lactic threshold vs ventilatory threshold during a ramp test on a cycle ergometer. J Sports Med Phys Fitness. 1997;37(2):117-21.
18. 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):1635-43.
19. Ribeiro J, Figueiredo P, Sousa M, De Jesus K, Keskinen K, Vilas-Boas JP, et al. Metabolic and ventilatory thresholds assessment in front crawl swimming. J Sports Med Phys Fitness. 2015;55(7-8):701-7.
20. Bernardi M, Guerra E, Di Giacinto B, Di Cesare A, Castellano V, Bhambhani Y. Field evaluation of paralympic athletes in selected sports: implications for training. Med Sci Sports Exerc. 2010;42(6):1200-8.

21. Bhambhani YN, Holland LJ, Steadward RD. Anaerobic threshold in wheelchair athletes with cerebral palsy: validity and reliability. Arch Phys Med Rehabil. 1993;74(3):305-11.

22. Coutts KD, McKenzie DC. Ventilatory thresholds during wheelchair exercise in individuals with spinal cord injuries. Paraplegia. 1995;33(7):419-22.

23. Davis JA, Vodak P, Wilmore JH, Vodak J, Kurtz P. Anaerobic threshold and maximal aerobic power for three modes of exercise. J Appl Physiol. 1976;41(4):544-50.

24. Keyser RE, Mor D, Andres FF. Cardiovascular responses and anaerobic threshold for bicycle and arm ergometer exercise. Arch Phys Med Rehabil. 1989;70(9):687-91.

25. Lin KH, Lai JS, Kao MJ, Lien IN. Anaerobic threshold and maximal oxygen consumption during arm cranking exercise in paraplegia. Arch Phys Med Rehabil. 1993;74(5):515-20.

26. Orr JL, Williamson P, Anderson W, Ross R, McCafferty S, Fettes P. Cardiopulmonary exercise testing: arm crank vs cycle ergometry. Anaesthesia. 2013;68(5):497-501.

27. Schneider DA, Sedlock DA, Gass E, Gass G. VO2peak 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):292-7.

28. Vinet A, Le Gallais D, Bernard PL, Poulain M, Varray A, Mercier J, et al. Aerobic metabolism and cardioventilatory responses in paraplegic athletes during an incremental wheelchair exercise. Eur J Appl Physiol Occup Physiol. 1997;76(5):455-61.

29. Yasuda N, Gaskill SE, Ruby BC. No gender-specific differences in mechanical efficiency during arm or leg exercise relative to ventilatory threshold. Scand J Med Sci Sports. 2008;18(2):205-12.

30. Hegge AM, Bucher E, Ettema G, Faude O, Holmberg HC, Sandbakk O. Gender differences in power production, energetic capacity and efficiency of elite cross-country skiers during whole-body, upper-body, and arm poling. Eur J Appl Physiol. 2015.

31. Borg GA. Psychophysical bases of perceived exertion. Med Sci Sports Exerc. 1982;14(5):377-81.

32. Leicht CA, Tolfrey K, Lenton JP, Bishop NC, Goosey-Tolfrey VL. The verification phase and reliability of physiological parameters in peak testing of elite wheelchair athletes. Eur J Appl Physiol. 2013;113(2):337-45.

33. Inbar O, Faina M, Demarie S, Whipp BJ. VO 2 Kinetics during Moderate Effort in Muscles of Different Masses and Training Level. ISRN Physiology. 2012;2013.

34. Powers SK, Dodd S, Garner R. Precision of ventilatory and gas exchange alterations as a predictor of the anaerobic threshold. Eur J Appl Physiol Occup Physiol. 1984;52(2):173-7.

35. Buckley JD, Bourdon PC, Woolford SM. Effect of measuring blood lactate concentrations using different automated lactate analysers on blood lactate transition thresholds. J Sci Med Sport. 2003;6(4):408-21.

36. Stegmann H, Kindermann W. Comparison of prolonged exercise tests at the individual anaerobic threshold and the fixed anaerobic threshold of 4 mmol.l(-1) lactate. Int J Sports Med. 1982;3(2):105-10.

37. Plichta SB, Kelvin EA, Munro BH. Munro's statistical methods for health care research: Wolters Kluwer Health/Lippincott Williams & Wilkins; 2013.

38. Bozdogan H. Model selection and Akaike's information criterion (AIC): The general theory and its analytical extensions. Psychometrika. 1987;52(3):345-70.

39. Burnham KP, Anderson DR. Model selection and multimodel inference: a practical information-theoretic approach: Springer Science & Business Media; 2003.

40. Wagenmakers E-J, Farrell S. AIC model selection using Akaike weights. Psychonomic bulletin & review. 2004;11(1):192-6.

41. Beneke R, Leithäuser R, Hütler M. Dependence of the maximal lactate steady state on the motor pattern of exercise. Br J Sports Med. 2001;35(3):192-6.

42. Leicht C, Perret C. Comparison of blood lactate elimination in individuals with paraplegia and able-bodied individuals during active recovery from exhaustive exercise. J Spinal Cord Med. 2008;31(1):60-4.

43. Rothman KJ. No adjustments are needed for multiple comparisons. Epidemiology. 1990;1(1):43-6.

44. Bhambhani Y. Physiology of wheelchair racing in athletes with spinal cord injury. Sports Med. 2002;32(1):23-51.

45. Dekerle J, Dupont L, Caby I, Marais G, Vanvelcenaher J, Lavoie JM, et al. Ventilatory thresholds in arm and leg exercises with spontaneously chosen crank and pedal rates. Percept Mot Skills. 2002;95(3 Pt 2):1035-46.

46. Pires FO, Hammond J, Lima-Silva AE, Bertuzzi RC, Kiss MA. Ventilation behavior during upper-body incremental exercise. J Strength Cond Res. 2011;25(1):225-30.

47. English Oxford Dictionaries 2017 [Available from: https://en.oxforddictionaries.com/definition/break_point.