Trinity College Dublin
Coláiste na Tríonóide, Baile Átha Cliath
The University of Dublin

A study of the repeatability of the Bruce Treadmill Protocol in measuring $\dot{V}O_{2max}$ and in estimating $\dot{V}O_{2max}$ through prediction equations

Kate Marie Macnamara

BSc Physiotherapy

Supervised by Prof. John Gormley

Submitted for the Degree of Research Masters in Science

Discipline of Physiotherapy

School of Medicine

University of Dublin, Trinity College

February 2021

Declaration of Online Access and the General Data Protection Regulation

This thesis is submitted by the undersigned to Trinity College Dublin for the examination of the degree of Research Masters in Science. I declare that this thesis has not been submitted as an exercise for a degree at this or any other university and it is entirely my own work. I agree to deposit this thesis in the University's open access institutional repository or allow the Library to do so on my behalf, subject to Irish Copyright Legislation and Trinity College Library conditions of use and acknowledgement. I consent to the examiner retaining a copy of the thesis beyond the examining period, should they so wish (EU GDPR May 2018).

I declare that this thesis has not been submitted as an exercise for a degree at this or any other university and it is entirely my own work. I agree to deposit this thesis in the University's open access institutional repository or allow the library to do so on my behalf, subject to Irish Copyright Legislation and Trinity College Library conditions of use and acknowledgement.

Signed Kate Macnamara Date **31-01-2021**

Kate Macnamara

Declaration and Statement of Plagiarism

I declare that the work submitted in this thesis is entirely my own. I have read and I understand the plagiarism provisions in the General Regulations of the University Calendar for the current year, found at http://www.tcd.ie/calendar. I have also completed the Online Tutorial on avoiding plagiarism 'Ready Steady Write', located at http://tcd-ie.libguides.com/plagiarism/ready-steady-write. I confirm that full and informed consent was obtained from each participant for the research in this thesis.

Signed Kate Macnamara Date **31-01-2021**

Kate Macnamara

Acknowledgements

Firstly, I would like to thank my supervisor, Professor John Gormley. Thank you for challenging my theories and ideas, for the debates and discussions around many topics, and for teaching and guiding me through the entire research process. I have learned so much in the past two and a half years, with regards to conducting research, but also in teaching and planning, in practical skills, and in how to question the norm and think outside the box.

Thank you to all of the participants across both studies – without you there would be no research. I am very grateful to all of the post-graduate physiotherapy researchers at Trinity College; thank you for the support, the good lunches and the laughs along the way. Thank you to the fourth year undergraduate physiotherapy students of 2019 with whom I conducted the research, and a special thank you to Caitríona Quinn for taking so much of your time to be my fellow researcher, keeping things so on-track and organised with each exercise test conducted. I would also like to thank Ann Monaghan for teaching me all the ways of the research lab, for reading many drafts of literature reviews before I knew what I was doing and helping me in my panics when equipment wouldn't behave. I am very grateful for the support of all my friends in this time, and especially for Sarah Purcell – your proof-reading skills and advice are excellent!

Finally, to my family, Ríonagh, Trevor and Jill, and to Seán Daly – thank you for the tea, the phone-calls, the distractions when they were needed and the jollying along when that was needed too! I am so grateful for all your support to me – during this masters, but also through my entire life in all I undertake. You have kept me going, especially when the going got tough during the pandemic and the many lockdowns, and I couldn't have done this without you all at my side.

Table of Contents

List of Figures and Tables

Figures

Tables

List of Abbreviations

95% CI – 95% Confidence Interval

AAA – abdominal aortic aneurysm

ACSM – American College of Sports Medicine

ANOVA – analysis of variance

AXIS – Appraisal tool for Cross-Sectional Studies

BMI – body mass index

CC – correlation coefficient (r)

CPET – Cardiopulmonary Exercise Testing

CV – coefficient of variation

CVD – cardiovascular disease

ECG – electrocardiogram

HR – heart rate

HR_{max} – maximal heart rate

Hrs – hours

IPAQ-SF – International Physical Activity Questionnaire – Short Form

LOA – limits of agreement

MCID – minimal clinically important difference

MET – metabolic equivalent of task

MRI – magnetic resonance imaging

MVPA – moderate to vigorous physical activity

NYHA – Hew York Heart Association

PA – physical activity

PAR-Q – Physical Activity Readiness Questionnaire

PRISMA Statement – The Preferred Reporting Items for Systematic Reviews and Meta-Analysis Statement

RPE – rating of perceived exertion

SD – standard deviation

$\dot{V}O_{2max}$ – maximal oxygen consumption

$\dot{V}O_{2peak}$ – highest achieved oxygen consumption in a given exercise test

Yrs – years

Summary

Introduction: Maximal oxygen consumption ($\dot{V}O_{2max}$) is a measure of cardiopulmonary fitness, and a strong indicator of risk of cardiovascular disease (CVD); with a lower $\dot{V}O_{2max}$ increasing a person's risk of CVD. The most reliable method of $\dot{V}O_{2max}$ measurement is via maximal exercise testing with direct analysis of expired air, commonly using the Bruce treadmill protocol. If, due to physical limitations of the participant, or a lack of testing equipment, $\dot{V}O_{2max}$ can be predicted through equations.

The first aim of this thesis was to determine the repeatability of the Bruce protocol in measuring $\dot{V}O_{2max}$. This was achieved through a literature review of previous Bruce protocol repeatability research, and by conducting maximal and submaximal repeatability studies using the Bruce protocol with healthy male adults. The second aim was to determine the reliability of prediction equations used with the Bruce protocol to calculate $\dot{V}O_{2max}$, again through a literature review of previous research in this area, and then comparing predicted and measured $\dot{V}O_{2max}$ results from the maximal testing study to determine their accuracy.

Study 1: The first literature review examined the previous research determining the repeatability of the Bruce protocol in measuring $\dot{V}O_{2max}$. Eleven studies were deemed eligible for inclusion. All studies concluded that the Bruce protocol was repeatable in measuring $\dot{V}O_{2max}$, although most findings were based on correlation coefficient analysis, which does not accurately represent agreement between repeated measures. A learning effect was also noted in some of the studies.

Study 2: The second literature review analysed previous literature examining different equations used to predict $\dot{V}O_{2max}$ following a Bruce protocol treadmill test. Seven studies were deemed eligible for inclusion. It was found that a number of variables were used to predict $\dot{V}O_{2max}$, which included exercise variables such as maximal test duration, heart rate, and submaximal $\dot{V}O_2$ extrapolated to maximal heart rate, as well as non-exercise data such as a person's age, body fat or body mass index. The main finding across the research was that the American College of Sports Medicine (ACSM) equations consistently over-predicted $\dot{V}O_{2max}$ and therefore should not be used.

Study 3: An initial research study was developed to examine the repeatability of the submaximal Bruce protocol in predicting $\dot{V}O_{2max}$ in health male adults. The submaximal test was chosen to establish a strict procedure for repeatability testing of the Bruce protocol, in a safe manner. Eighteen participants completed the study, and completed three submaximal Bruce protocol

treadmill tests, each a week apart. $\dot{V}O_{2max}$ was predicted using the Fitmate equation, chosen from the results of the previous literature review. Using limits of agreement (LOA) analysis, a wide range in differences between repeated $\dot{V}O_{2max}$ predictions was found (LOA between Test 1 and Test 2: 4.86 to -10.35 $ml \cdot kg^{-1} \cdot min^{-1}$; LOA between Test 2 and Test 3: 6.67 to -7.96 $ml \cdot kg^{-1} \cdot min^{-1}$). A learning effect was noted when the mean differences between repeated $\dot{V}O_{2max}$ predictions were examined, as the mean difference between Test 2 and Test 3 was much lower than that between Test 1 and Test 2. Due to the large LOA across all repeated tests, this study concluded that the submaximal Bruce protocol was not repeatable in measuring $\dot{V}O_{2max}$.

Study 4: The final study of this thesis aimed to examine the repeatability of the Bruce maximal treadmill protocol in measuring $\dot{V}O_{2max}$, as well as to determine the accuracy of several $\dot{V}O_{2max}$ prediction equations, identified through the second literature review. Fifteen healthy male adults completed three repetitions of the maximal Bruce protocol, with a further four participants completing two repetitions. $\dot{V}O_{2max}$ was measured using a metabolic cart. Six prediction equations were used to predict $\dot{V}O_{2max}$, and these results were compared to the measured $\dot{V}O_{2max}$ values. It was found that LOA gave a large range for $\dot{V}O_{2max}$ differences between Tests 1 and 2, but a smaller range between the second and third tests. The maximal Bruce protocol was deemed not repeatable from the first to second test, but due to learning effect noted, is likely repeatable at a third test. The most accurate prediction equation was that by Bruce et al.[1] for active males ("Bruce 2" equation), while the ACSM equations were found to be unreliable in predicting $\dot{V}O_{2max}$.

Conclusion: The current research has shown that the maximal Bruce protocol is repeatable in measuring $\dot{V}O_{2max}$ but only after a full familiarisation session with the test. The submaximal protocol, with $\dot{V}O_{2max}$ predicted through an equation is not repeatable, based on the current findings. Regarding $\dot{V}O_{2max}$ prediction equations, the ACSM equations should be avoided, and the "Bruce 2" equation appears to provide the most accurate result for healthy male participants.

Chapter 1: Introduction

Cardiorespiratory fitness, or a person's maximal aerobic capacity, is measured as maximal oxygen consumption ($\dot{V}O_{2max}$)[2]. A person's aerobic capacity is a strong contributor to predicting a person's risk of developing cardiovascular disease (CVD)[2,3], with a high $\dot{V}O_{2max}$ indicating lower risk. $\dot{V}O_{2max}$ is most commonly measured with a cardiopulmonary exercise test[2,4,5]. The $\dot{V}O_{2max}$ results are used to aid prescription of general and clinical exercise programmes[6] and measure cardiorespiratory fitness improvement post exercise intervention[7,8]. The test can also predict cardiovascular disease risk[2] and determine normative cardiorespiratory fitness in healthy and clinical populations[9-11].

$\dot{V}O_{2max}$ is usually measured using a treadmill or cycle ergometer. Research has shown that the cycle ergometer yields lower peak $\dot{V}O_2$ values compared to treadmill testing[12,13], likely due to a larger muscle mass being utilised during running compared to cycling, which increases oxygen uptake[14]. While there are many cardiopulmonary exercise test protocols described in the literature, the Bruce Protocol for the treadmill is the most commonly used protocol to determine $\dot{V}O_{2max}$[15,16]. It was first described by Bruce et al.[17] as a tool to aid diagnosis of coronary insufficiency and measure aerobic capacity in people with cardiac disease, mainly to test those with mild to moderate cardiac disease. At that time, single stage exercise tests could measure aerobic fitness in those with severe cardiac disease, as one stage already brought these patients to their maximal exertion; but single stage tests were not stressing enough for other patients only mildly or moderately impaired by cardiac diseases. Therefore, a step-wise protocol was devised to gradually increase workload for these individuals, and determine their maximal effort safely. This original step-wise protocol was detailed with large speed and treadmill incline increments[17], but it was refined by Bruce et al.[1] to the protocol commonly used today, which starts at a grade of 10% and a speed of 2.7 $km \cdot h^{-1}$, and increases by 2% incline after every three minutes, along with speed increments to 4.0 $km \cdot h^{-1}$, 5.5 $km \cdot h^{-1}$, 6.8 $km \cdot h^{-1}$, 8.0 $km \cdot h^{-1}$, 8.9 $km \cdot h^{-1}$, 9.7 $km \cdot h^{-1}$ and 10.5 $km \cdot h^{-1}$ at each stage.

As well as its original purpose, the Bruce Protocol has evolved to be used as a tool for assessing aerobic capacity in various populations including healthy athletes[18,19], children[20,21] and those with other health conditions such as diabetes[22] and chronic obstructive pulmonary disease[23]. It has also been used to validate other $\dot{V}O_{2max}$ testing protocols in both paediatric and adult populations[24-28].

Due to its common usage clinically and in research, the validity and reliability of the Bruce protocol is highly important. Repeatability is an element of reliability, where good repeatability means that a test participant will get the same test result on repeated tests, if they are conducted under the same testing conditions by the same tester, over a short period of time[29]. The addition of considering cardiorespiratory fitness as a factor in determining a person's risk of CVD greatly improved the accuracy of risk classification and chance of all-cause mortality[2]. This risk classification contributes to intervention planning, physical activity counselling, and prescription of exercise to improve a person's low $\dot{V}O_{2max}$. Given that a person's $\dot{V}O_{2max}$ is used to determine their risk of developing CVD, it is important that the $\dot{V}O_{2max}$ is measured accurately. Researchers and clinicians must be sure that they can rely on the $\dot{V}O_{2max}$ measurement obtained when using the Bruce treadmill protocol to determine a person's cardiorespiratory fitness. Furthermore, if the Bruce protocol is used to determine $\dot{V}O_{2max}$ before and after an intervention, it is important that the protocol is repeatable, so that any change noted can be attributed to true physiological change in the person, and not merely due to measurement variation in the protocol itself. Likewise, if the Bruce protocol is used to validate new $\dot{V}O_{2max}$ testing protocols, but it itself is not repeatable, then the validity of these other protocols may be called into question.

For certain clinical populations completing a maximal exercise test is not safe or feasible – perhaps due to cardiac conditions, balance difficulties or joint and muscle pain[30]. In these cases, submaximal tests can be conducted, and prediction equations used to calculate the person's $\dot{V}O_{2max}$. Prediction equations may also be used in situations where exercise testing equipment or trained testing staff are not available, or to save time and cost[31-33]. Determining an accurate $\dot{V}O_{2max}$ value is important, but there is a risk with prediction equations that the $\dot{V}O_{2max}$ value will be under- or over-estimated[34-36]. By using a prediction equation to determine a patient's $\dot{V}O_{2max}$, the risk of incorrectly

classifying their CVD risk level, or incorrectly estimating their response to an intervention, is higher. This could subsequently negatively impact future interventions and management for that patient[2]. While the use of prediction equations may be unavoidable (if patients are unable to complete maximal exercise testing and $\dot{V}O_{2max}$ measurement) the risk of incorrect prediction of cardiorespiratory fitness should always be taken into account.

In the literature, a person's highest oxygen consumption can be referred to in two ways: $\dot{V}O_{2max}$ and $\dot{V}O_{2peak}$, leading to debate on how $\dot{V}O_{2max}$ should be defined. Some researchers say that no change (or very small increase) in $\dot{V}O_2$ despite an increase in workload (i.e. a plateau in $\dot{V}O_2$) is required to confirm that a person reached their $\dot{V}O_{2max}$[4,37,38]. The occurrence of a plateau in $\dot{V}O_2$ is relatively low, however, with only approximately 60% of adults reaching a $\dot{V}O_2$ plateau, and even fewer in children, elderly and clinical populations[39], as well as only 25-47% of elite athletes achieving the $\dot{V}O_2$ plateau during testing[40]. Other criteria have been established to identify when a person has reached their $\dot{V}O_{2max}$, in the absence of a $\dot{V}O_2$ plateau[4,41] – but these criteria have in turn been criticised for inconsistencies and difficulties in applying them across different populations[42,43]. Therefore, some authors use the term "$\dot{V}O_{2peak}$" to define the highest $\dot{V}O_2$ achieved by participants during a maximal exercise test, which may or may not be their $\dot{V}O_{2max}$. "$\dot{V}O_{2peak}$" has been described to best suit clinical populations, while "$\dot{V}O_{2max}$" is more appropriate for healthy individuals undergoing exercise testing[42,44]. As this thesis will focus on healthy participants (with the exception in the repeatability literature review where clinical populations were included) the term "$\dot{V}O_{2max}$" will be used throughout in relation to maximal oxygen consumption. Exceptions to this will be for referenced studies that specifically used the term "$\dot{V}O_{2peak}$" to describe their results.

The first aim of this thesis is to determine the repeatability of the Bruce treadmill protocol in measuring the $\dot{V}O_{2max}$ of healthy males. This will be achieved by conducting a literature review of previous studies of Bruce repeatability, and conducting a maximal Bruce treadmill study with the protocol being repeated on a number of occasions. The second aim of the thesis is to examine the currently available prediction equations used alongside the Bruce protocol to predict $\dot{V}O_{2max}$, and to examine their predictive accuracy with the data collected during the maximal Bruce testing study.

Chapter 2: Literature Review: The current evidence for the repeatability of the Bruce Protocol in measuring $\dot{V}O_{2max}$

2.1 Introduction

As the Bruce protocol is the most commonly used treadmill protocol for determining maximal aerobic capacity ($\dot{V}O_{2max}$)[15,45,46], it is important that it is repeatable so that whether used in research or clinically, the results are reliable and accurate. A reduced $\dot{V}O_{2max}$ increases risk of cardiovascular disease[2,3] and so to ensure accurate determination of risk, an accurate measurement protocol must be used. Repeatability demonstrates that a measurement gives an accurate and unchanging reading when conducted with the same participant under the same clinical settings by the same tester[29]. To determine this repeatability, the statistical analysis conducted on the data is important. A number of statistical methods have been used for testing repeatability, such as correlation coefficients[47], coefficients of variation[48], and limits of agreement[49].

The aim of this literature review was to examine the previous studies of repeatability into the Bruce protocol in measuring $\dot{V}O_{2max}$, in both healthy and clinical populations.

2.2 Methods

2.2.1 Search Strategy

The Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) statement was followed throughout the development of this review[50]. The online databases of Embase, Medline, CINAHL and Web of Science were initially searched in November 2018 to find studies relating to the repeatability of the Bruce protocol. The search strategy was designed to combine the search terms “treadmill” or “exercise test” with “Bruce” and “protocol”, along with validity and reliability terms such as “repeatability”, “reproducibility”, “predict”, “specificity”, “test-retest” and “evaluation” (full list of search terms in Appendix 1, pg. 104).

2.2.2 Eligibility criteria

The inclusion criteria for the repeatability review were as follows: studies published in English, using the standard Bruce protocol or modified Bruce protocol, repeated on two or more occasions where results looked at repeatability of the protocol (and not change in $\dot{V}O_{2max}$ over time), in the same sample of participants. Participants could be healthy or from a clinical population. The Standard Bruce protocol was defined as per the study by Bruce et al.[1] with the treadmill beginning at a grade of 10% and a speed of 2.7 $km \cdot h^{-1}$, increasing by 2% incline at every three-minute stage, along with speed increments to 4.0 $km \cdot h^{-1}$, 5.5 $km \cdot h^{-1}$, 6.8 $km \cdot h^{-1}$, 8.0 $km \cdot h^{-1}$, 8.9 $km \cdot h^{-1}$, 9.7 $km \cdot h^{-1}$ and 10.5 $km \cdot h^{-1}$ at each stage. The Modified Bruce protocol included some form of warm-up stage, generally beginning with two 3-minute stages of 2.7 $km \cdot h^{-1}$ at 0% incline and 2.7 $km \cdot h^{-1}$ at 5% incline before continuing as the full standard Bruce protocol[51]. Studies that examined the effect of a drug, dietary or exercise intervention, or that used the Bruce protocol as a stress test in conjunction with imaging or electrocardiogram (ECG) monitoring to diagnose cardiac conditions were excluded. Studies that were not conducted with human participants, that had only paediatric participants, that used the Bruce protocol only once to compare to or validate another protocol, or that used a prediction equation to estimate $\dot{V}O_{2max}$ from submaximal Bruce protocol testing were also excluded.

2.2.3 Study selection and data extraction

The search was initially conducted in November 2018, and the search was repeated again in May 2020 to identify any new publications. The titles and abstracts of all search results were screened by two independent researchers against the inclusion and exclusion criteria. Full texts of selected articles were then read and included in the review if appropriate. Reference lists from included articles were reviewed to identify any further appropriate studies. The full texts of all included studies were analysed, with data extracted into a standardised template.

2.2.4 Quality appraisal

Reporting quality and risk of bias was assessed in each included study, following the Appraisal tool for Cross-Sectional Studies (AXIS)[52] to assess the quality of reporting in

observational longitudinal research (Appendix 2, pg. 105). This tool comprises twenty questions relating to the information reported in the study, covering study aims, participant recruitment, description of methods and statistical analysis used, consistency of results reported, discussion and limitations of the study, and ethical considerations. A score out of twenty was given to each study, with a higher the score indicating better quality of reporting.

2.3 Results

2.3.1 Study selection

The search strategy yielded a total of 1,502 studies, and following the screening process, eleven studies were identified for inclusion in this review[1,17,53-61]. The PRISMA flow diagram[50] detailing the full screening and selection process can be seen in Figure 2.1.

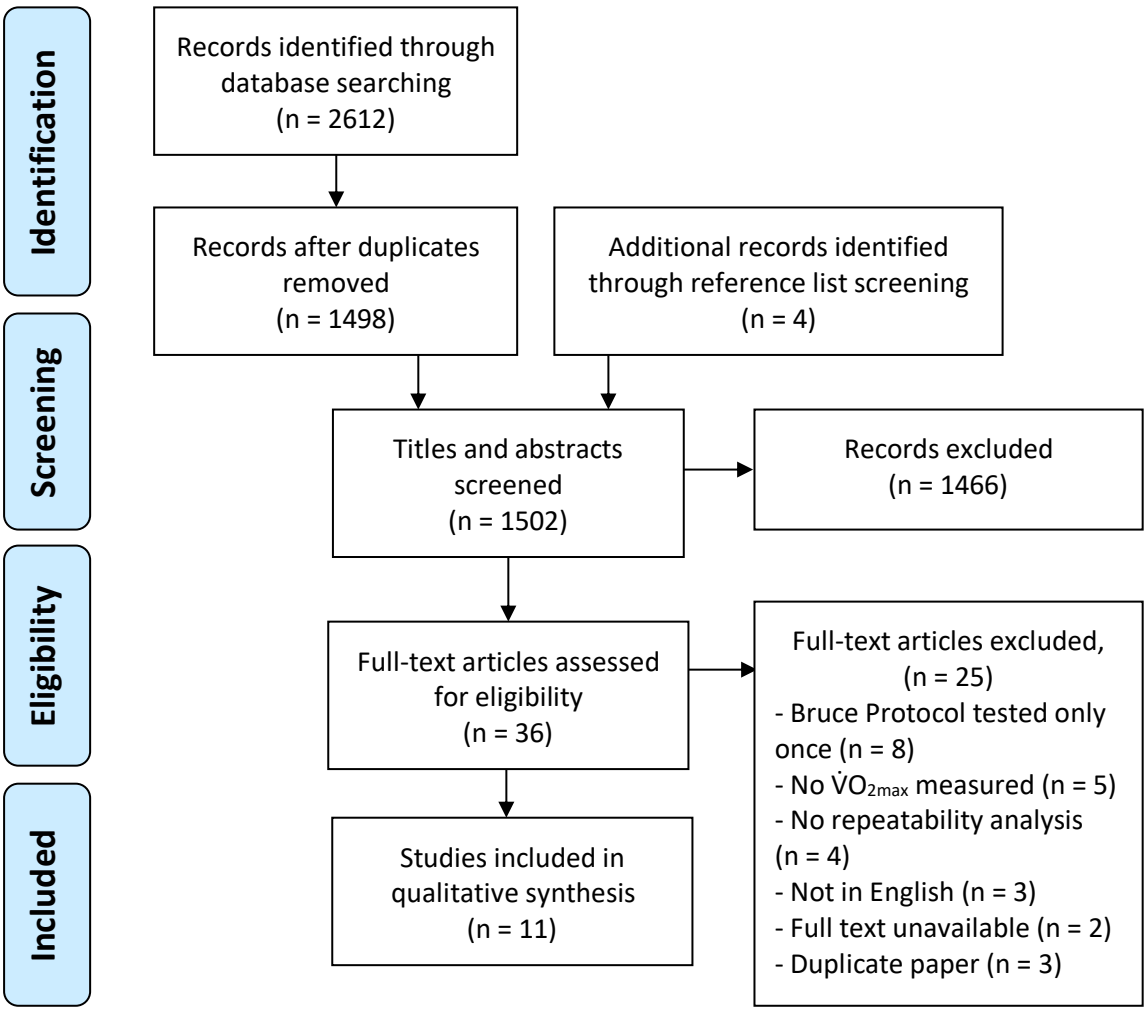


Figure 2.1 Flowchart for literature selection process

2.3.2 Quality Assessment

The mean AXIS score for the studies was 10.9/20 (standard deviation ±2.6; range = 8-17). The studies scored highly in areas relating to suitability of study design and measurement methods used, as well in providing appropriate and complete results in relation to the described methods. The studies generally scored low in areas relating to participants: identifying and describing non-responders, justifying the sample size, and having an appropriate selection process. Scores were also low in relation to determining statistical significance and precision estimates. Only two studies[59,61] scored higher than 65% in the assessment, indicating that the risk of bias was high across the included studies. The full results can be seen in Table 2.1.

2.3.3 Participants

From the eleven studies examining the repeatability of the Bruce protocol, seven tested only healthy populations[17,53-57,60], three examined people with cardiac conditions[58,59,61] and one examined both healthy people and people with cardiac conditions[1]. Two studies[17,58] did not give gender details, but examining all available gender data, it was found that 58.4% of subjects in the studies were male. The number of participants per study ranged from six[60] to seventy-nine[1]. The full breakdown of study details can be seen in Tables 2.2 and 2.3. “Nil specified” has been written in these tables for sections where no information regarding this category was provided in the study articles.

2.3.4 Pre-test instruction and familiarisation

There was a wide variation in the instructions given prior to the exercise test. Four studies provided no information regarding instruction[1,17,53,61]. The remaining studies gave direction regarding the duration of fasting required before testing, which ranged from two hours[56,58,59] to a full night[57], as well as instruction about levels of physical activity allowed before testing. Two studies encouraged participants to maintain a stable weight throughout their testing period[54,55]. There was also variation between studies regarding the level of familiarisation with equipment allowed prior to testing.

Table 2.1 AXIS risk of bias assessment[52]

	Bruce et al.[17]	Profant et al.[53]	Bruce et al.[1]	Froelicher Jr et al.[54]	Ho[55]	Nordrehaug et al.[56]	Fielding et al.[57]	Cooke et al.[58]	Jakovljevic et al.[59]	Hall-Lopez et al.[60]	Harwood et al.[61]
1.	Yes	No	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes
2.	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes
3.	No	No	No	No	No	No	No	No	No	No	Yes
4.	Yes	Yes	Yes	No	Yes	No	Yes	Yes	Yes	No	Yes
5.	Yes	Yes	Unknown	Unknown	Yes	Unknown	Unknown	Yes	Unknown	Unknown	Yes
6.	Unknown	Unknown	Unknown	Unknown	Unknown	Unknown	Unknown	Unknown	Unknown	Unknown	Yes
7.	No	No	No	No	No	No	No	No	No	No	No
8.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
9.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
10.	No	No	No	No	No	No	Yes	No	Yes	No	Yes
11.	No	No	No	No	No	Yes	Yes	No	Yes	No	No
12.	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes
13.*	Unknown	Unknown	Unknown	Unknown	Unknown	Unknown	Unknown	Yes	Unknown	Yes	No
14.	No	No	No	No	No	No	No	No	No	No	No
15.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes
16.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
17.	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes
18.	No	No	No	No	No	No	No	Yes	Yes	Yes	Yes
19.*	No	No	No	Unknown	No	Unknown	Unknown	No	No	No	No
20.	Unknown	Yes	Yes	Unknown	Unknown	Unknown	Yes	Yes	Yes	Yes	Yes
Total Positives	9/20 45%	11/20 55%	10/20 50%	8/20 40%	11/20 55%	8/20 40%	12/20 60%	12/20 60%	13/20 65%	9/20 45%	17/20 85%

*For Questions 13 and 19, Yes is a negative answer, no is a positive answer.

Table 2.2 Comparison of results across eleven repeatability studies (a)

	Participant numbers/ descriptions	Type of protocol	Instructions prior to testing	Familiarisation or warm up (extra to protocol)	Handrail use
Bruce et al.[17]	17 healthy participants (no further specifics given)	3-minute stages: Increments in Stages 1-4: speed 2.7 km·h^{-1} → 5.5 km·h^{-1} → 8.0 km·h^{-1} → 9.7 km·h^{-1}. Grade:10% → 14% → 18% → 22%	*Nil specified*	*Nil specified*	*Nil specified*
Profant et al.[53]	37 healthy middle-aged women (age range 29-70 yrs)	Standard Bruce	*Nil specified*	*Nil specified*	*Nil specified*
Bruce et al.[1]	35 healthy males, 32 healthy females, 12 male patients with cardiac disease	Standard Bruce	*Nil specified*	*Nil specified*	1-2 fingers support as required to maintain position
Froelicher jr. et al.[54]	15 healthy males. Mean age 31.5 yrs (± 11.4 yrs)	Standard Bruce	Keep weight/ PA stable. Fasting x4hrs before test	*Nil specified*	No support allowed
Ho[55]	24 healthy males. Mean age 43.1 yrs (±11.5 yrs)	Standard Bruce	Keep weight/ PA stable throughout study	Familiarisation with treadmill walking and Douglas bag	No support allowed
Nordrehaug et al.[56]	10 healthy males. Mean age 35 yrs (±11 yrs).	Standard Bruce	Fasting x 2hrs before test	Two full Bruce protocols completed for familiarisation 1 week before testing.	*Nil specified*
Fielding et al.[57]	17 healthy females. Mean age 59 yrs (± 1year). No regular PA.	Modified Bruce (start 3.2 km·h^{-1} at 10% incline instead of 2.7 km·h^{-1})	Fasting overnight. 1 caffeine-free snack 1 hour before testing.	Warm up: 2 mins (3.2 km·h^{-1}, 0% grade) and 2 mins (3.2 km·h^{-1}, 10% grade). Rest x 5mins.	No support allowed
Cooke et al.[58]	12 heart failure patients (NYHA II-III) (no further specifics given)	Both Standard and modified Bruce protocols (not specified which participants did which protocol)	Fast x2hrs. No alcohol/ caffeine before tests.	*Nil specified*	*Nil specified*
Jakovljevic et al.[59]	15 male & 4 female patients with chronic heart failure (NYHA Classes I-II). Mean age 62 yrs (±11 yrs).	Modified Bruce protocol (no further details given)	Fast x2hrs. No alcohol/ caffeine before tests	Familiarisation on treadmill, few days before testing.	*Nil specified*
Hall-Lopez et al.[60]	6 healthy males. Mean age 23.4 yrs (±1.3 yrs).	Standard Bruce protocol with warm up beforehand	No PA x24hrs, fast x4hrs, be well hydrated	5-minute warm up, 5km·h^{-1}, 0% incline	*Nil specified*
Harwood et al.[61]	11 males and 1 female with known AAA. Mean age 75.4 yrs (±5.5 yrs)	Modified Bruce, starting 2.7 km·h^{-1} at 0% incline	*Nil specified*	Familiarisation, not specified if on same or different day to testing	*Nil specified*

Key: yrs = years; PA = physical activity; NYHA = New York Heart Association; AAA = abdominal aortic aneurysm; hrs = hours; s = seconds; km·h^{-1} = kilometres per hour

Table 2.3 Comparison of results across eleven repeatability studies (b)

	Number of repetitions	Duration between repetitions	Statistical analysis used	Statistical results for $\dot{V}O_{2max}$	Statistical results for test duration
Bruce et al.[17]	2	Several days	*Nil specified*	"Virtually identical"; "Change was not significant"	*Nil specified*
Profant et al.[53]	2	Weeks	CC	"Satisfactory" reproducibility. CC: r=0.86	*Nil specified*
Bruce et al.[1]	2	Days to weeks	CC	Healthy population CC: r=0.990. Cardiac patients CC: r=0.945. "Excellent reproducibility" for both.	*Nil specified*
Froelicher Jr et al.[54]	3	Minimum 1 week, maximum 5 weeks	CV; statistical test (not specified)	CV=4.4% (range 1.2-8.5%). No difference between repeated tests (no *p* values given).	Longer duration in 2nd and 3rd tests, (p<0.001)
Ho[55]	3	Minimum 1 week, maximum 9 weeks	CV; 2-way ANOVA	CV 9.7%. Significant difference between 1st and 2nd tests, and between 1st and 3rd tests (p<0.001)	CV = 5.5%. No statistical difference between repeated tests (no *p* value given).
Nordrehaug et al.[56]	2	1 week	CC; CV	CC: r=0.94 (no *p* values given). CV=5%.	*Nil specified*
Fielding et al.[57]	5	At least 1 week	CC; CV; ANOVA	CC test 1 to test 2: r=0.75, p=0.0005 (significant correlation). CV=6.5%. No significant difference over 5 tests (p=0.78).	Test 1 to test 2: r=0.81, p<0.0001 (significant correlation).No significant difference over 5 tests (no *p* value given)
Cooke et al.[58]	2	At least 4 weeks	CV; Student t-test; LOA; Repeatability coefficient	CV = 4.7%. No significant difference between 1st and 2nd tests (p=0.1). LOA: -25ml·min^{-1} to 55.6 ml·min^{-1} (mean difference: 15.8ml·min^{-1})	*Nil specified*
Jakovljevic et al.[59]	2	Within one week	CC; CV; LOA	CC: r=0.97. CV=3.8%. Significant difference between tests: p=0.03. LOA: -5.3 to 2.7 ml·kg^{-1}·min^{-1} (mean difference: -1.3 ml·kg^{-1}·min^{-1})	Strong correlation: CC: r=0.95. Significant difference between tests: p=0.00. CV=7.1%. LOA: 183s to 61s (mean difference = 61s)
Hall-Lopez et al.[60]	2	10 minutes, tested on same day	CC (r); Determination coefficient (r^2)	"High reproducibility". CC: r = 0.907; r^2 = 0.823 (no *p* values given)	*Nil specified*
Harwood et al.[61]	2	Maximum one week	CC; (ICC) CV; Paired t-test	Significant correlation: CC: r=0.927 CV = 7.1% No significant difference between tests (p=0.001).	No significant difference between tests, with significant correlation: CC: r=0.967 (p = 0.000)

Key: CC = correlation coefficient (r); CV = coefficient of variation; ANOVA = analysis of variance; s = seconds; HR = heart rate; LOA = limits of agreement

2.3.5 Protocols and procedures during testing

The first use of the Bruce protocol was in 1963[17] where the test started at the standard 2.7 km·h^{-1} speed and 10% grade but had larger increments in speed and grade in subsequent three-minute stages. The Standard Bruce protocol[1] established in 1973 was used for five studies with no warm up[1,53-56], while the study by Hall-López et al.[60] used a five-minute warm up followed by a rest prior to testing with the standard protocol. Various modifications of the protocol were followed by three studies[57,59,61] (Table 2.2). Finally, Cooke et al.[58] stated that either the standard or modified Bruce protocol was completed and repeated by their participants, but did not specify which participants followed which protocol. The number of repetitions to determine repeatability was two repetitions in eight studies[1,17,53,56,58-61], while Froelicher Jr et al.[54] and Ho[55] held three repetitions each, and Fielding et al.[57] conducted five repetitions in their study. There was variation in the gap between repeated tests, with studies ranging from ten minutes[60] to a possible nine weeks[55].

2.3.6 Repeatability of $\dot{V}O_{2max}$ measurement

The original study by Bruce et al.[17] gave no specifics regarding repeatability, merely a statement that the $\dot{V}O_{2max}$ values measured were "virtually identical" between the first and second tests. Seven of the studies[1,53,56,57,59-61] reported correlation coefficients varying between 0.75[57] to 0.99[1]. The coefficient of variation, used in seven of the studies to determine repeatability of $\dot{V}O_{2max}$ measured[54-59,61], ranged from 4.4%[54] to 9.7%[55].

Five of the studies used paired-samples t-tests or analysis of variance (ANOVA) as part of their analysis to determine repeatability of $\dot{V}O_{2max}$/$\dot{V}O_{2peak}$ measurements between tests[55,57-59,61]. The studies by Ho[55] and Jakovljevic et al.[59] did find significant differences between the repeated $\dot{V}O_{2max}$/ $\dot{V}O_{2peak}$ measurements, while the studies by Fielding et al.[57], Cooke et al.[58] and Harwood et al.[61] did not. Only two studies used limits of agreement[58,59], with Jakovljevic et al.[59] reporting a mean difference of -1.3 ml·kg^{-1}·min^{-1} in $\dot{V}O_{2max}$ between repeated tests, and a range in limits of agreement of -5.3 to 2.7 ml·kg^{-1}·min^{-1} which are assumed to be the 95% confidence intervals (95% CI). Cooke et al.[58] reported limits of agreement to be -25 to 55.6ml·min^{-1} – however, the results are difficult to interpret as the graphical representation of the limits of

agreement was much larger than the stated limits in the text. All studies concluded that the Bruce protocol had good repeatability in measuring $\dot{V}O_{2max}$.

2.3.7 Duration of repeated tests

The same statistical tests were used to analyse the duration of the repeated exercise tests in five of the studies, with varying results[54,55,57,59,61]. Correlation coefficients showed a strong relationship between repeated exercise test length in four studies[56,57,59,61], with r values ranging from 0.74[56] to 0.967[61]. The coefficient of variation between test durations in three studies were found to be 5.5%[55], 7%[56] and 7.1%[59]. Harwood et al.[61] reported "no significant difference" between test durations, while the study by Froelicher Jr et al.[54] concluded that test duration significantly increased from the first to second test, and from the second to third test – however, neither study specified which test of significance was used. A paired-samples t-test showed a significant increase from the first to the second test in the study by Jakovljevic et al.[59] (p<0.05), and this study was again the only one to use limits of agreement to analyse the difference between test durations; with a mean difference of 61 seconds, with assumed 95% confidence interval limits of agreement between -183s to 61s.

2.4 Discussion

2.4.1 Search strategy and quality assessment

The studies included in this review span six decades, reflecting the long period of time that the Bruce protocol has been in use. Despite this, only eleven studies were identified, through a detailed and in-depth search, that examined the protocol's repeatability. The quality of reporting was generally low, especially in the earlier studies[1,17,53-56] and that by Hall-López et al.[60]. Relating to selection of participants, only one study[61] justified their sample size, as well as providing a flowchart of participants included and excluded. No study discussed "non-responders", described by Downes et al.[52] as a difficult area to address due to the very limited information available to researchers about those who do not respond to recruitment. The authors of the AXIS tool stated that some baseline characteristics may be available, and that methods of recruitment should be clearly documented, to allow readers to determine whether the

study sample is representative of the target population: something that none of the included studies described.

Most studies did not clarify the level of statistical significance set in their analysis, as well as being unclear in the statistical methods used. As the repeatability of the Bruce protocol is dependent on the statistical results comparing repeated measures, this shortfall in the risk of bias assessment lowers the quality and reliability of results from included studies.

2.4.2 Participants

There was a majority of male participants across the eleven studies, compared to females. Participant numbers also ranged greatly across the studies, with eight of the eleven studies including less than twenty participants[17,54,56-61]. Smaller sample sizes can lower the weight or impact that statistical results provide, and call into question the validity of the study findings[62-64]. These small studies were included in this review despite their sample sizes, due to the low volume of research available examining the repeatability of the Bruce protocol.

2.4.3 Heterogeneity of pre-test instructions

There was wide variation in pre-test instructions given to participants across the included studies. Generally, earlier studies[1,17,53] as well as the later Harwood et al.[61] study gave no details on pre-test instructions or considerations. There were clearer instructions regarding duration of fasting, levels of physical activity allowed during the testing period and resting on the testing days given in the remaining studies[54-60]. To ensure accurate estimation of repeatability, it is important that there is consistency between each test repetition regarding how a participant prepares for the day, and what they do in the time between tests. This allows for clearer analysis of results of the repeatability of the treadmill test itself, as differences and changes in results can be attributed to the test, rather than variation in the participants from one day to the next. By failing to address these variables, the reliability and validity of the earlier studies' repeated tests was weakened, as it could not be confirmed that identical test procedures and conditions were met at each repetition.

2.4.4 Variation in protocol design and procedures

Consistency in testing procedures and protocols is also vital when testing repeatability. Variation was again noted between studies in the protocols followed, mainly with the first stage of the test varying in speed and grade. There was also some inconsistency within the descriptions of the study protocols in the studies themselves, regarding handrail use, durations between repeated tests, and or if familiarisation with the protocol was allowed.

It is known that $\dot{V}O_{2max}$ increases with exercise training[65-67] or can deteriorate over time if no exercise training is maintained[68-70]. Therefore, a short, consistent duration between repeated tests is preferable. The later studies[56,57,59,61] strove to achieve this, with repeat testing generally one week apart and no test repeated on the same day. The earlier studies were vague when describing the duration between tests, stating only "several days"[17], "weeks"[53], and "days to several weeks"[1], as was the study by Cooke et al.[58] which stated there were "at least four weeks" between repeated tests. It is not clear if participants all had an equal duration between their repeated tests, and so findings could have been influenced by this inconsistency. In repeatability testing, every attempt should be made to keep consistency within variables wherever possible. However, participant availability, change in schedule, and availability of the equipment to be used can account for some small variation in duration between testing, as with human participants and with exercise testing it can be difficult to keep every variable controlled.

The shortest duration between tests was in the study by Hall-López et al.[60], who conducted both repetitions of the maximal test on the same day with ten minutes rest between them. No adverse outcomes were reported from doing two maximal treadmill tests in one day; however, it is difficult to see how fatigue would not impact the repeatability measured in this study.

Another issue with the described protocols related to handrail use during testing, which was not clarified in most of the studies[17,53,56,58-61], despite handrail support being known to increase exercise test duration and reduce submaximal $\dot{V}O_2$ and heart rate (HR) measurements[71,72].

2.4.5 Familiarisation and Learning effect of treadmill testing

A learning effect is when the duration of the test, or the physiological results measured, are improved in a second test, because the subject has become familiar with what the test entails[73,74]. This is therefore an important factor that needs to be considered in this review when exploring the repeatability of the Bruce protocol. The earlier studies[1,17,53], and that by Cooke et al.[58], did not take learning effect into consideration when describing their protocols. Participants had an opportunity to walk on the treadmill in the studies by Nordrehaug et al.[56], Fielding et al.[57] and Hall-López et al.[60], for "familiarisation" with the testing equipment, but no further discussion on the topic was provided. Harwood et al.[61] allowed familiarisation with the treadmill prior to testing, stating that this was done to avoid any learning effect in the results. The authors attributed the small differences in results between repeated tests to variations in participant motivation and participant effort, and did not further discuss learning effect.

Three studies examined learning effect with the Bruce protocol[54,55,59], each dealing with the topic differently. Froelicher Jr et al.[54] did not take any measures prior to testing to compensate for learning effect, with no warm-up or familiarisation provided. They noted an increase in treadmill test duration from first test to second and third tests, and reduced heart rate and oxygen consumption for similar workloads between the first test, and the second and third tests completed. They hypothesised that this was caused by reduced anxiety in participants when completing a familiar test and noted that "habituation" must be considered when examining the repeatability of a treadmill test. In contrast, Ho[55] allowed their participants to familiarise with the treadmill and the gas analysis equipment before testing, so that "test results would not be affected". Their participants had a significant increase in $\dot{V}O_{2max}$ and duration between their first and second tests. The authors put these differences down to learning effect, as there was no significant difference between second and third repeats of the Bruce protocol, and concluded that the test was repeatable. The studies by both Froelicher Jr et al.[54] and Ho[55] examined other protocols besides the Bruce protocol, with participants completing multiple treadmill tests over nine[54] or fifteen[55] weeks. As each participant becomes more familiar with treadmill exercise testing over the testing period, it is possible that learning effect could have caused improvement in results, rather than a true variation

in the measurement ability of the protocol itself. Jakovljevic et al.[59] also allowed walking practice on the treadmill as a familiarisation method and found an increase in treadmill test duration from first to second test. They reasoned that this was due to a learning effect, but concluded that as the change in $\dot{V}O_{2max}$ measured between repeated tests was minimal (1.3 $ml^{-1} \cdot kg^{-1} \cdot min$), this learning effect was most unlikely to have "clinical implications".

Other studies looking at the learning effect in exercise testing found varying results. Some found a definite learning effect for increased duration or oxygen consumption on repeated exercise testing[74-78]. Other studies queried the influence of a learning effect in cardiopulmonary exercise testing, regarding its effect as minimal and not clinically important[79,80]. These studies examined cycle ergometer testing instead of treadmill testing, however, and the difference in mode of exercise could play a role in the size of the learning effect. A participant's expectations of testing may also influence their exercise test outcomes. Several studies compared outcomes during known and unknown durations of exercise in their participants. The rating of perceived exertion (RPE) increased, while heart rate and submaximal $\dot{V}O_2$ measurements reduced, during exercise of unknown duration[81,82]. The authors hypothesised that with an unknown exercise duration, participants become physiologically more economical, to prepare for the exercise duration that may be to come. Arguably, once a person has completed one maximal Bruce protocol test, they may have higher or improved results on a second test, simply by now understanding what to expect, how long the test may last, and how difficult it may feel to them during the test. The attention paid to learning effect in previous Bruce protocol studies has been minimal, although clearly it can have a significant influence on results.

2.4.6 Differences in statistical analysis of data

Correlation coefficients (CC) are used to determine the strength and direction of the relationship between two variables[83]. This statistic was used in seven of the ten studies in this review to compare repeated $\dot{V}O_{2max}$ measurements[1,53,56,57,59-61]. High correlation results led the authors to conclude that there was a high level of repeatability for the Bruce protocol. The use of correlation coefficients to determine agreement between

two variables was challenged by Bland and Altman[84], as correlation coefficients do not demonstrate the difference or variation occurring between the two sets of results. Since the same variable (i.e. $\dot{V}O_{2max}$) is being measured in the same participant using the same test, it is expected that there will be a strong relationship between both test results – as the participant's $\dot{V}O_{2max}$ should not have changed between the two test sessions. As correlation coefficients do not have the ability to determine repeatability, claims of repeatability based on correlation coefficients are hard to support.

The coefficient of variation (CV) describes how much a set of values varies from its mean, and is calculated by dividing the mean by the standard deviation, then multiplying this by 100 to get a percentage of variation[85]. The CV was another popular method of analysis used to examine repeatability. The within-participant variance is most important for repeatability studies, to see the change in measurement from a person's first test to their second[86]. If the coefficient of variation is calculated between the participants, this will not give any information regarding the repeatability – merely how the participants vary from one another. Of the seven studies that used CV, four did not explain their methods of calculation[55,57,58,61]. Froelicher Jr et al.[54] obtained the CV for each individual across their three trials of the Bruce protocol, then averaged the fifteen CVs to get an overall CV for the Bruce protocol's repeatability. The researchers of the final two studies did explain their methods, however there is some variation between the descriptors used. Jakovljevic et al.[59] used the within-person standard deviation (defined by Bland and Altman[87] as the common standard deviation of repeated measurements from a group of people) in their calculation, while Nordrehaug et al.[56] used just a standard deviation of repeated measurements – not specifying if this is within-subject, or an average of the subjects standard deviations. A further comment on the use of CVs by Froelicher Jr et al.[54] and Ho[55] is that, as the authors were analysing other protocols alongside the Bruce protocol, they focused more on the similarities of the CVs between the different protocols and how all were equally repeatable than on the significance of the Bruce protocol's CV itself. With differing methods of calculating the coefficient of variation, and a difficulty in confirming which methods were used by each author, it is challenging to compare the results from each study, and come to a firm conclusion on the repeatability of the Bruce protocol in measuring $\dot{V}O_{2max}$.

Bland and Altman[84] described how to plot the limits of agreement (LOA) of two data sets, by comparing the differences between two measurements to the mean of both measurements, and determining the range of variation in the repeated data. They stated that 95% of all data points should lie between ± 1.96 standard deviations from the mean of the collected data; therefore, outliers as well as any variation between repeated tests will be clearly identifiable, and will give a better interpretation of the repeatability of the test. Jakovljevic et al.[59] and Cooke et al.[58] were the only authors that looked at LOA. Jakovljevic et al.[59] found a 95% LOA of -5.3 $ml \cdot kg^{-1} \cdot min^{-1}$ to 2.7 $ml \cdot kg^{-1} \cdot min^{-1}$, with a mean difference of -1.3 $ml \cdot kg^{-1} \cdot min^{-1}$. Despite this range of 8 $ml \cdot kg^{-1} \cdot min^{-1}$, the authors concluded that their protocol was sufficiently repeatable. Cooke et al.[58] reported their results in $ml \cdot min^{-1}$ for $\dot{V}O_{2max}$ and found LOA between -25 to 55.6 $ml \cdot min^{-1}$ (mean difference 15.8$ml \cdot min^{-1}$). However, the graph presented by Cooke et al.[58] demonstrating the LOA of $\dot{V}O_{2max}$ does not match the results in the text, therefore making comparison to the other studies difficult – as well as reducing the reliability of the results from the Cooke et al.[58] study. As the results were presented in $ml \cdot min^{-1}$ without details of bodyweight by Cooke et al.[58] it is also difficult to convert results to the same units to compare these limits of agreement for this review. However, the range in $\dot{V}O_{2max}$ differences demonstrated in both studies is wide, and neither study considered the possibility that this range is due to inconsistency in repeatability of the Bruce protocol itself.

2.5 Conclusion

Research into the reliability and repeatability of the Bruce protocol spans several decades. Evidence based on correlation coefficients and statistical tests has led to conclusions that the Bruce protocol in repeatable. Both studies that used limits of agreement to determine repeatability, however, found a wide range in $\dot{V}O_{2max}$[58,59] and duration[59] between the repeated measurements. This indicates a further need for investigation into the repeatability of the Bruce protocol. A consistent, standardised protocol for how to conduct the Bruce protocol treadmill test, with clear instructions and processes is required to ensure accurate testing and repeatability every time it is

used; this currently does not exist, as can be seen by the variations from study to study in the methods for conducting the Bruce protocol. Learning effect and its role in repeatability studies should also be more closely monitored and studied in the future, as it has been shown to affect test duration, and submaximal $\dot{V}O_2$ and heart rate values. This may have an influence on $\dot{V}O_{2max}$ measurements and on the repeatability conclusions formed regarding the Bruce protocol.

Chapter 3: Literature Review – A review of the reliability of equations used to predict $\dot{V}O_{2max}$ used alongside the Bruce protocol

3.1 Introduction

Cardiorespiratory fitness is commonly viewed as one of the most important indicators of cardiovascular disease risk and a measure of all-cause mortality[2]. The gold standard method of determining a person's $\dot{V}O_{2max}$ is through maximal exercise testing and analysis of expired air[2,4,88]. In certain cases a maximal test is not possible, due to medical conditions or physical limitations[30], or if testing equipment is not available, or to save on time and cost[31-33]. Prediction equations can be used in these cases to calculate $\dot{V}O_{2max}$ either using submaximal exercise test data[89-91], or by using non-exercise variables such as gender, body mass index or physical activity levels[92-94]. However, there is the risk that predicting $\dot{V}O_{2max}$ instead of directly measuring it through maximal exercise testing, that the value will be over- or under-estimated[34-36]. This could influence future interventions and management for the tested person[2].

There are many varying equations used to predict $\dot{V}O_{2max}$, incorporating both exercise and non-exercise variables to determine a person's cardiorespiratory fitness. It has been shown that equations can be population-specific, with varying accuracy and reliability depending on the population with which they are used, and therefore equations should only be used to estimate $\dot{V}O_{2max}$ in the population for which it was devised[35,95-99]. To control for the differing effect of exercise on healthy populations versus those with clinical conditions, the aim of the current review was to examine prediction equations developed for use with the Bruce protocol, or compared to data taken from the Bruce protocol, to determine which specific equation, or which variables, best influenced the accurate prediction of $\dot{V}O_{2max}$. The Bruce protocol was chosen, as it is the most commonly used protocol for cardiopulmonary exercise testing.

3.2 Methods

3.2.1 Search Strategy

The Preferred Reporting items for Systematic Reviews and Meta-Analysis (PRISMA) statement was used to guide development of this literature review[50]. A search strategy was developed to search four databases (Embase, CINAHL, Web of Science, and Medline) for articles that compared $\dot{V}O_{2max}$ predicted through equations, to $\dot{V}O_{2max}$ measured using the Bruce maximal treadmill test protocol. Search terms included "exercise test" and "treadmill test" combined with "Bruce", and these were cross-searched with terms such as "evaluation study", "prediction", and "validation study" (full list of search terms in Appendix 1, pg. 104).

3.2.2 Eligibility criteria

The inclusion criteria were as follows:

- articles written in English;
- at least 20 healthy participants aged 18 years or older;
- $\dot{V}O_{2max}$ directly measured with participants completing a maximal Bruce treadmill protocol;
- $\dot{V}O_{2max}$ predicted using equations that require either data taken from the Bruce protocol exercise test, or resting/anthropometric data;
- inclusion of the specific equation used, or clear citation of where the equation originated.

Studies were excluded if:

- participants had any cardiac condition;
- the study examined the effect of exercise, drug, or dietary interventions;
- the study used magnetic resonance imaging (MRI) scanning during testing, or assessed the diagnostic value of ECG changes;
- the study used non-human participants; or
- the study formulated regression equations to predict $\dot{V}O_{2max}$ but did not validate their equation.

3.2.3 Study selection and data extraction

The search was initially carried out in February 2019, and repeated in July 2019, to identify any new publications. Using the screening and data extraction software, Covidence, the titles and abstracts of the full list of search results were reviewed independently by two researchers and compared to the inclusion and exclusion criteria. The full texts of selected articles were then read and analysed by the same two

researchers, and included for the literature review if deemed appropriate. Any conflicts in decisions were discussed between the two researchers until agreement was reached. Reference lists from included articles were then reviewed, and any studies fitting the inclusion criteria were also included.

Following this process, the studies chosen for inclusion underwent quality appraisal, and were then analysed, with data extracted into a standardised template.

3.2.4 Quality appraisal

Risk of bias and the quality of reporting was assessed using the Appraisal tool for Cross-Sectional Studies (AXIS)[52] (Appendix 2, pg. 105). This tool has twenty questions regarding reporting, including areas such as study aims, participant recruitment, description of methods used and statistical analysis done, as well as consistency of results and discussion provided. No scoring system is recommended by Downes et al.[52]; therefore, for the purposes of this review, a point was awarded for every "yes" answer (excluding questions 13 and 19, for which a "no" answer was considered positive, and was awarded one point). Each study was given a score out of twenty, and the higher the score, the better the quality of reporting in the study.

3.3 Results

3.3.1 Study selection

The PRISMA flow diagram (Figure 3.1) demonstrates the results of the study screening and selection process. After duplicates were removed, a total of 1358 studies were identified. Following screening, a final total of seven studies were included for review[31,100-105]. The full breakdown of the data extracted from each study can be seen in Tables 3.1 and 3.2.

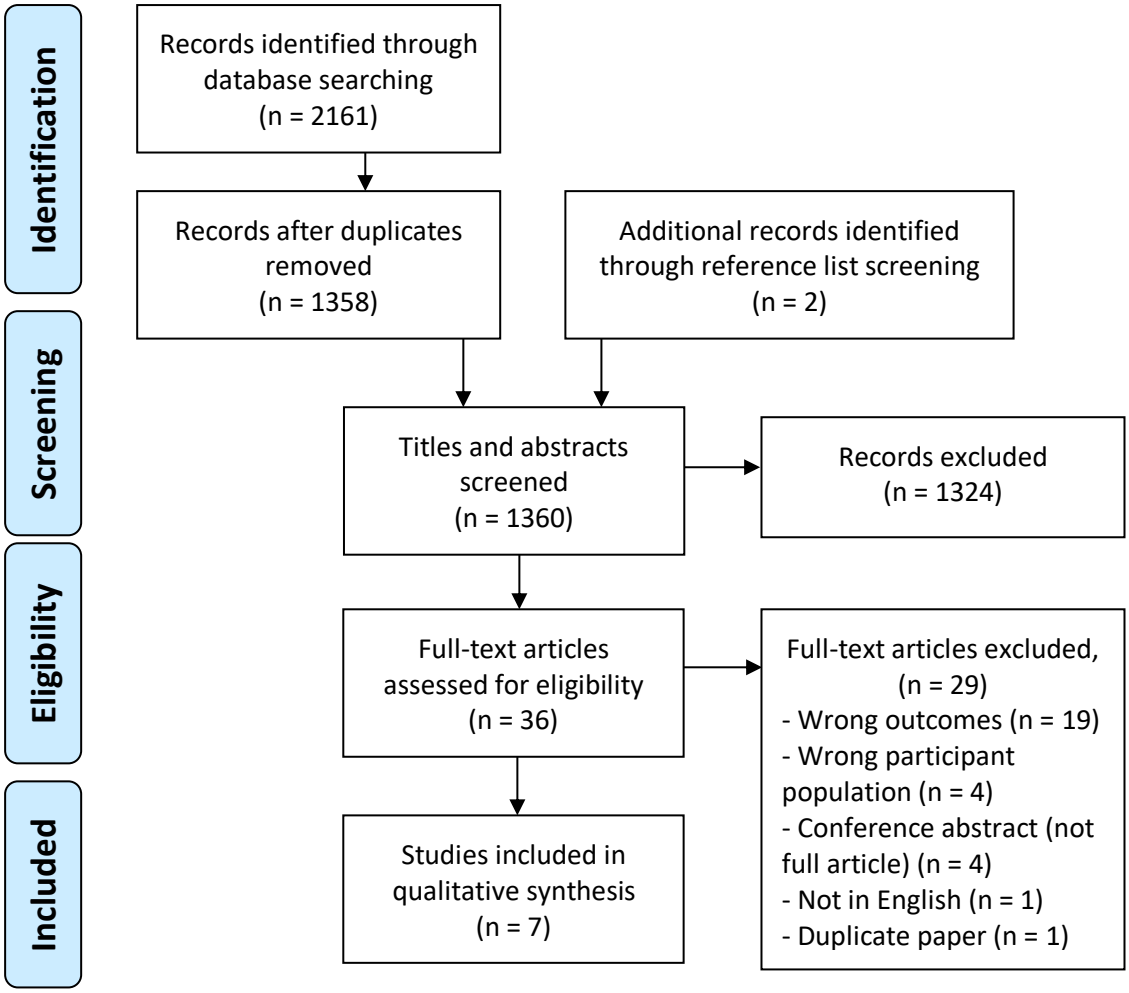


Figure 3.1 Flowchart for literature selection process

3.3.2 Quality Assessment

The mean score across the seven studies from assessment with the AXIS tool[52] was 11.14/20 (±2.97) (Table 3.3). Across the board, the studies were strong in identifying their aims, choosing an appropriate study design, measuring suitable variables, and providing internal consistency when displaying their results. They were all poor, however, in justifying their sample size, and discussing "non-responders". There was no discussion in any study regarding drop-out of participants, the specific process of inviting volunteers, or the selection process utilised. The study by Jackson et al.[100] had the lowest score, and was poor in defining its target population and included participants, as well as lacking sufficient detail regarding testing methods to allow for the study to be repeated. The study by Koutlianos et al.[103], in contrast, scored relatively highly, only losing marks in areas where all seven studies were lacking (justifying sample size and defining "non-responder" information).

Table 3.1 Comparison of study details across seven predictability equation studies (a)

	Participant information	Instructions / familiarisation	Equation used; *(maximal or submaximal Bruce treadmill test)*	Variables used in Equations
Jackson et al.[100]	1,604 males aged 25-70 years	*Nil specified*	Regression eq.: $\dot{V}O_{2max}$ = 47.9 – (0.27 x age) + (3.41 x PA questionnaire) – (0.2 x fat % x PA questionnaire); ***(variables are non-exercise, maximal or submaximal test not applicable)***	● Age (years) ● PA questionnaire (score) ● Fat % (body fat as %)
Grant et al.[101]	15 males, and 15 females, aged 18-35 years	3 minutes walking on treadmill for familiarisation	Eq. 1: Male $\dot{V}O_{2max}$ = 3.88 + 0.056 x duration[107]; ***(maximal treadmill test)*** Eq. 2: Female $\dot{V}O_{2max}$ = 1.06 + 0.056 x duration[107]; ***(maximal treadmill test)*** Eqs. as above for male (Eq. 3)/female (Eq. 4) for 85% $\dot{V}O_2$ *(use duration from submaximal Bruce protocol to 85% HR Reserve).* $\dot{V}O_{2max}$ = (85% $\dot{V}O_2$ x 0.174) + 85% $\dot{V}O_2$; ***(submaximal treadmill test)***	Duration (seconds)
Maeder et al.[102]	33 males, 10 females, aged 30-41 years	No vigorous PA on day of testing	American College of Sports Medicine (ACSM) eq.; ***(maximal treadmill test)*** *Not specified if 'walking' or 'running' eq. used.*	● TM speed (m·min^{-1}) ● Grade (% TM incline, decimal)
Lee et al.[31]	32 males, and 16 females, aged 18-59 years	No food/PA x 4 hours before testing. ? familiarisation: submaximal Bruce completed first.	1. ACSM walking eq.: $\dot{V}O_2$ = (speed x 0.1) + (speed x 1.8 x grade) + 3.5; ***(submaximal treadmill test).*** *Calculate the above for stages 1-3 of protocol, then calculate line of best fit through 3 data points relating to HR and predicted* $\dot{V}O_{2max}$*. Extrapolate to predicted* HR_{max} *for* $\dot{V}O_{2max}$ 2. Fitmate eq.: $\dot{V}O_{2max}$ = linear regression between measured HR and measured $\dot{V}O_2$ during submaximal test, extrapolated to HR_{max} predicted; ***(submaximal treadmill test)*** 3. Fitmate eq. with measured HR_{max}: Linear regression of submaximal measured HR and $\dot{V}O_2$ data (as above), extrapolated to measured HR_{max} (measured during maximal Bruce protocol) to estimate $\dot{V}O_{2max}$***; (submaximal treadmill test)***	● TM speed (m·min^{-1}) ● Grade (% TM incline, decimal) ● HR (beats per minute) ● HR_{max} predicted = 220 - age (in years)
Koutliano s et al.[103]	55 male athletes, mean age 28.3 years	Avoid caffeine & alcohol x 2 hours before testing	1. ACSM running eq.: $\dot{V}O_2$ = (0.2 x speed) + (0.9 x speed x grade) + 3.5; ***(maximal treadmill test)*** 2. Enter regression eq.: $\dot{V}O_{2max}$ = 58.443 – (0.215 x age) – (0.632 x BMI) – (68.639 x grade) + (1.597 x duration); ***(maximal treadmill test)*** 3. Stepwise regression eq.: $\dot{V}O_{2max}$ = 33.971 – (0.291 x age) + (1.481 x duration); ***(maximal treadmill test)***	● TM speed (m·min^{-1}) ● Grade (% TM incline, decimal) ● Age (years) ● BMI (kg·m^{-2}) ● Duration (minutes)
Nitin et al.[104]	20 males aged 18-30 years	No food x 3 hours before testing	1. $\dot{V}O_{2max}$ = (0.046 x weight) – 0.012 2. $\dot{V}O_{2max}$ = (0.04 x weight) + 0.232 3. $\dot{V}O_{2max}$ = (0.018 x weight) + 1.212 ***(Variables are non-exercise, maximal or submaximal test not applicable)***	Body weight (kg)
Crouse et al.[105]	472 male college students aged 17-25 years	*Nil specified*	1. Bruce active males eq.: $\dot{V}O_{2max}$ = (3.788 x duration) + 0.19; 2. Foster eq.: $\dot{V}O_{2max}$ = 14.8 – (1.379 x duration) + (0.451 x duration^2) – (0.012 x duration^3); 3. Football eq. (regression eq. developed in current study): $\dot{V}O_{2max}$ = (4.017 x duration) – 4.644; ***(all three eqs. are for maximal treadmill test)***	Duration (minutes)

Key: PA = physical activity; Eq. = equation; Eqs. = equations HR = heart rate; HR_{max} = maximal heart rate; TM = treadmill; BMI = body mass index.

Table 3.2 Comparison of study details across seven predictability equation studies (b)

	Statistics used: Measured vs Predicted $\dot{V}O_{2max}$	**Results comparing Measured and Predicted $\dot{V}O_{2max}$**
Jackson et al.[100]	CC	CC: r = 0.79. **Conclusion**: satisfactory correlation / equation predicts accurately
Grant et al.[101]	● CC; ● Standard error of estimate used to calculate CV	1. Males Bruce max test: CC: r = 0.49 ($p < 0.1$); CV = 8.8%. 2. Females Bruce max test: CC: r = 0.92 ($p < 0.01$); CV = 7.4%. 3. Males Bruce Submax test: CC: r = 0.59 ($p < 0.05$); CV = 8.1%. 4. Females Bruce Submax test: CC: r = 0.82 ($p < 0.01$); CV = 11.9% **Conclusion:** Bruce equations valid for use in females; inconclusive results for use in males
Maeder et al.[102]	CC	CC: r = 0.52 **Conclusion:** Bruce protocol with ACSM equation gives poor prediction of $\dot{V}O_{2max}$
Lee et al.[31]	● Repeated measures ANOVA to compare measured $\dot{V}O_{2max}$ and all predicted $\dot{V}O_{2max}$ values ● CC to individually compare each predicted $\dot{V}O_{2max}$ to measured $\dot{V}O_{2max}$ ● LOA plots comparing each equation individually to the measured $\dot{V}O_{2max}$	1. ACSM equation: ANOVA: significant difference ($p = 0.01$); CC: r = 0.758 ($p < 0.01$) 2. Fitmate equation: ANOVA: no significant difference ($p = 0.152$); CC: r = 0.897 ($p < 0.01$) 3. Fitmate with measured HR_{max}: ANOVA: significant difference ($p = 0.01$); CC: r = 0.894 ($p < 0.01$) **Conclusion:** Better estimate for $\dot{V}O_{2max}$ with Fitmate equation/method, rather than with ACSM equation
Koutlianos et al.[103]	CC	CC: 1. ACSM equation: r = 0.27 2. Enter regression equation: r = 0.64 3. Stepwise regression equation: r = 0.61 **Conclusion:** ACSM equation over-estimates $\dot{V}O_{2max}$; other equations are accurate
Nitin et al.[104]	LOA, results compared using paired t-tests	CC: Equation 1: r=0.15; Equation 2: r=0.29; Equation 3: r=0.81 (systematic error noted) **Conclusion:** Equations 1 and 2 significantly overestimated $\dot{V}O_{2max}$ (based on LOA plots). Equation 3 had systematic error (based on LOA plot).
Crouse et al.[105]	Repeated measures ANOVA to compare between equations. *Not specified: statistics used to compare measured vs predicted $\dot{V}O_{2max}$*	Significant difference in predicted $\dot{V}O_{2max}$ between equations ($p < 0.001$). Bruce 4.4% higher, and Foster 2.2% lower than Football equation $\dot{V}O_{2max}$ prediction. Bruce and Foster equations: significantly different predicted $\dot{V}O_{2max}$ compared to measured ($p = 0.001$).

Key: CC = correlation coefficients; CV = coefficient of variation; ANOVA = analysis of variance; LOA = limits of agreement; max = maximal; submax = submaximal; ACSM = American College of Sports Medicine; HR_{max} = maximal heart rate.

Table 3.3 AXIS risk of bias assessment[52]

	Jackson et al.[100]	Grant et al.[101]	Maeder et al.[102]	Lee et al.[31]	Koutlianos et al.[103]	Nitin et al.[104]	Crouse et al.[105]
1.	Yes	Yes	Yes	Yes	Yes	Yes	Yes
2.	Yes	Yes	Yes	Yes	Yes	Yes	Yes
3.	No	No	No	No	No	No	No
4.	No	No	Yes	No	Yes	Yes	Yes
5.	No	Unknown	Yes	Unknown	Yes	No	Yes
6.	No	Unknown	Unknown	Unknown	Unknown	Unknown	Unknown
7.	Unknown	Unknown	Unknown	Unknown	Unknown	Unknown	Unknown
8.	Yes	Yes	Yes	Yes	Yes	Yes	Yes
9.	No	Yes	Yes	Yes	Yes	Unknown	Yes
10.	No	Yes	Yes	Yes	Yes	Yes	Yes
11.	No	Yes	Yes	Yes	Yes	No	Yes
12.	Yes	No	Yes	Yes	Yes	Yes	Yes
13.*	Unknown	Unknown	Unknown	Unknown	Unknown	Unknown	Unknown
14.	No	No	No	No	No	No	No
15.	Yes	Yes	Yes	Yes	Yes	Yes	Yes
16.	Yes	No	Yes	Yes	Yes	No	Yes
17.	Yes	Yes	Yes	Yes	Yes	Yes	Yes
18.	No	No	Yes	No	Yes	No	No
19.*	Unknown	Unknown	Unknown	Unknown	No	Unknown	No
20.	Unknown	Yes	Unknown	Yes	Yes	Yes	Yes
Total Positives	7/20	9/20	13/20	11/20	15/20	9/20	14/20

"Yes" = 1 point, "No" and "Unknown" = 0 points.

*For Questions 13 and 19, "Yes" is a negative answer (therefore awarded 0 points), and "No" is a positive answer (therefore awarded 1 point).

3.3.3 Participants

The number of participants included in each study varied greatly, from twenty participants[104] to 1,604 participants[100]. The majority of participants were male (98%) and the ages ranged from 17 to 70 years, across the seven studies. No participants had any cardiovascular disease, or other health conditions.

3.3.4 Study Design and protocols followed.

All seven studies followed a cross-sectional study design. The standard Bruce protocol was used in all studies, excluding that by Nitin et al.[104] which modified it to stop increasing the incline after the seventh stage (due to treadmill technical difficulties). The study by Grant et al.[101] followed the standard Bruce protocol, but also took data at the point in testing when 85% of maximal heart rate reserve was reached. All studies used an indirect breath-by-breath method to analysed expired gases, except that by Lee et al.[31] which used the Douglas bag method. Some researchers provided specific instructions to participants in preparation for exercise testing; Maeder et al.[102] asked participants to refrain from vigorous physical activity on the day of testing, Lee et al.[31] instructed participants to fast and refrain from physical activity for four hours prior to testing, and Koutlianos et al.[103] requested that participants avoid caffeine and alcohol for two hours before testing. To allow familiarisation to the testing protocol, Grant et al.[101] allowed participants three minutes of treadmill walking prior to testing, and all participants in the study by Lee et al.[31] completed a submaximal Bruce treadmill test on a separate day, prior to their maximal treadmill test. No other studies addressed familiarisation. All studies used the $\dot{V}O_{2max}$ measured from the Bruce protocol to determine the accuracy of their chosen $\dot{V}O_{2max}$ prediction equations.

3.3.5 Equations used to predict $\dot{V}O_{2max}$

Three studies generated their own $\dot{V}O_{2max}$ prediction equations and examined their accuracy against a measured $\dot{V}O_{2max}$[100,103,105]. They all found, through analysis by correlation coefficients, that their own equations accurately predicted $\dot{V}O_{2max}$ in their participants (r = 0.61 – 0.82). Lee et al.[31] examined the equations used by the Fitmate device to predict $\dot{V}O_{2max}$[106] which is based on submaximal heart rate and $\dot{V}O_2$ measurements taken during an exercise test. They found that the Fitmate equations

accurately predicted $\dot{V}O_{2max}$ in their participants. Grant et al.[101] reviewed the accuracy of two equations by Poole[107] and two equations by Heyward[108], and concluded that these equations were accurate and valid for use in females, but had inconclusive results for use with males.

The American College of Sports Medicine (ACSM) walking and running equations[109] were examined in three studies[31,102,103], and all researchers concluded that the ACSM equations overestimated $\dot{V}O_{2max}$ for all participants. Nitin et al.[104] studied the accuracy of three previously formulated equations and found that two of these equations significantly overestimated $\dot{V}O_{2max}$[110,111], and that the third equation[112] yielded a systematic error when tested. Finally, Crouse et al.[105] examined two previously established prediction equations and determined that one significantly underpredicted $\dot{V}O_{2max}$[113] and the other significantly overpredicted $\dot{V}O_{2max}$[1].

3.3.6 Variables used to predict $\dot{V}O_{2max}$

There was a mixture of exercise and non-exercise data used to predict $\dot{V}O_{2max}$ across the studies in this review. Non-exercise variables such as a participant's age, percentage body fat, body weight, body mass index (BMI) or their results in a physical activity questionnaire were utilised by several researchers[100,103,104]. Duration of a maximal treadmill test was a common variable, used in three studies[101,103,105]. The speed and percentage grade of the treadmill were variables utilised by the ACSM equations, examined by three studies[31,102,103].

3.4 Discussion

To complete this literature review accurately and thoroughly, strict inclusion/exclusion criteria were developed for study results in the database search. Therefore, only seven articles from the large number (n = 1360) of articles screened initially were suitable for inclusion here. Even with this process of elimination, there was still much variation between the seven included studies, making comparison between methods and results more difficult. This review examines the bias associated with each study included, as well as looking at the different equations used and the variables required for the

equations, and the statistical analysis used to compare the measured $\dot{V}O_{2max}$ to the predicted value of $\dot{V}O_{2max}$.

3.4.1 Evidence of Biases

The quality of reporting in each study was assessed through the AXIS tool[52]. A higher score indicated a higher quality of reporting. There was a wide range between the results found, with the lowest being 7/20 and the highest 15/20. Risk of bias assessments are important to determine the credibility of a study, and in establishing any error or bias in study reporting[114]. They can influence the final conclusion of literature and systematic reviews, with the reliability of results being called into question because of bias, or being commended and supported because of acceptable reporting. While the studies included were all strong in reporting their aims, and ensuring a suitable study design and appropriate study methods, they were poor in justifying their sample size, describing their participant selection process, and discussing "non-responders". This can raise some questions about selection bias across the studies. Three of the studies[100,103,105] clearly defined their target populations, and it was evident from their described participants that their sample was representative of this population. The other studies, however, either did not state to whom they wished to apply their findings[31,101], or their selected participants did not appear to accurately represent their target population[102,104]. For example, Nitin et al.[104] aimed to estimate $\dot{V}O_{2max}$ in a "healthy Indian population", yet their participants were healthy males recruited from a college campus, between eighteen to thirty years of age – which this review argues is not representative of an entire country's healthy population. Often those who volunteer for research studies, do so for specific reasons or outcomes: for example, with exercise studies, it may be reasoned that the volunteers are likely to be those already physically active, or (if sampled from a college population) involved in exercise-science departments[115,116]. With little information regarding recruitment processes, and no information given in any included study regarding "non-responders", it is hard to define whether attempts were made to sample wider, more representative groups, or if there were any specific reasons or patterns notable in those not volunteering to participate in these exercise studies.

3.4.2 Equations and Variables

There were a variety of equations used within the seven studies in this review. They were divided into four main categories:

1. those using duration of maximal test alone;
2. those using grade and speed of treadmill (with or without extrapolation to maximal heart rate);
3. those using non-exercise data (age, Body Mass Index (BMI), body weight, body fat percentage, physical activity levels);
4. those using a combination of the above.

Duration was a common variable across the equations used. Both Grant et al.[101] and Crouse et al.[105] used equations solely based on the maximal test duration. Grant et al.[101] additionally looked at predicting $\dot{V}O_{2max}$ from their submaximal data collected, based on submaximal test duration. Both studies found that their prediction equations on male subjects incorrectly predicted $\dot{V}O_{2max}$ when compared to the measured value. The equation by Bruce et al.[1] over-predicted the result, while the equation by Foster et al.[113] under-predicted the $\dot{V}O_{2max}$ value (both used in the Crouse et al.[105] study). Grant et al.[101] used differing equations for men and women and did find that the results for the female predicted $\dot{V}O_{2max}$ was strongly correlated to the measured $\dot{V}O_{2max}$ value. It has been shown that the duration of a treadmill test can vary with familiarisation and experience, as well as with encouragement given by the tester to the participant[74,77,78,117]. With this variability, it is not surprising that equations basing their predictions solely on test duration fail to accurately estimate the maximal $\dot{V}O_2$. The only other study using the test duration as part of its prediction was that by Koutlianos et al.[103]. The authors created two of their own equations (by enter regression and stepwise regression) and combined duration with other variables not relating to the exercise test completed. They concluded that these two equations adequately predicted the $\dot{V}O_{2max}$; however, they based this on correlation coefficients between measured and predicted data that equal 0.64 and 0.61 (p<0.001 for both) for enter and stepwise regression equations (respectively), which is considered only moderate strength correlation[118,119].

The American College of Sports Medicine (ACSM) equations calculate $\dot{V}O_2$ based on the speed of the treadmill and the incline, or grade, at which the treadmill is set. Following this method, the $\dot{V}O_2$ for each 3-minute stage will be the same for each participant, and

will only vary depending on which stage each participant reaches. Each of the three studies included here that examined the ACSM walking or running equations found that the ACSM equations over-predicted $\dot{V}O_{2max}$[31,102,103]. This is consistent with findings from other studies that examine the ACSM equations in different populations or by using different treadmill protocols[35,96,120,121]. As $\dot{V}O_{2max}$ is a strong indicator of risk of cardiovascular disease, as well as used to measure improvement in fitness after exercise or medical intervention, having an over-prediction of a person's cardiorespiratory fitness can result in a reduced estimation of their cardiovascular risk, which could greatly impact the intervention they receive.

There is another set of equations that do not use data from the exercise test, but rather base their predictions on non-exercise data such as age, body weight or fat percentage, or on BMI. Nitin et al.[104] limited their three separate equations to using body weight as the only variable. They found that their first and second equations over-estimated the $\dot{V}O_{2max}$ compared to the measured value, and that their third equation had systematic error with its results. Jackson et al.[100] used age and fat percentage to predict the $\dot{V}O_{2max}$ of their participants, along with their scores of the NASA Self Report Physical Activity Scale. They concluded that there was satisfactory correlation between measured and predicted $\dot{V}O_{2max}$ using their equation. Some non-exercise variables can influence the predicted $\dot{V}O_{2max}$ more than others, such as age, or central adiposity[89,122]. Physical activity is another variable that influences $\dot{V}O_{2max}$ – however there is differing opinion on how strong an impact it can have on predicting $\dot{V}O_{2max}$. Cardiorespiratory fitness can be maintained and improved with physical activity – particularly vigorous-intensity physical activity (at approximately 60% of a person's $\dot{V}O_{2max}$[123] or greater than six metabolic equivalents (METs)[124]. Wang et al.[125] suggested that equations including physical activity levels in their calculations have better validity than those without physical activity information, citing that physical activity is the most important influencer of cardiorespiratory fitness. In contrast, a study by Bradshaw et al.[126] showed that the physical activity rating was the least effective variable in predicting $\dot{V}O_{2max}$ (using standardised β-weight scores). Dyrstad et al.[127] found that $\dot{V}O_{2max}$ was influenced more by other factors such as gender, age and BMI, rather than physical activity levels. The varying effect of physical activity levels on the prediction of $\dot{V}O_{2max}$ could be due to

this variable being mainly self-reported, and therefore at risk of recall bias or over- or under-estimation of a person's own activity levels[99,125]. In the current review, Jackson et al.[100] were the only researchers that incorporated physical activity data, in self-reported form, into their prediction equation, finding that physical activity did influence the prediction of $\dot{V}O_{2max}$, and provided an accurate prediction equation.

Estimating $\dot{V}O_{2max}$ using non-exercise equations may be done to save time or money, as this method is considerably more convenient than having participants complete a maximal or even submaximal exercise test[126]. It has been shown that for cycle ergometer tests, equations that utilise *exercise* data are more accurate than equations using *non-exercise* data when tested in the same participants[128]. Looking at treadmill tests, one study examining the effect of multiple types of variables on the prediction of $\dot{V}O_{2max}$ found that equations combining maximal, submaximal and non-exercise data in their predictions were more accurate than equations that combined only two of the three types of data[129]. On the other hand, many authors have stated that non-exercise equations are at least as effective at predicting $\dot{V}O_{2max}$ as equations using submaximal exercise test physiological data[99,126,130-132].

For some populations, such as athletes, or college students, there can be variation between their $\dot{V}O_{2max}$ values, while having similar body composition, age, or gender (i.e. a homogeneous group or population)[133-135]. Equations utilising non-exercise data may fail to predict accurately the differing physiologic responses to maximal or submaximal exercise that people of similar anthropometric make-up might experience. There is also variation in body composition between people of different ethnicities, meaning that non-exercise equations developed for one ethnic group of people may not be applicable or accurate when used with another ethnic group[34,136].

Since it is known that the gold standard method of measuring $\dot{V}O_{2max}$ is through a maximal exercise test with direct analysis of the expired air[4,30], it appears logical to prefer a $\dot{V}O_{2max}$ prediction equation that requires exercise data, over one that bases its result solely on anthropometric data. Prediction equations using non-exercise data are generally favoured specifically because exercise testing is not required to obtain the $\dot{V}O_{2max}$ result; their convenience and simplicity of use in a clinical setting are the reasons

many researchers promote their use[125,131,132]. However, with the wide range of variables available for use, and the differing physiologic effects of exercise and non-exercise variables on people from different anthropometric make-up or different ethnicities, further research is still warranted to determine what type of equation best suits the prediction of cardiorespiratory fitness.

Reviewing the results from equations used in the current review, it can be seen that equations based solely on body weight were inaccurate[104], while an equation incorporating age, physical activity level and body fat percentage yielded a more accurate prediction[100]. The only other study included in this review that used non-exercise data in its equations was by Koutlianos et al.[103], where the two equations used combined non-exercise data relating to age and BMI, with test duration and treadmill grade. The authors concluded that both of their equations accurately predicted $\dot{V}O_{2max}$. It can be seen that including more variables in the equation may lead to a better prediction. Comparing specifically the correlation coefficients reported in the studies by Koutlianos et al.[103] and Jackson et al.[100], it can be seen that the latter study found a stronger correlation for its measurement. This would seem to contradict the theory that equations combining more data types (exercise and non-exercise) predict $\dot{V}O_{2max}$ more accurately than equations with less variables[129]. However, comparing the two studies as a whole, the study by Koutlianos et al.[103] was of a much higher standard than that by Jackson et al.[100]. As seen in Table 3.3, the study by Koutlianos et al.[103] scored much higher on the risk of bias assessment, with the study by Jackson et al.[100] falling down in detailing its statistical methods and analysis, defining its target population, and providing clear details on the study design and methods. This must be taken into account when comparing the results from both studies.

From the above analysis, it is difficult to say which variable, or combination of variables, is the most effective in predicting an accurate $\dot{V}O_{2max}$. The only definitive results are that the ACSM equation based on speed and grade consistently over-predicts $\dot{V}O_{2max}$, and that equations using the duration of the treadmill test completed may not provide fully accurate results either. The best combination of variables for predicting $\dot{V}O_{2max}$ along with the Bruce protocol – or any treadmill protocol – is still to be found.

3.4.3 Statistical Analysis

Across the seven studies included in this review, there was variety in the methods of statistical analysis. Most commonly, correlation coefficients were used to determine the strength of relationship between measured and predicted $\dot{V}O_{2max}$[31,100-103]. Standard error of estimate and coefficients of variation were used to analyse data in the study by Grant et al.[101]. The study by Lee et al.[31] utilised repeated measures Analysis of Variance (ANOVA), standard error of estimate and limits of agreement in its data analysis, and finally Nitin et al.[104] also looked at the limits of agreement in their statistical analysis.

With regard to correlation coefficients, there is debate that these are not the most accurate way to determine agreement between two measurements[84,137]. By the fact that two methods are used to determine the same variable (in this case, measuring $\dot{V}O_{2max}$ with a Douglas bag, or metabolic cart, versus predicting it with an equation in the same sample of participants), it can be expected that the results from both will be related – therefore giving a high correlation coefficient[138]. A high correlation coefficient does not mean that no difference or variation will be found between the $\dot{V}O_{2max}$ results obtained from each assessment method. Concluding that an equation provides an accurate prediction based on correlation coefficients alone, as was done by Jackson et al.[100], Maeder et al.[102] and Koutlianos et al.[103], is therefore not a fully reliable conclusion[139].

The standard error of estimate can be used to determine the accuracy of an estimated mean value, or how much variation is between estimated values[140]. The coefficient of variation (CV) is another measure of how certain values differ from their mean: it is the standard deviation divided by the mean, multiplied by 100, presented as a percentage[141]. Grant et al.[101] used the standard error of estimate of their results (instead of the standard deviation) to calculate the coefficient of variation. This is not the conventional method; however, they state they used this calculation to measure the amount of error occurring with each equation used to predict $\dot{V}O_{2max}$. Generally, a CV below 10% is considered acceptable for normal variation in $\dot{V}O_{2max}$[30,142], and in the study by Grant et al.[101], it was found that their "error", or CV, was 8.8% for the males and 7.4% for the females. The authors concluded that while their equations used for the Bruce

protocol tended towards an over-prediction of $\dot{V}O_{2max}$, they were adequate in relation to previous literature and to the other equations and protocols examined within their study. As they combined several different statistical methods and used varying terms for these, however, the reliability of their results may be questioned.

Bland and Altman[84] disagreed with using correlation coefficients to compare two measurement methods of the same variable. Instead they suggested looking at the limits of agreement, by plotting the difference between two measurements against the mean of those two measurements. Nitin et al.[104] used limits of agreement plots to compare their measured and predicted $\dot{V}O_{2max}$ values. They found that for their first two equations, there was random error shown with the limits of agreement plots, and systematic error with their third equation. This was stated in their results and discussion, but not expanded on or explained further, and it is difficult to determine from their presented plots how exactly they came to their conclusion. The authors presented their $\dot{V}O_{2max}$ values in $L \cdot min^{-1}$, which does not take body weight of participants into account and therefore makes comparison of results between different participants, and across populations, less accurate[143,144].

For each correlation coefficient they calculated, Lee et al.[31] also gave a limits of agreement plot to demonstrate the range of values and bias found for each equation. It is clear to see by comparing the limits of agreement for each of their three equations that the ACSM equation had the highest bias, and the widest range of differences between results, whereas their Fitmate equation using age-predicted $\dot{V}O_{2max}$ had a much smaller bias and range of difference in $\dot{V}O_{2max}$ values.

Repeated measures ANOVA (analysis of variance) was used in the studies by both Lee et al.[31] and Crouse et al.[105]. The former used this to compare between all predicted values of $\dot{V}O_{2max}$ as well as the measured value, and found that there was significant difference between the results of the prediction equations. The authors then used correlation coefficients and limits of agreement to look at the relationship between each individual prediction and the measured $\dot{V}O_{2max}$ (as discussed above). Crouse et al.[105] also used the repeated measures ANOVA to compare the predicted $\dot{V}O_{2max}$ values from their three equations. They did not, however, include the measured $\dot{V}O_{2max}$ in this

comparison, and therefore only concluded that there was a significant difference between the prediction equations. They stated later that the equations by both Foster et al.[113] and Bruce et al.[1] gave a significantly difference $\dot{V}O_{2max}$ prediction compared to their measured $\dot{V}O_{2max}$, but did not explain the statistics they used for this, or elaborate on this point further.

3.4.4 Limitations of Current Review

As previously stated, the Bruce protocol is generally viewed as the most popularly used maximal treadmill protocol to measure $\dot{V}O_{2max}$. Many researchers, however, have explored the accuracy of numerous prediction equations used with other treadmill or cycle ergometer protocols, in both healthy and clinical populations[128,145,146]. To achieve the goal of this literature review in identifying and examining prediction equations specifically used with the Bruce protocol, a strict and detailed inclusion and exclusion criteria for studies was required. There may be other equations, or combinations of variables, that more accurately predict $\dot{V}O_{2max}$ when used with other protocols, that have not been examined in the current review.

Following on from this, the current review looked only at a healthy population. As it has been shown that prediction equations should only be specifically used with the population for which they were developed[120,121,146-148], comparing results between healthy and clinical populations would not be meaningful. As prediction equations are more commonly used with clinical population who may be unable to complete a maximal exercise test, it would be beneficial to review the reliability of equations and of the variables used in populations with cardiovascular disease, or those on cardiac medications, or with other medical conditions.

With regards to the content of the studies included, there is little consistency regarding the specific equations used to estimate the $\dot{V}O_{2max}$ in participants, as well as the variables they utilised. It is hard, therefore, to draw a conclusion about which variables have the greatest effect on predicting $\dot{V}O_{2max}$, especially when those studies that do use the same variables use different statistical analysis methods to examine the variables, or have very different levels of bias in their reporting. Several of the studies scored very low in their risk of bias assessment, and none achieved higher than 15/20 (75%);

therefore reporting bias must be taken into account when looking at which variables are recommended by each study, especially with those examining non-exercise data and how well they predict $\dot{V}O_{2max}$.

3.4.5 Future Recommendations

With so much variety in methods of predicting maximal oxygen consumption currently available, as well as differing opinions and evidence for their accuracy and reliability, there are numerous future directions that can be taken in this area of research. As the Bruce protocol is the most commonly used and known treadmill protocol, it is recommended that future studies look at how exercise data gained from maximal and submaximal treadmill tests following the Bruce protocol relate to measured $\dot{V}O_{2max}$ values. Prediction equations are often used with people unable to complete a maximal test, and therefore a review into the accuracy of prediction equations used with specific clinical populations would also be beneficial. Examining the difference between 'exercise' and 'non-exercise' prediction equations in homogenous populations, where anthropometric data is similar for all participants would help to determine how effective non-exercise equations are in differentiating between anthropometrically similar people.

Any study looking at the accuracy of either a new $\dot{V}O_{2max}$ prediction equation, or one that has been previously established, must ensure that $\dot{V}O_{2max}$ is also measured via analysis of expired air during a maximal exercise test – this is the gold standard of determining cardiorespiratory fitness, and therefore predicted data should always be compared to the measured $\dot{V}O_{2max}$ from the same participants, to accurately determine the equation's reliability.

Care must be taken in any future studies to use the correct statistical analysis when looking at the difference between the measured and predicted $\dot{V}O_{2max}$ values. Using the method described by Bland and Altman[84] to plot the limits of agreement may be the best way to demonstrate the true difference between both values. Authors should also ensure that statistical methods are clearly described, as in the studies reviewed here it has proved difficult at times to determine if the choice and execution of statistical tests were accurate.

3.5 Conclusion

To the author's knowledge, this is the first review to examine $\dot{V}O_{2max}$ prediction equations used in conjunction with a Bruce protocol maximal treadmill test. A wide range of variables can be used to create prediction equations for $\dot{V}O_{2max}$. Exercise test duration is a popular variable, but provides mixed results, whether used alone or in combination with other variables. Non-exercise equations may provide a reliable prediction of $\dot{V}O_{2max}$ for certain populations, but the best combination of non-exercise variables is still unclear. Non-exercise prediction equations are convenient to use, as little effort and time is required to gain a $\dot{V}O_{2max}$ value, but this convenience should not overshadow the fact that with no exercise data, it is more difficult to predict how a person will respond to exertion. As prediction equations are usually used with participants who cannot safely complete a maximal test, it may be beneficial in the future to examine equations using submaximal data versus equations using non-exercise data in predicting $\dot{V}O_{2max}$, as this could give further insight into which types of variables are the most effective $\dot{V}O_{2max}$ predictors.

The only definitive finding regarding these equations, is that the ACSM equation consistently over-predicts $\dot{V}O_{2max}$ and therefore, its use for predicting maximal oxygen consumption should be discontinued.

Chapter 4: A study of the repeatability of the submaximal Bruce protocol graded treadmill test in measuring $\dot{V}O_2$ and predicting maximal $\dot{V}O_2$.

4.1 Introduction

The best and most reliable way to measure a person's cardiorespiratory fitness ($\dot{V}O_{2max}$) is to have them complete a maximal exercise test while measuring and analysing their expired breaths to determine the highest value for $\dot{V}O_2$ that they achieve[2,4,88]. This can, however, pose problems in both clinical and research settings, depending on equipment availability, or the physical condition and capabilities of the participants[41]. Often, $\dot{V}O_{2max}$ values are required for people with cardiac or respiratory conditions, or for people with physical limitations that would prevent them exercising safely to their maximum capacity. In these cases, a submaximal exercise test can be completed, and a prediction of their $\dot{V}O_{2max}$ be calculated, based on their submaximal exercise data[30].

As the Bruce protocol is one of the most commonly used treadmill protocols for determining $\dot{V}O_{2max}$[45,103,149,150], it is the focus of the current research. The repeatability of the Bruce protocol was questioned in the results of the literature review from Chapter 2, due to inconsistent control of certain variables (such as duration between repeated tests, state of fasting or resting before testing, or opportunity for familiarisation with testing equipment), not accounting for learning effect, and for the use of inappropriate statistics to determine repeatability. To examine the Bruce protocol's repeatability more thoroughly, it is important to design a study that comprehensively controls for the many variables that can affect $\dot{V}O_{2max}$, such as timing of sessions, environmental set up, fasting state of participants and levels of physical activity undertaken by participants during the study[151]. The aim of this study was to establish a protocol for examining the repeatability of the submaximal Bruce treadmill protocol in in predicting $\dot{V}O_{2max}$ in healthy male adults. The submaximal protocol was chosen as it poses less risk of cardiovascular complications associated with exercising to exhaustion[41,152], therefore making it a safe option with which to establish a repeatability testing protocol with healthy adult males, and to identify any patterns or challenges that could arise in future repeatability studies of the maximal Bruce protocol.

4.2 Methods

4.2.1 Study design

The study design was an observational, longitudinal study. Ethical approval for this study was granted by the Faculty of Health Sciences Ethics Committee at Trinity College, Dublin in February 2019 (Appendix 3, pg. 106). The design of this study, the data analysis and data interpretation were conducted by the thesis author. The data was collected by final year physiotherapy students.

4.2.2 Participants and Recruitment

A sample size calculation for this study was undertaken, with a result of twenty-three participants required (Appendix 4, pg. 107). A poster and e-mail recruitment campaign was launched from February to April 2019, to request volunteers from the staff and students of Trinity College Dublin. Those expressing interest were sent a participant information leaflet and allowed seven days prior to obtaining informed consent to allow for contemplation of study details. Inclusion criteria specified healthy, English-speaking male adults between 18 and 35 years of age. Volunteers were excluded if they had any cardiac, respiratory, metabolic or neurological condition, had a fitted electronic device (e.g. pacemaker), had any musculoskeletal injury in the previous three months, were a smoker, had a BMI ≥30, or had any intellectual disability/cognitive impairment that impaired their ability to give informed consent, follow instructions during testing, or to exercise adequately (full list of exclusion criteria in Appendix 5, pg. 108). All participants were required to have a low risk of cardiovascular event, as per the American College of Sports Medicine's (ACSM) Guidelines for Exercise Testing and Prescription[153] (Appendix 6, pg. 109).

4.2.3 Pre-test Screening and Assessment

Volunteers were provided with the Participant Information Leaflet (Appendix 7, pg. 112) at least seven days prior to their first testing date. Participants attended the exercise laboratory at the Trinity Centre for Health Sciences three times, with one week between each session. Participants were instructed to fast from midnight the night before their testing day, not to take any caffeine or alcohol in the 24 hours prior to testing, to take motorised transport to the testing centre, to wear light, comfortable exercise clothes to

each testing session, and to continue their normal training regime without increasing their physical activity for the duration of the testing period. The first session began with a screening process reviewing the inclusion and exclusion criteria, as well as explaining the study aims, procedure, risks and benefits. Eligible participants gave informed signed consent at the first testing session (Appendix 8, pg. 119), and then completed the Physical Activity Readiness Questionnaire (PAR-Q; Appendix 9, pg. 123).

Participants' height was measured using a Seca 213 stadiometer, and body composition analysis was done using a body composition analyser (Tanita MC 180-MA), to measure body weight, fat percentage, body mass index (BMI), fat mass, muscle mass and bone mass. Blood pressure was measured in sitting position using an OMROM M3 Comfort electronic blood pressure monitor. To ensure that participants' cholesterol and blood glucose levels were normal, these were measured via finger-prick test using the PRIMA 3-in-1 Self-Testing Kit.

4.2.4 Treadmill $\dot{V}O_2$ testing

The $\dot{V}O_2$ measuring equipment COSMED K4b^2 was calibrated on every testing day following the manufacturer's instructions. Each participant was verbally familiarised with the testing equipment and protocol, including the Rating of Perceived Exertion (RPE) scale from 0-10, and the signs and symptoms requiring test termination:

- Pain, discomfort, or anginal equivalent in the chest, neck, jaw, arms or other areas that may have resulted from ischemia;
- Failure of heart rate to increase with increased exercise intensity;
- Signs of poor perfusion (light-headedness, confusion, ataxia, cyanosis, pallor, nausea or cold/clammy skin);
- Shortness of breath, wheezing, leg cramps or claudication;
- Physical or verbal manifestations of severe fatigue;
- Participant desired to stop;
- Technical difficulties with monitoring heart rate or pulmonary gases, or any other testing equipment failure;
- Participant reached 85% of predicted HR_{max}.

Participants were informed not to speak or cough during test, unless they needed to urgently stop the test. Participants were instructed not to use the handrail on the treadmill (Viasys LE 300CE) at any stage during testing. The K4b^2 mask and strap were fitted, and the Bruce treadmill protocol was conducted (Table 4.1).

Table 4.1 Progression of the Bruce protocol

Stage	Duration (minutes)	Speed ($km \cdot h^{-1}$)	Incline (%)
1	3	2.7	10
2	3	4.0	12
3	3	5.5	14
4	3	6.8	16
5	3	8.1	18
6	3	8.9	20
7	3	9.7	22

RPE was taken at the end of each stage, with the participant pointing at their score on a print-out of the scale held in front of them. Heart rate and $\dot{V}O_2$ were monitored throughout testing. If no adverse events occurred, the test was terminated when the participant reached 85% of their age-predicted maximal heart rate (HR_{max}). The formula by Tanaka et al.[154] was used to calculate HR_{max} (HR_{max} = 208 – (0.7 x age)), as this equation was formulated based on a similar population to the current study, and as it has been recommended as a suitable HR_{max} predictive equation by the ACSM[41]. Each test was led and monitored by the same researcher for the entire study to avoid inter-rater discrepancies.

A five-minute walking cool-down at 0% incline and 2.7$km \cdot h^{-1}$ was completed by participants after reaching 85% of their HR_{max} while HR and $\dot{V}O_2$ were monitored. Participants were scheduled for a second and third appointment, with the best effort made in all cases to schedule the follow up appointments on the same weekday at the same time for each of the three sessions, to eliminate variation that might affect the repeatability test results.

The second and third sessions included measurement of body weight and then a repeat of the Bruce treadmill protocol as described above. Following completion of testing, participants were sent an overview of their health and test results and invited to ask any questions from the researchers.

4.2.5 Predicting $\dot{V}O_{2max}$

The $\dot{V}O_{2max}$ prediction equation was chosen from the prediction equations found in the literature review in Chapter Three. Only three equations were found suitable for the submaximal test protocol: one used by Grant et al.[101], the ACSM walking equation[109],

and the Fitmate equation[31]. The literature review concluded that the ACSM equation over-estimated $\dot{V}O_{2max}$, therefore it was not used for this study. As the prediction of $\dot{V}O_{2max}$ in the study by Grant et al.[101] was done by combining two equations from different sources[107,108] which were not originally designed together, this method was not chosen. The Fitmate equation, based on extrapolating measured $\dot{V}O_2$ to the age-predicted HR_{max} by linear regression, was the most suitable for this study. HR_{max} was predicted using the Tanaka et al.[154] formula.

4.2.6 Statistical Analysis

Mean and standard deviation (SD) values were obtained for all recorded data. Correlation coefficients, coefficient of variation and paired-samples t-tests were calculated to compare predicted $\dot{V}O_{2max}$ results for Test 1 and Test 2, and from Test 2 and Test3. The limits of agreement using Bland and Altman plots were also calculated for the predicted $\dot{V}O_{2max}$ values between each repeated test.

4.3 Results

4.3.1 Participants

Twenty-two people expressed interest in the study, with eighteen participants completing the full testing procedure (Figure 4.1). Baseline participant characteristics can be seen in Table 4.2. For one participant, some values from the body composition analysis were mistakenly not recorded (Appendix 10.1, pg. 124). All data was normally distributed. All included participants were healthy, with no cardiac or other medical conditions. Participants' ages ranged from 19 to 34 years.

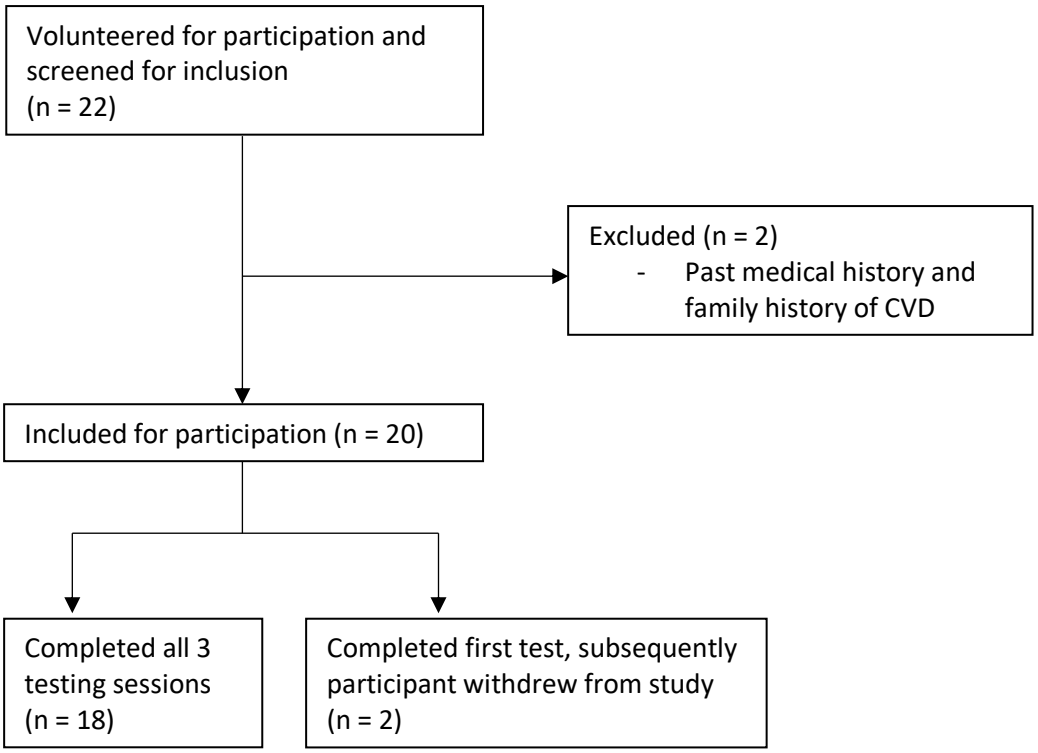


Figure 4.1: Participant inclusion flowchart

Table 4.2 Baseline characteristics of participants (n=18)

Characteristic	Mean (standard deviation)
Age (years)	22.9 (4.4)
Height (cm)	181.3 (6.9)
Weight (kg)	76.29 (10.1)
BMI ($kg \cdot m^{-2}$)	23.3 (2.9)
Basal metabolic rate (kcal)	1938 (193.3)
Fat mass (kg)	12.14 (4.8)
Fat Percent (%)	15.25 (4.2)
Free fat mass (kg)	64.78 (7.2)
Muscle mass (kg)	61.31 (6.8)
Bone mass (kg)	3.23 (0.3)

4.3.2 Submaximal Exercise Tests

Fifty-four tests were completed during the study. All tests were terminated due to participants reaching 85% of predicted HR_{max}, except for two participants in their Test 1 phase: one stopped due to leg cramp, and one stopped due to incorrect readings from the HR monitor. As both participants came near to their 85% of predicted HR_{max}, they were considered to have completed one full test and so the data for their second and

third tests were included in analysis as learning effect would not be compromised. Therefore, in analysis of the data, Test 1 results consist of 16 participants, and Test 2 and Test 3 results both consist of 18 test results.

The duration between the majority of tests was one week (between Test 1 and Test 2: 72.2%, n=13; between Test 2 and Test 3: 83.3%, n=15); however, the duration was shorter or longer between tests for certain participants, due to participants' personal scheduling difficulties (between Tests 1 and 2: duration longer than one week: n=5; between Tests 2 and 3: duration longer than one week: n=1, duration shorter than 1 week: n=2. Shortest duration = 3 days, longest duration = 14 days).

Heyward and Gibson[155] provided normative $\dot{V}O_{2max}$ values divided into age and gender categories, with $\dot{V}O_{2max}$ values ranked as "superior", "excellent", "good", "fair" or "poor". The predicted $\dot{V}O_{2max}$ values for each test completed in this study were compared to this normative $\dot{V}O_{2max}$ data. It was found that 73.1% of all $\dot{V}O_{2max}$ test results ranked as "superior" and 19.2% were "excellent", for the relative age group, while 5.8% were classed as "good" results, and only 1.9% of tests had a "fair" or "poor" result.

4.3.3 Repeatability of Bruce protocol

To analyse the repeatability of the submaximal Bruce protocol in predicting $\dot{V}O_{2max}$ values, correlation coefficients, coefficients of variation, paired-samples t-tests (displayed in Table 4.3), and limits of agreement (LOA; Figures 4.2 and 4.3) were calculated for the predicted $\dot{V}O_{2max}$ values to compare Test 1 and Test 2 results, and Test 2 and Test 3 results.

Table 4.3 Statistical analysis between tests for $\dot{V}O_{2max}$ predicted

	Test 1 – Test 2	**Test 2 – Test 3**
Correlation Coefficient	r = 0.82	r = 0.78
Mean Coefficient of Variation	Mean: 4.5% (SD: ±4.2%; Range: 0.5-15.6%)	Mean: 3.6% (SD: ±2.5%; Range: 0.9-9.3%)
Paired-samples t-test	p = 0.013	p = 0.473

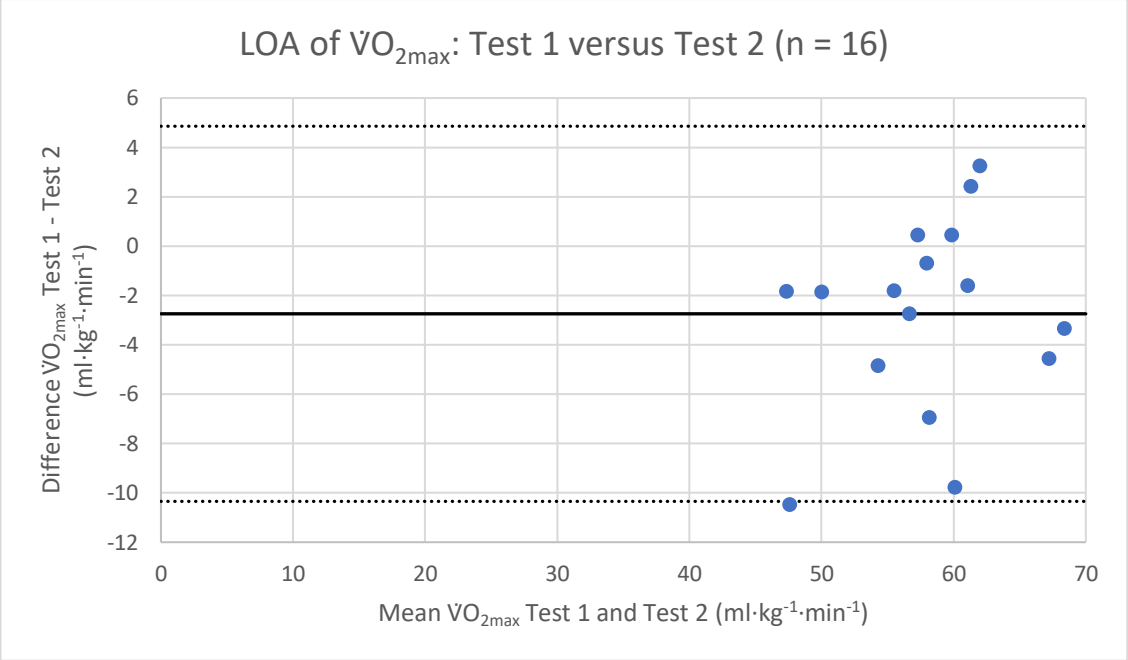


Figure 4.2 Limits of agreement (LOA) of predicted $\dot{V}O_{2max}$ between Test 1 and Test 2

The limits of agreement in $\dot{V}O_{2max}$ between the first and second tests ranged from 4.86 to -10.35 ml·kg^{-1}·min^{-1}, spanning a range of 15.21 ml·kg^{-1}·min^{-1} (Figure 4.2). The mean difference between Test 1 and Test 2 measurements was -2.74 ml·kg^{-1}·min^{-1}. The lower and upper 95% confidence intervals (95% CI) for the mean difference were -6.02 ml·kg^{-1}·min^{-1} and 0.17 ml·kg^{-1}·min^{-1}, respectively (ranging 6.19 ml·kg^{-1}·min^{-1}).

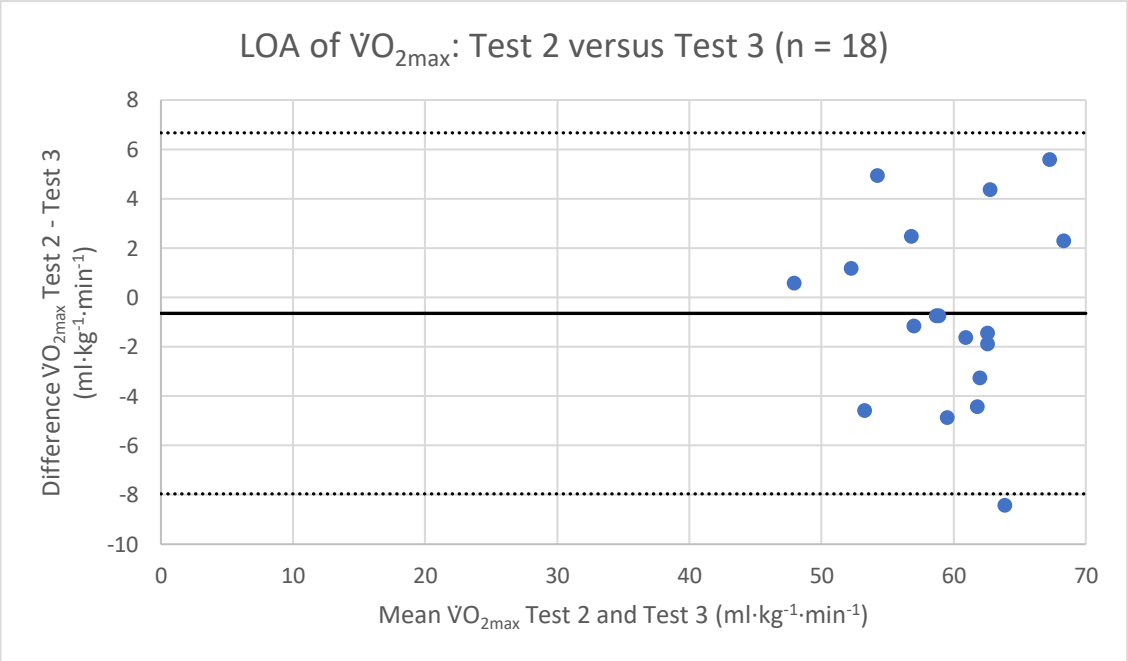


Figure 4.3 Limits of agreement (LOA) of predicted $\dot{V}O_{2max}$ between Test 2 and Test 3

Comparing $\dot{V}O_{2max}$ results from the second and third tests, the limits of agreement range from 6.67 to -7.96 ml·kg^{-1}·min^{-1} (Figure 4.3), giving a difference range of 14.63 ml·kg^{-1}·min^{-1}. The mean difference between results from Test 2 and Test 3 was -0.65 ml·kg^{-1}·min^{-1}. The lower and upper 95% CI for the mean difference were -3.87 ml·kg^{-1}·min^{-1} and 1.94 ml·kg^{-1}·min^{-1}, respectively (ranging 5.81 ml·kg^{-1}·min^{-1}).

4.3.4 Repeatability of submaximal test duration

Each test was stopped when the participant reached 85% of their age-predicted maximal heart rate. The repeatability of the submaximal test durations was also analysed with correlation coefficients, coefficients of variation and paired-samples t-tests, comparing Test and Test 2, and Test 2 and Test 3 in turn (Table 4.4). Limits of agreement were also analysed for the submaximal test durations.

Table 4.4 Statistical analysis between repeated tests for submaximal test duration

	Test 1 – Test 2	**Test 2 – Test 3**
Correlation Coefficient	r = 0.85	r = 0.79
Mean Coefficient of Variation	Mean: 4.2% (SD: ±5.0%; Range: 0.5-19.6%)	Mean: 4.8% (SD: ±5.0%; Range: 0.6-19.0%)
Paired-samples t-test	*p* = 0.671	*p* = 0.819

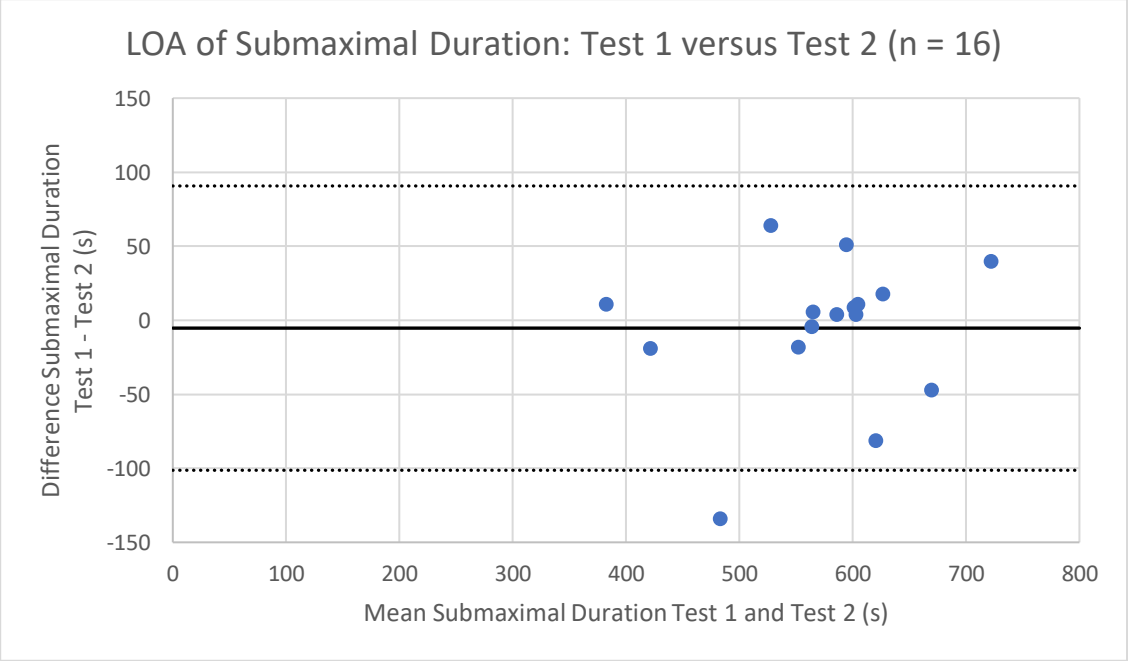


Figure 4.4 Limits of agreement (LOA) of submaximal test duration between Test 1 and Test 2

The limits of agreement for submaximal duration comparing Tests 1 and 2 ranged from -101.32s to 90.69s, which spans a total of 192.01s (Figure 4.4). The mean difference was -5.31 seconds. The lower and upper 95% CI for the mean difference were -31.41s and 20.78s, respectively (ranging 52.19s).

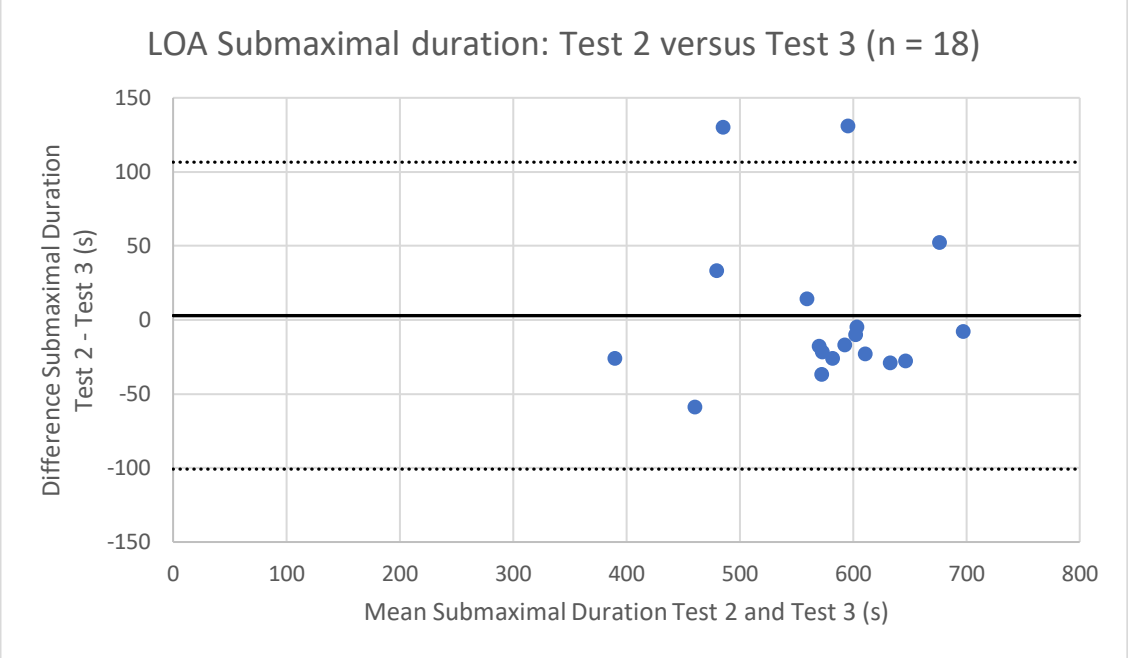


Figure 4.5 Limits of agreement (LOA) of submaximal test duration between Test 2 and Test 3

Comparing the submaximal test durations from Tests 2 and 3, the limits of agreement were found to be from -100.76s to 106.54s, which spans a range of 207.3s (Figure 4.5). The mean difference between Test 2 and Test 3 test durations was 2.89 seconds. The lower and upper 95% CI for the mean difference were -23.41s and 29.19s, respectively (ranging 52.60s).

All raw data for this submaximal Bruce protocol repeatability study can be found in Appendix 10 (pg. 124).

4.4 Discussion

4.4.1 Repeatability of the Submaximal Bruce protocol in predicting $\dot{V}O_{2max}$

Looking firstly at the analysis by correlation coefficients, this study shows similar results to those found in previous Bruce protocol repeatability studies (from the literature review in Chapter 2): a strong correlation between repeated $\dot{V}O_{2max}$ results from Bruce protocol testing (Test 1 versus Test 2 r = 0.82; Test 2 versus Test 3 r = 0.78). This indicates

that repeated $\dot{V}O_{2max}$ results are strongly related to each other; however this does not translate into agreement, or repeatability[84,138]. The mean coefficient of variation for both comparisons (Test 1 v Test 2, and Test 2 v Test 3) were less than 10%, which is generally considered to be an acceptable level of variation for $\dot{V}O_{2max}$ measurement[30,142]. As all repeated results came from the same participants having undergone the same testing procedure, it is expected that the results would be related to each other. The limits of agreement give a more in-depth analysis of the differences between test results, compared to correlation coefficients[84,156,157]. It can be seen from the current test results that between Test 1 and Test 2, 95% of differences between repeated $\dot{V}O_{2max}$ measurements could range as much as 15.21 $ml \cdot kg^{-1} \cdot min^{-1}$. Participants were tested under the same conditions from the first to the second test, with only a week between tests and having had no change in usual physical activity levels that could account for a change in $\dot{V}O_{2max}$ levels. The limits of agreement were similar when comparing Test 2 and Test 3 results, with a total range in limits of agreement of 14.63 $ml \cdot kg^{-1} \cdot min^{-1}$. Considering that the minimal clinically important difference (MCID) for $\dot{V}O_{2max}$ for healthy adults is generally taken as 3.5 $ml \cdot kg^{-1} \cdot min^{-1}$ [158,159], the variation in the current study could be attributed to poor repeatability of the Bruce protocol itself. This wide possible range in $\dot{V}O_{2max}$ from one test to a repetition a week later cannot be accounted for by MCID alone, and calls into question the repeatability of the Bruce protocol.

4.4.2 Learning Effect

In a number of previous repeatability studies of the Bruce protocol, a learning effect was noted and examined[54,55]. The effect was mainly noted in treadmill test duration, with participants lasting for longer durations to their maximum effort on their repeated tests compared to their initial test. Although the tests in the current study were submaximal, these submaximal durations were examined, and it was noted that 62.5% of submaximal durations decreased from Test 1 to Test 2 – a result in contrast with previous studies. The mean difference was -5.31 seconds, while the correlation coefficient for test durations of the first and second tests was 0.85, the mean coefficient of variation was 4.2% and there was no significant difference between the two test durations ($p = 0.671$). Comparing the second and third test submaximal durations, it was

found that 72.2% of participants had an increased test duration in their third compared to their second submaximal test, which more strongly suggests a learning effect in this stage of the study. With increased experience of the testing procedure and understanding what to expect from the protocol, participants may have become more physiologically economical and improved their test performance. This psychological influence of exercise test expectations has been previously identified and study by other authors, also[81,82].

Regarding $\dot{V}O_{2max}$ predicted values using the Fitmate equation, there was generally an increase in values from Test 1 to Test 2. While there was strong correlation between repeated $\dot{V}O_{2max}$ values across all three test repetitions, the mean coefficient of variation was larger between the first and second repetitions, compared to the second and third repetitions. A significant difference was also noted between repeated results from Test 1 to Test 2 when analysed with paired-samples t-tests (p = 0.013) but not on the repetition from Test 2 to Test 3 (p = 0.473). Finally, the mean difference identified through the limits of agreement plots was much larger between Tests 1 and 2 (-2.74 $ml \cdot kg^{-1} \cdot min^{-1}$), compared to between Tests 2 and 3 (-0.65 $ml \cdot kg^{-1} \cdot min^{-1}$). These differences indicate that learning effect played an important role in participants' performance in these submaximal tests, and on their $\dot{V}O_{2max}$ predicted values. Although the limits of agreement showed a wide range in differences between all repeated $\dot{V}O_{2max}$ predictions, the repeatability is somewhat better from second to third attempts, rather than after a person's first experience with the protocol. This can be explained through a learning effect. As this study was submaximal, further research into repeatability and learning effect in this population using the maximal Bruce protocol would be of great benefit.

4.4.3 Limitations

Although measures were taken to control for environmental variables throughout the study, such as testing centre, equipment and research personnel, in hindsight several other actions could have been taken to further ensure that variation in results would reflect repeatability of the Bruce protocol itself, and not be due to other variables. Motivational variables were not considered and may have influenced the obtained data.

Verbal encouragement was not standardised during the exercise testing, and it has been shown that the style and frequency of verbal encouragement can influence a person's exercise test results[117,160,161]. Participants could also see their progress regarding stage and duration during the test on the treadmill. This could have influenced their determination to compete with their previous result in their second or third tests, and have had an impact on learning effect and reliability of comparing the repeated test results.

Another limitation is that the participant sample was relatively homogenous, coming from a population of healthy male university students and staff. The age span ranged from 18 to 34 years, with the majority of participants (77.8%) aged 20 to 29 years. It was noted that the mean $\dot{V}O_{2max}$ predicted across the three tests was in the excellent-superior categories, as per Heyward and Gibson[155] (92.3% of all test results). No data was gathered regarding normal levels of physical activity in the participants, such as type of physical activity undertaken, how often, or at what level. Therefore, it is difficult to determine if the participants themselves were physically very fit, or if the prediction equation was likely to be overestimating the $\dot{V}O_{2max}$ of these participants. Future studies would benefit from gathering more detailed data regarding physical activity levels, as well as gaining measured $\dot{V}O_{2max}$ data, to compare this to the predicted values.

A further limitation of this study is that the sample size was smaller than that required as per the sample size calculation. This may have impacted the statistical results of the study, as having a small sample size reduces the power of the study[62-64]. Using limits of agreement, a wide confidence interval for the mean difference may reflect a small sample size – when a larger sample size is used, the 95% confidence interval should become narrower if there is no systematic error and the measure is in fact repeatable[84, 137]. In this study, the 95% CI for the mean difference in $\dot{V}O_{2max}$ testing between Test 1 and Test 2 ranges 6.19 $ml \cdot kg^{-1} \cdot min^{-1}$, wider than the MCID for $\dot{V}O_{2max}$ testing (3.5 $ml \cdot kg^{-1} \cdot min^{-1}$). Similarly, it ranges 5.81 $ml \cdot kg^{-1} \cdot min^{-1}$ or the analysis between Test 2 and Test 3 results. These wide ranges could be caused by the small sample size, or could be demonstrating that the submaximal Bruce protocol is not repeatable in measuring $\dot{V}O_{2max}$. The results from this study should be taken to show that a trend towards the submaximal Bruce protocol not being a repeatable method of determining $\dot{V}O_{2max}$, but

that to confirm these results future studies should take into consideration all limitations discussed above and conduct further repeatability studies with a larger sample size to improve the power and validity of these findings.

4.5 Conclusions and Future Recommendations

Assessing the submaximal test results from this study using a variety of different statistical methods, it can be seen that predicting $\dot{V}O_{2max}$ in healthy male participants using the Bruce treadmill protocol is not definitively repeatable. The wide range in limits of agreement between repeated $\dot{V}O_{2max}$ results pulls the repeatability of the protocol into question. A learning effect was noted to influence the repeated test duration and predicted $\dot{V}O_{2max}$ results at this submaximal level. Future studies in this area should have a larger and more heterogenous group of participants, and should control more specifically for variables other than the protocol itself which could influence $\dot{V}O_{2max}$ results (such as verbal encouragement and blinding of participants to their results prior to the end of testing). Maximal testing using the Bruce protocol should be carried out, with focus on the influence of learning effect between several test repetitions, to determine if this effect is as strong with maximal testing as it appears to be with submaximal treadmill testing. The use of the maximal Bruce protocol to measure $\dot{V}O_{2max}$ should also be conducted to better determine the accuracy of prediction equations used to calculate $\dot{V}O_{2max}$, by comparing predicted with measured values.

Chapter 5: A study of the repeatability of the Bruce protocol graded treadmill test in measuring maximal $\dot{V}O_2$, and in predicting $\dot{V}O_{2max}$ from exercise data.

5.1 Introduction

As previously discussed, maximal aerobic capacity ($\dot{V}O_{2max}$) is an important factor in determining a person's cardiovascular risk[2]. The most accurate and reliable method for obtaining a $\dot{V}O_{2max}$ value for an individual is to have them complete a maximal exercise test while measuring and analysing their breath-by-breath expired gases and determining their highest $\dot{V}O_{2max}$ value[101,103,162]. The most commonly used protocol for testing $\dot{V}O_{2max}$ is the Bruce protocol[15,16] and it is therefore the focus of the current research. The repeatability of any measurement tool is highly important, to ensure results are reliable[163,164] and as shown in the literature review in Chapter 2, the previous repeatability studies examining the Bruce protocol have been flawed with regards to the statistical analysis used, standardised instructions, and controlling for learning effect of the Bruce protocol.

In many scenarios, completing a maximal exercise test while measuring expired air is not possible, whether that be due to a person's physical limitations such as pain or fatigue, or to a lack of specific testing equipment[30,33]. In these situations, prediction equations are used to calculate the person's predicted $\dot{V}O_{2max}$, based on certain variables such as age, weight, submaximal heart rates or exercise test durations[1,41,100]. However, predictions equations may over- or under-predict a person's $\dot{V}O_{2max}$[34,35], and choosing the correct prediction equation and method of testing is highly important[41].

There is also poor consensus regarding the type of data sampling to be used to determine the $\dot{V}O_{2max}$ from the collected test data[165,166]. Measurement of $\dot{V}O_2$ during exercise via indirect calorimetry provides a large volume of data. A number of methods for averaging this data are available to researchers, such as time averaging the last 30 seconds of data[167], or averaging the highest available 30 seconds of $\dot{V}O_2$ data[44], or by obtaining a rolling average of breath-by-breath $\dot{V}O_2$ data for any number of breaths and choosing the highest average[168-170]. Due to these varying options, no set conclusion is

available as to which provides the most accurate $\dot{V}O_{2max}$ value, and therefore ensuring a repeatable measurement of $\dot{V}O_{2max}$ from one test to the next is difficult.

The first aim of the current study was to examine the repeatability of the Bruce protocol in measuring $\dot{V}O_{2max}$ for male university students and staff members, by following a strict protocol to repeat the test with the participants on three occasions, and comparing results using a variety of statistical analysis methods. The second aim of the study was to examine the results from various prediction equations in calculating $\dot{V}O_{2max}$ and comparing these results to the known $\dot{V}O_{2max}$ values for the participants. The varying methods of data averaging to determine a $\dot{V}O_{2max}$ were also examined for each of the participant's collected data, to identify and compare any differences between their results.

5.2 Methods

5.2.1 Recruitment and Screening

Ethical approval was obtained in June 2019 for this study from the Faculty of Health Sciences Ethics Committee, at Trinity College Dublin (Appendix 11, pg. 127). A sample size calculation for this study was undertaken, with a result of twenty-three participants required (Appendix 4, pg. 107). An email and poster campaign to recruit healthy male volunteers from the staff and students of Trinity College Dublin was launched, and ran from October 2019 to March 2020, with the recruitment email being sent to all staff and students under the School of Medicine, as well as posters being displayed around the university campus. Volunteers expressed interest via email to the lead investigator and were screened against the list of inclusion and exclusion criteria. Volunteers were eligible to participate if they were healthy non-smoking males between ages 18 to 45 years, and currently a staff or student member of Trinity College Dublin. Volunteers were excluded if they were female, non-fluent in English, a current smoker or had quit within the past six months, had any physical or cognitive disability affecting their ability to give informed consent or safely exercise on a treadmill, or had any medical condition or past medical history including cardiac, pulmonary and metabolic conditions (full list of conditions in Appendix 5, pg. 108). All participants were required to have a low risk

of cardiovascular event, as per the American College of Sports Medicine (ACSM) Guidelines for Exercise Testing and Prescription[153] (Appendix 6, pg. 109).

5.2.2 Pre-test Instructions and Baseline Assessments

Included volunteers were given the Participant Information Leaflet (Appendix 12, pg. 128) at least seven days before their first testing session. They were informed not to complete any vigorous activity in the 24 hours preceding the testing session, and to fast from midnight the night before the testing session. They were also requested to take public transport or drive to the test centre, to avoid being fatigued due to an active commute. At the start of the first session, the main aims and processes of the study were explained, as well as the risks and benefits of participating in the study and details relating to data processing and data protection. Agreeable volunteers gave their informed, signed consent before the testing session began (Appendix 13, pg. 135).

Each participant completed the Physical Activity Readiness Questionnaire (PAR-Q[171]; Appendix 9, pg. 123) as a screen for safety to exercise, as well as the International Physical Activity Questionnaire – Short Form (IPAQ-SF[172], Appendix 14, pg. 138) to obtain a subjective measure of their usual physical activity levels. Following this, the participant's height was measured using a Seca 213 stadiometer, and they completed a bioimpedance analysis on a body composition analyser (Tanita MC 180-MA), which included body weight measurement. As in the previous study, to ensure participants' blood glucose and blood cholesterol were normal, these were measured via a finger-prick blood test using the PRIMA 3-in-1 Self-Testing Kit. Resting blood pressure was tested using an OMROM M3 Comfort electronic blood pressure monitor, after a five-minute seated rest. Prior to each test, the COSMED Quark CPET (Cardiopulmonary Exercise Testing) was calibrated according to the manufacturer's specifications. The COSMED heart rate monitor was attached at the apex beat on the chest of the participant, and resting heart rate was measured. The final explanation of the test, including the termination criteria and a reminder of the risks, as well as instructions for measuring the rating of perceived exertion (RPE) and lactate measurement, was given. The face mask was secured, as well as the safety harness on the treadmill (treadmill make: Viasys LE 300CE). The treadmill display and the COSMED Quark data

output were hidden from participants' view to blind them throughout testing to their test duration, HR_{max} and $\dot{V}O_{2max}$ achieved.

5.2.3 Exercise Testing

The standard Bruce protocol was followed for the $\dot{V}O_{2max}$ testing (Table 5.1).

Table 5.1 Standard Bruce protocol

Stage	Duration	Speed ($km \cdot h^{-1}$)	Incline (%)
1	3 minutes	2.7	10
2	3 minutes	4.0	12
3	3 minutes	5.5	14
4	3 minutes	6.8	16
5	3 minutes	8.1	18
6	3 minutes	8.9	20
7	3 minutes	9.7	22

RPE was recorded in the last thirty seconds of each stage, and the finger-prick lactate test was measured at the start of each stage, from Stage 4 onwards, using the Lactate Plus Lactate Meter (Nova Biomedical). Heart rate and $\dot{V}O_2$ were monitored throughout the test. Standardised verbal encouragement (written following research by Andreacci et al.[117] and Midgley et al.[173]) was given every twenty seconds from the 9th minute onwards during each test (Appendix 15, pg. 140). Termination criteria for the maximal test were as follows:

- Pain, discomfort, or anginal equivalent in the chest, neck, jaw, arms or other areas that may have resulted from ischemia;
- Failure of heart rate to increase with increased exercise intensity;
- Signs of poor perfusion (light-headedness, confusion, ataxia, cyanosis, pallor, nausea or cold/clammy skin);
- Shortness of breath, wheezing, leg cramps or claudication;
- Physical or verbal manifestations of severe fatigue;
- Participant desired to stop;
- Technical difficulties with monitoring heart rate or pulmonary gases, or any other testing equipment failure;
- Participant completed all seven stages of the protocol.

Once the test was terminated, the participant completed a five-minute walking cool-down at $2.7 km \cdot h^{-1}$ before removing the harness, facemask and heart rate monitor, and having a seated rest. Post-test blood pressure and heart rate were taken, and the participant was asked to define their main reason for stopping the test.

Participants were requested to return for two more sessions on the same day of the week, each a week apart, to repeat the maximal treadmill testing. Participants were asked to keep their physical activity levels unchanged for the duration of their involvement in the study. The second and third testing sessions consisted only of body weight, resting blood pressure and resting heart rate measurements, followed by the maximal treadmill testing as described above for the first session. As for the first test, participants were requested to fast from midnight the night before their test, not to complete any vigorous physical activity in the 24 hours prior to testing, and to take public transport or drive to the test centre, for their second and third testing sessions.

5.2.4 Physical Activity Monitoring

Participants were given an ActiGraph activity monitor after their first testing session, to wear for the following seven days. This was to determine their usual average weekly physical activity levels. They were given verbal instructions and provided with an instruction booklet and an activity diary (Appendix 16, pg. 141), and were requested to return the activity monitor at their next testing session.

At the end of the third and final session, the participants were given the results from their tests and activity monitoring in the form of a personalised health report (Appendix 17, pg. 145) and were given the opportunity to ask any questions regarding the testing and their results.

5.2.5 Time and Breath-by-Breath Averaging to Calculate $\dot{V}O_{2max}$

The $\dot{V}O_2$ data output from each test completed was analysed in Excel following several different methods of time and breath-by-breath sampling, to calculate the $\dot{V}O_{2max}$ value, as described by previous authors:

- the last 30-seconds of $\dot{V}O_2$ data at maximal exertion (LAST30S)[167,170];
- the highest 30-second rolling average during the test (HIGH30S)[44,174];
- the highest 15-second block average (BLOCK15S)[169,175] and highest 30-second block average (BLOCK30S)[169,170] during the test;
- the highest 15-breath rolling average (ROLL15BR)[168,170] during the test.

5.2.6 Prediction equations

Six prediction equations were used to calculate $\dot{V}O_{2max}$ for each test by each participant. Table 5.2 outlines each equation. "Bruce 1" and "Bruce 2" equations were from the same study[1]: one quoted in the text of the study by Bruce et al.[1] ("Bruce 1"); the other presented as an equation of the line and cited by Crouse et al.[105] when examining prediction equations ("Bruce 2"). Two methods of interpreting the ACSM equations were found: one from the ACSM's Guidelines for Exercise Testing and Prescription book[41] and the other described by Lee et al.[31] when examining prediction equations. The final equation was the Fitmate equation as described by Lee et al.[31] and was included as it was the equation used in the previous submaximal study (Chapter 4). To calculate age-predicted maximal heart rate (HR_{max}), the Tanaka et al.[154] equation was used.

Table 5.2 Prediction Equations

Equation Name	Equation	Explanation
"Bruce 1"[1]	$\dot{V}O_{2max}$ = 6.70 - (2.82 x sex weighting) + (0.056 x D)	- Regression equation for healthy participants - Sex weighting: men = 1; 0 = women - D = Duration of maximal treadmill test, in seconds
"Bruce 2"[1]	$\dot{V}O_{2max}$ = (3.778 x D) + 0.19	- Regression equation for active men - D = Duration of maximal treadmill test, in minutes
"Foster"[113]	$\dot{V}O_{2max}$ = 14.8 - (1.379 x D) + (0.451 x D^2) - (0.012 x D^3)	- Regression equation for any population - D = Duration of maximal treadmill test, in minutes
ACSM running equation: "ACSM2018"[41]	$\dot{V}O_2$ = 3.5 + (0.2 x speed) + (0.9 x speed x grade)	- Conduct treadmill test to 85% of age-predicted HR_{max} [*Tanaka et al.[154] equation: 208 – (0.7 x age)]* - Extrapolate HR values from exercise test to age-predicted HR_{max}, against test time, to determine which stage of Bruce protocol would have been achieved at maximal exertion. - Use equation to calculate $\dot{V}O_{2max}$ at that stage: speed of that stage in $m \cdot min^{-1}$; grade that stage as treadmill grade percentage, in decimal form.
ACSM walking equation "ACSMLee"[31]	$\dot{V}O_2$ = 3.5 + (0.1 x speed) + (1.8 x speed x grade)	- Calculate submaximal $\dot{V}O_2$ for Stages 1, 2, and 3 of Bruce protocol. - Conduct treadmill test to 85% of age predicted HR_{max} [154]. - Average HR for last 30s of each stage 1, 2, 3. If participant did not complete stage 3, average HR for last 30s of test, instead of Stage 3. - Determine equation of the line of best fit through submaximal $\dot{V}O_2$ and HR for the three points. - Extrapolate line to age-predicted HR_{max}, and calculate predicted $\dot{V}O_{2max}$.
"Fitmate"[31]	Linear equation of the line, plotting $\dot{V}O_2$ against heart rate *(equation different for each individual)*	- Exercise to 85% age-predicted HR_{max}. - Plot the measured $\dot{V}O_2$ against measured HR up to the 85% age-predicted HR_{max} end-point. - Get the linear equation of the line. - Extrapolate $\dot{V}O_2$ to the age-predicted HR_{max}, to get $\dot{V}O_{2max}$. - HR_{max} predicted with Tanaka et al.[154] equation.

5.2.7 Statistical Analysis

Paired-samples t-tests were used to compare the results from different sampling methods for $\dot{V}O_{2max}$ for each test. Limits of agreement were used to compare the difference in $\dot{V}O_{2max}$ from the first and second tests, as well as the second and third tests, and between different methods of sampling. Limits of agreement were also used to compare test durations. Correlation coefficients, the coefficient of variation, and paired-samples t-tests for significance were calculated for $\dot{V}O_{2max}$ and duration values between repeated tests. Paired-samples t-tests, correlation coefficients and coefficients of variation were used to compare the prediction equation results to each other, and to the $\dot{V}O_{2max}$ measured with maximal testing. Significance was set at $p < 0.05$ for paired-samples t-tests.

5.3 Results

5.3.1 Participants

Forty-three people expressed interest in participating, via e-mail. At the completion of testing, fifteen participants had completed all three testing sessions, and a further four participants had completed their first and second tests, but were unable to complete their final test session due to the closure of the university during the Covid-19 pandemic. Their results are included in the analysis of comparison between Test 1 and Test 2. Figure 5.1 shows the full participant inclusion flow-chart.

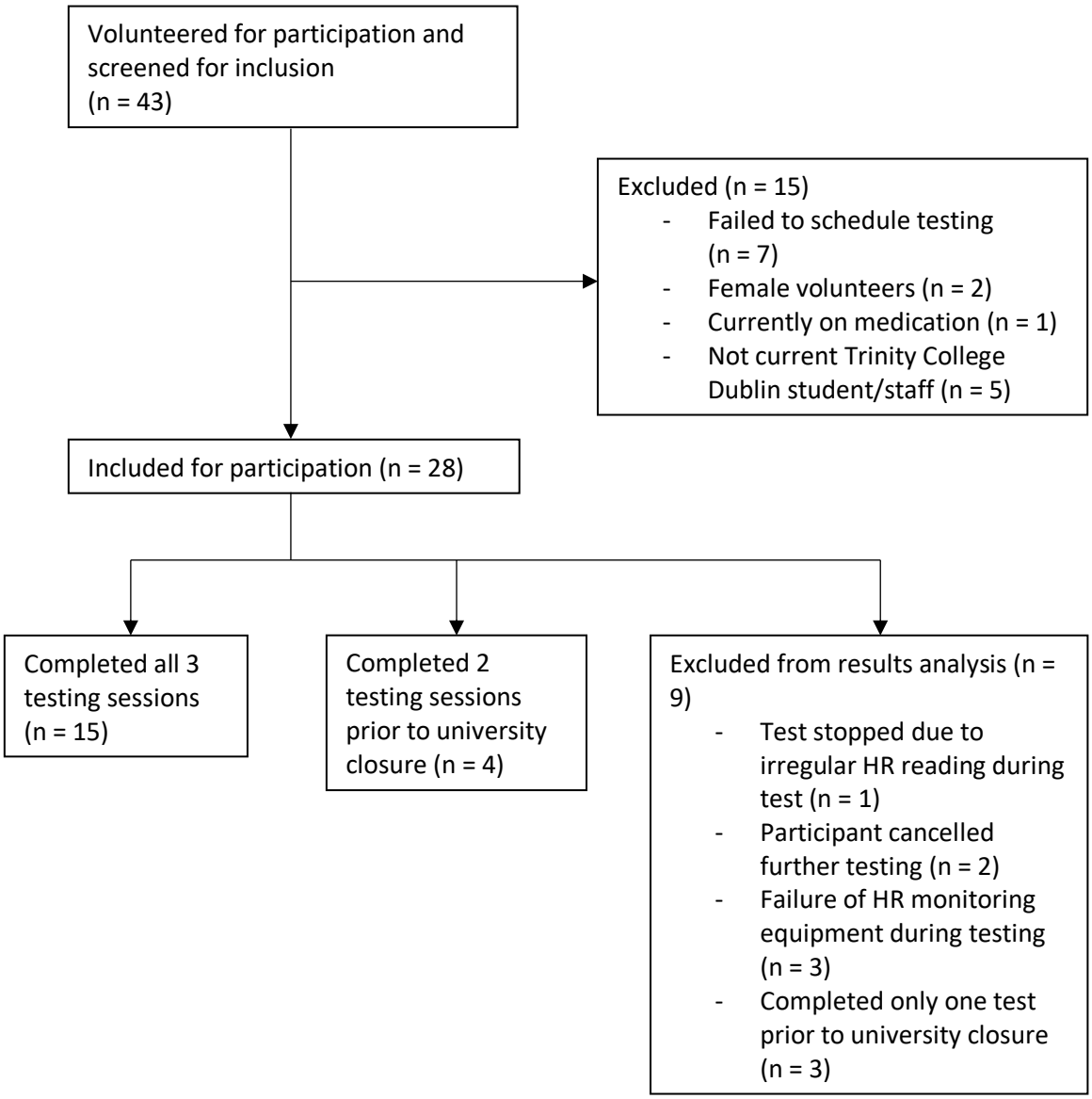


Figure 5.1 Participant inclusion flowchart

Baseline characteristics of the participants can be seen in Table 5.3. All data was normally distributed. All included participants were healthy, with no cardiac or other medical conditions. Participants' ages ranged from 18 to 41 years, with the majority of participants under 25 years of age (63.2%). According to self-reported physical activity levels, 79.0% of participants reported achieving greater than 150 minutes of moderate to vigorous physical activity (MVPA) per week. ActiGraph data showed that all participants achieved greater than 150 minutes of MVPA during the week in which they wore the activity monitor. When analysing bouts of physical activity lasting longer than

ten minutes, however, only 55.6% of the participants achieved the recommended guidelines of greater than or equal to 150 minutes of moderate-vigorous physical activity.

Table 5.3 Baseline characteristics of participants (n=19)

Characteristic	Mean (standard deviation)
Age (years)	26.3 (7.7)
Height (cm)	179.9 (5.7)
Weight (kg)	76.14 (9.4)
BMI ($kg \cdot m^{-2}$)	23.4 (2.1)
Basal metabolic rate (kcal)	1886 (209.6)
Fat mass (kg)	11.95 (4.6)
Fat Percent (%)	15.43 (4.8)
Free fat mass (kg)	63.98 (6.7)
Muscle mass (kg)	60.79 (6.4)
Bone mass (kg)	3.18 (0.3)

5.3.2 Maximal Exercise Tests

Fifty-three maximal treadmill tests were completed in total. All participants achieved maximal exercise test criteria for the tests completed. Common reasons given for stopping the test were leg fatigue or cramping (26/53, or 49.1%), breathlessness (16/53, or 30.2%), or for both breathlessness and leg fatigue (7/53, or 13.2%). Two tests were stopped by the researcher, as the participant plateaued at heart rate max; one participant reported stopping due to “taste of blood in his mouth” and one participant did not give a reason for stopping one test. The duration between the majority of tests was one week (79.4%), however some participants had shorter (5.9%) or longer (14.7%) durations between their testing sessions, due to personal scheduling difficulties. The longest duration between a subsequent test was 23 days, and the shortest duration was 6 days.

5.3.3 Analysis of $\dot{V}O_2$ data

Five methods of time and breath-by-breath sampling were used to determine $\dot{V}O_{2max}$ from the indirect calorimetry data output of $\dot{V}O_2$. Paired-samples t-tests comparing different methods showed that there were significant differences between all methods of data averaging ($p < 0.05$), excluding between BLOCK15S and ROLL15BR for both Tests

1 (p = 0.414) and 2 (p = 0.084), or between HIGH30S and BLOCK15S for both Tests 2 (p = 0.139) and 3 (p = 0.670). The full results are in Appendix 18.7 (pg. 158). Absolute differences between each data sampling method for each participant were also calculated (Appendix 18.8, pg. 159) and it was found that across the three tests, the mean value of $\dot{V}O_{2max}$ differences between methods was small (Test 1 = 0.60 $ml \cdot kg^{-1} \cdot min^{-1}$; Test 2 = 0.61 $ml \cdot kg^{-1} \cdot min^{-1}$; Test 3 = 0.72 $ml \cdot kg^{-1} \cdot min^{-1}$).

5.3.4 $\dot{V}O_{2max}$ values compared to Normative Data

Looking at results from the highest rolling 30-second average $\dot{V}O_2$ data, it was found that 41.5% of results were categorised as "superior" when compared to normative age-based data[155]. Following this, 18.9% were "excellent" results, 30.2% were "good", and 9.4% were classed as "fair" or "poor" results.

5.3.5 Repeatability of Bruce protocol in measuring $\dot{V}O_{2max}$ using highest 30-second average

To analyse the repeatability of the Bruce protocol, the highest rolling 30-second average (HIGH30S) sampling method of $\dot{V}O_2$ was chosen as it is commonly used, and recommended by the American Heart Association[44]. The LAST30S sampling method was also analysed for repeatability, as taking data from the maximal effort point, or last 30-60 seconds of an exercise test, was a common method among the studies examined in the Chapter 2 literature review[54,57,59,61].

For HIGH30S repeatability analysis, the limits of agreement (LOA), as described by Bland and Altman[84] were used to compare the difference in $\dot{V}O_{2max}$ between the participants' first and second tests, as well as their second and third tests (Figures 5.2 and 5.3).

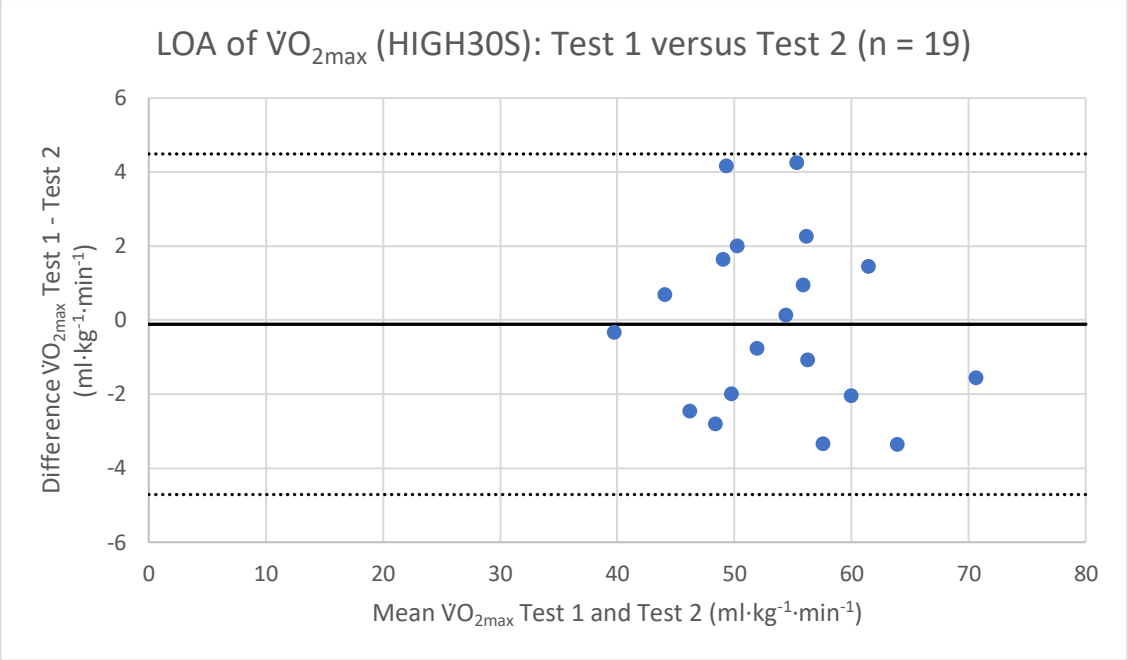


Figure 5.2 Limits of agreement (LOA) of $\dot{V}O_{2max}$ (HIGH30S) between Test 1 and Test 2

The limits of agreement in $\dot{V}O_{2max}$ between the first and second tests ranged from 4.49 to -4.71 ml·kg^{-1}·min^{-1} (Figure 5.2). This gave a difference range of 9.2 ml·kg^{-1}·min^{-1} between the first and second time that participants completed the maximal Bruce protocol. The mean difference between Test 1 and Test 2 measurements was -0.11 ml·kg^{-1}·min^{-1}. The lower and upper 95% confidence intervals (95% CI) for the mean difference were -1.42 ml·kg^{-1}·min^{-1} and 1.02 ml·kg^{-1}·min^{-1}, respectively (ranging 2.26 ml·kg^{-1}·min^{-1}).

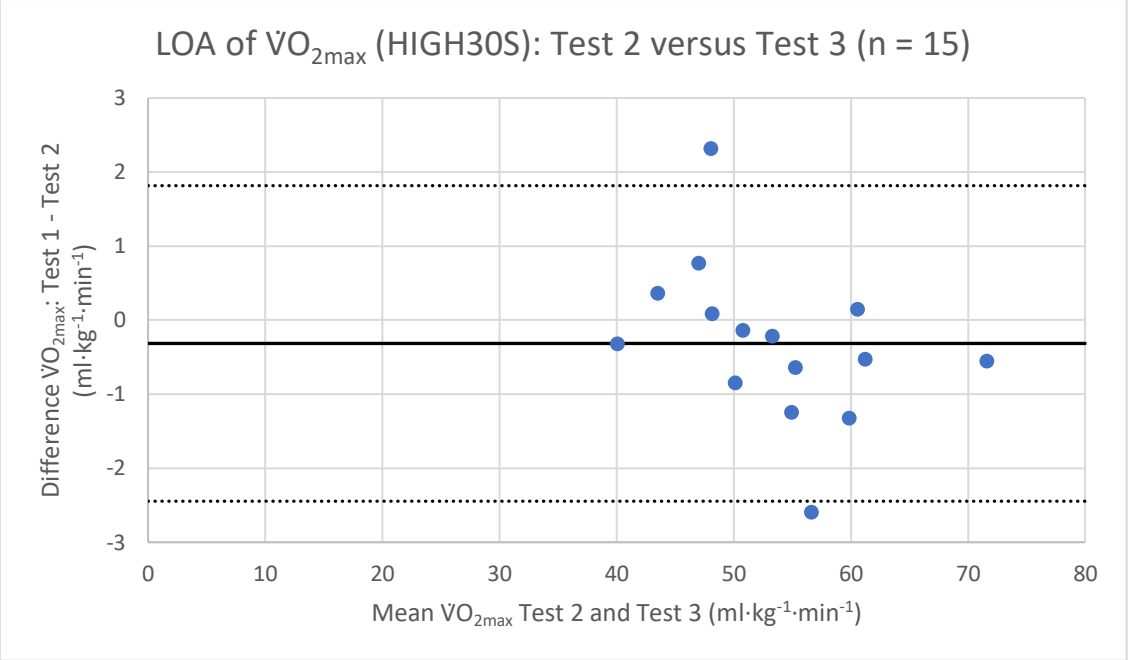


Figure 5.3 Limits of agreement (LOA) of $\dot{V}O_{2max}$ (HIGH30S) between Test 2 and Test 3

Comparing $\dot{V}O_{2max}$ results from the second and third test, the limits of agreement were found to range from 1.82 to -2.45 ml·kg^{-1}·min^{-1} (Figure 5.3), giving a difference range of 4.27 ml·kg^{-1}·min^{-1}. The mean difference between these repeated tests was -0.32 ml·kg^{-1}·min^{-1}. The lower and upper 95% CI for the mean difference were -0.92 ml·kg^{-1}·min^{-1} and 0.29 ml·kg^{-1}·min^{-1}, respectively (ranging 1.20 ml·kg^{-1}·min^{-1}).

Correlation coefficients, coefficients of variation and paired-samples t-tests were also assessed for $\dot{V}O_{2max}$ (Table 5.4) between the repeated tests. To calculate the coefficient of variation, within-participant SD was divided by within-participant mean, and multiplied by 100, for each repeated measure for each individual. The mean CVs for all participants are displayed in Table 5.4. For all comparisons, correlation was high (r > 0.9), percentage variation was low (≤2.6%) and no significant difference was found in any of the tests (p >0.05).

Table 5.4 Statistical analysis between repeated tests for $\dot{V}O_{2max}$ (HIGH30S)

	Test 1 vs Test 2	**Test 2 vs Test 3**
Correlation Coefficient	r = 0.95	r = 0.99
Mean Coefficient of Variation	2.6% (SD ±1.6; range 0.2-6.0%)	1.1% (SD ±1.0; range 0.1-3.4%)
Paired-sample t-test	p = 0.838	p = 0.280

5.3.6 Repeatability of Bruce protocol in measuring $\dot{V}O_{2max}$ averaged from last 30s of exercise test

Limits of agreement were used to analyse the difference between $\dot{V}O_{2max}$ averaged from the last 30s of testing (LAST30S) also, from Test 1 and Test 2, as well as between Test 2 and Test 3 results (Figures 5.4 and 5.5).

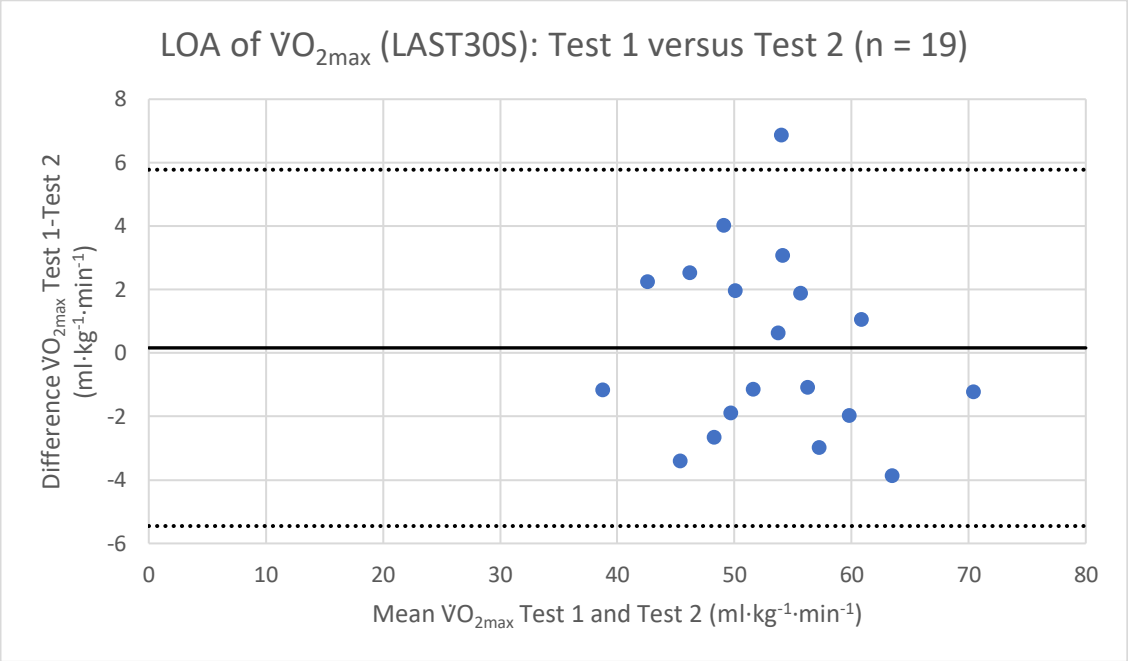


Figure 5.4 Limits of agreement (LOA) of $\dot{V}O_{2max}$ (LAST30S) between Test 1 and Test 2

For LAST30S data, the limits of agreement in $\dot{V}O_{2max}$ between the first and second tests ranged from 5.78 to -5.45 ml·kg^{-1}·min^{-1} (Figure 5.4), with a difference range of 11.23 ml·kg^{-1}·min^{-1}. The mean difference in $\dot{V}O_{2max}$ values was 0.16 ml·kg^{-1}·min^{-1}. The lower and upper 95% CI for the mean difference were -1.22 ml·kg^{-1}·min^{-1} and 1.54 ml·kg^{-1}·min^{-1}, respectively (ranging 2.76 ml·kg^{-1}·min^{-1}).

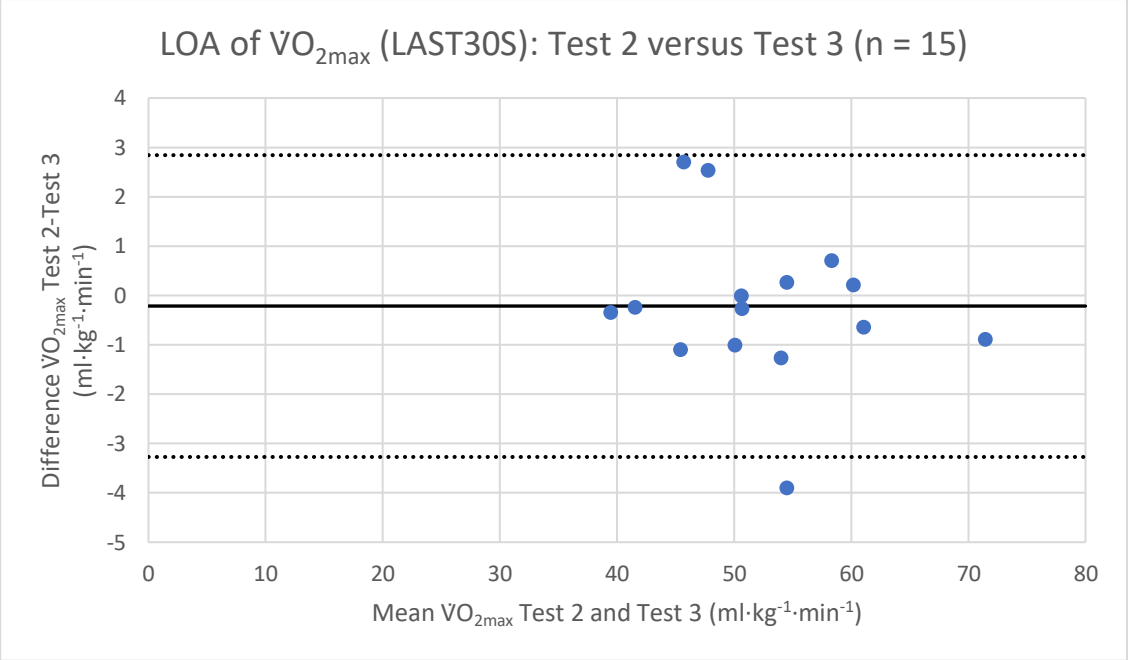


Figure 5.5 Limits of Agreement (LOA) of $\dot{V}O_{2max}$ (LAST30S) between Test 2 and Test 3

Comparing $\dot{V}O_{2max}$ results from the second and third test for LAST30S data, it was found that the limits of agreement were from 2.84 to -3.27 ml·kg^{-1}·min^{-1} (Figure 5.5), giving a difference range of 6.11 ml·kg^{-1}·min^{-1}. The mean difference was -0.22 ml·kg^{-1}·min^{-1}. The lower and upper 95% CI for the mean difference were -1.08 ml·kg^{-1}·min^{-1} and 0.65 ml·kg^{-1}·min^{-1}, respectively (ranging 1.73 ml·kg^{-1}·min^{-1}).

Correlation coefficients, coefficients of variation and paired-sample t-tests were again assessed for this set of $\dot{V}O_{2max}$ data (Table 5.5) between the repeated tests. For all comparisons, correlation was high ($r > 0.9$), percentage variation was low (≤3.3%) and no significant difference was found in any of the tests ($p > 0.05$).

Table 5.5 Statistical analysis between repeated tests for $\dot{V}O_{2max}$ (LAST30S)

	Test 1 vs Test 2	**Test 2 vs Test 3**
Correlation Coefficient	r = 0.93	r = 0.98
Mean Coefficient of Variation	3.3% (SD ±2.0; range 0.9-9.0%)	1.5% (SD ±1.6; range 0.01-5.1%)
Paired-samples t-test	$p = 0.806$	$p = 0.602$

Regardless of sampling method (HIGH30s or LAST30s), the number of participants with an increasing $\dot{V}O_{2max}$ from Test 1 to Test 2 (52.6%) was almost equal to the number of participants decreasing in $\dot{V}O_{2max}$ between their first and seconds tests (47.4%).

Comparing Test 2 to Test 3, 66.7% of participants had an increase in $\dot{V}O_{2max}$ measurement in Test 3 compared to Test 2.

5.3.7 Duration of exercise tests

LOA were also used for assessing difference in maximal test duration between the first and second test repetitions (Figure 5.6) and the second and third repetitions (Figure 5.7).

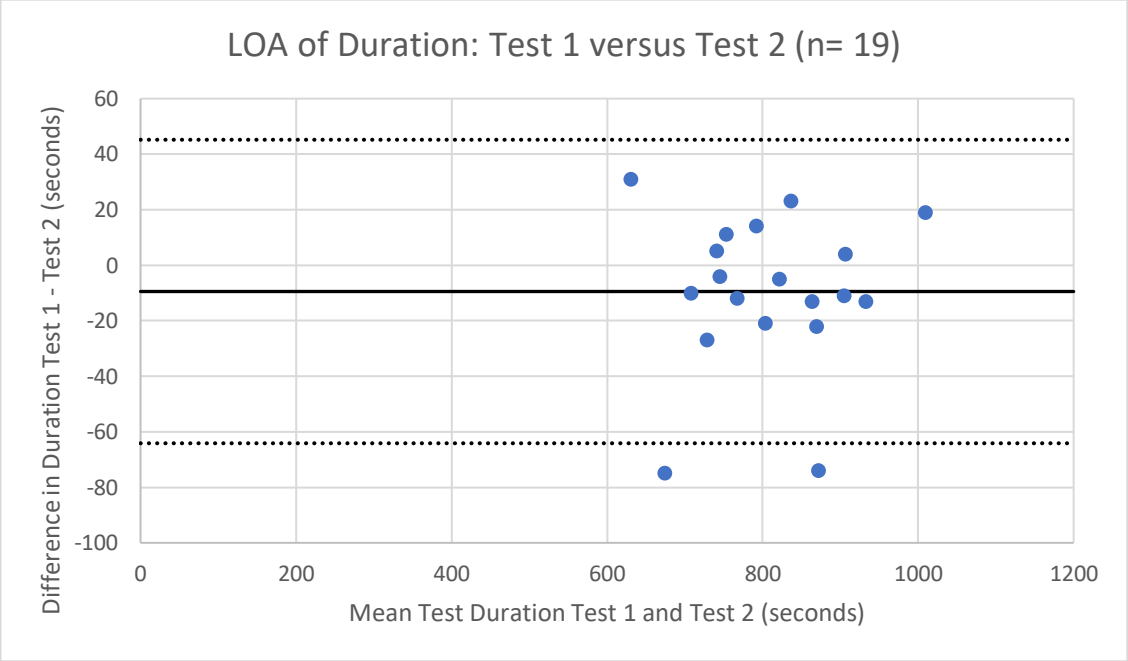


Figure 5.6 Limits of agreement (LOA) for maximal test duration between Test 1 and Test 2

The limits of agreement for difference in duration between Test 1 and Test 2 ranged from 45.19s to -64.14s (Figure 5.6). This gave a range of 109.33s. The mean difference was -9.47s. The lower and upper 95% CI for the mean difference were -22.92s and 3.97s, respectively (ranging 26.89s).

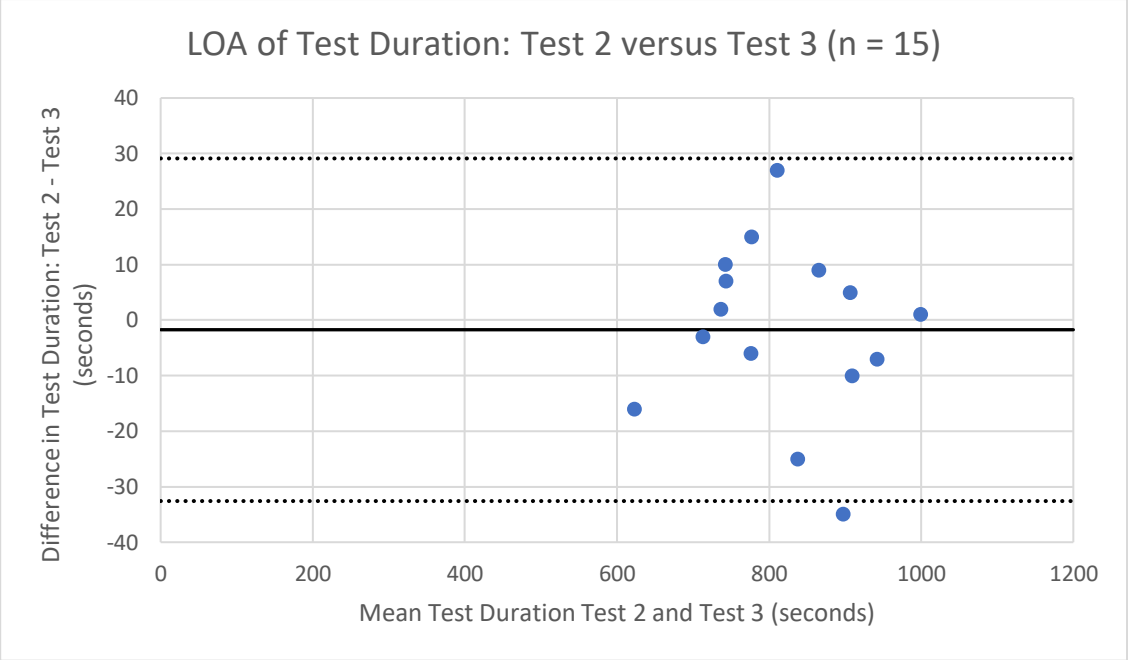


Figure 5.7 Limits of agreement (LOA) for maximal test duration between Test 2 and Test 3

The limits of agreement comparing test durations from Test 2 and Test 3 ranged from 29.12s to -32.59s (Figure 5.7). The range between these two was 61.71s. The mean difference was -1.73s. The lower and upper 95% CI for the mean difference were -10.45s and 6.98s, respectively (ranging 17.44s).

Correlation coefficients, coefficients of variation and paired-samples t-tests were analysed for duration data also (Table 5.6), and again the results indicated high correlation (r>0.9), low variation (CV ≤1.9%) and no statistical difference between tests ($p > 0.05$).

Table 5.6 Statistical analysis between repeated tests for test duration

	Test 1 vs Test 2	**Test 2 vs Test 3**
Correlation Coefficient	r = 0.96	r = 0.99
Mean Coefficient of Variation	1.9% (SD ±2.0; range 0.3-7.9%)	1.0% (SD ±0.8; range 0.1-2.8)
Paired-samples t-tests	$p = 0.156$	$p = 0.676$

The majority of participants achieved a longer test duration in their second test compared to their first (63.2%), while 46.7% of participants increased their test duration from Test 2 to Test 3.

5.3.8 Prediction equations

Paired-samples t-tests were used to compare predicted $\dot{V}O_{2max}$ results to the measured $\dot{V}O_{2max}$ from the HIGH30s method (Table 5.7). All equations, excluding the "ACSMLee" and "Fitmate" equations, were significantly different to measured $\dot{V}O_{2max}$ for both Tests 1 and 2. The "Bruce 1" and "Foster" results were the only significantly different results compared to measured $\dot{V}O_{2max}$ in Test 3.

Table 5.7 Paired-samples t-tests comparing predicted $\dot{V}O_{2max}$ to HIGH30S measured $\dot{V}O_{2max}$

	***p*-values Test 1**	***p*-values Test 2**	***p*-values Test 3**
"Bruce 1"	0.000*	0.000*	0.001*
"Bruce 2"	0.000*	0.005*	0.057
"Foster"	0.000*	0.000*	0.000*
"ACSM2018"	0.003*	0.009*	0.215
"ACSMLee"	0.064	0.869	0.084
"Fitmate"	0.093	0.601	0.253

*Significant p-values indicated with **

Coefficients of variation and correlation coefficients between the measured (HIGH30s) and predicted $\dot{V}O_{2max}$ results are displayed in Table 5.8. Both ACSM equations generally had the highest CVs and the lowest correlation coefficient values, while the other four equations were more highly correlated to the measured $\dot{V}O_{2max}$, and had less percentage variation.

Table 5.8 Mean coefficients of variation (CV) and correlation coefficients (CC) between measured $\dot{V}O_{2max}$ (HIGH30S) and each predicted $\dot{V}O_{2max}$

	CV% Test 1	**CV% Test 2**	**CV% Test 3**	**CC Test 1**	**CC Test 2**	**CC Test 3**
"Bruce 1"	6.7	5.9	5.4	0.92	0.91	0.92
"Bruce 2"	4.6	4.3	3.8	0.92	0.91	0.92
"Foster"	7.5	6.9	6.0	0.92	0.91	0.92
"ACSM2018"	7.8	8.5	8.4	0.72	0.69	0.40
"ACSMLee"	7.8	8.0	8.5	0.43	0.32	0.56
"Fitmate "	5.4	7.1	5.9	0.75	0.64	0.72

Looking at absolute values of measured and predicted $\dot{V}O_{2max}$, and the mean differences between these, it was found that on average the difference between measured and predicted $\dot{V}O_{2max}$ ranged from 4.80 to 5.04 $ml \cdot kg^{-1} \cdot min^{-1}$ (Table 5.9).

Table 5.9 Mean difference (±SD) between measured $\dot{V}O_{2max}$ (HIGH30S) and predicted $\dot{V}O_{2max}$ (absolute values, ml·kg^{-1}·min^{-1})

	Test 1	Test 2	Test 3
"Bruce 1"	4.95 (±2.7)	4.49 (±3.3)	4.14 (±3.2)
"Bruce 2"	3.37 (±2.2)	3.30 (±2.2)	2.95 (±2.3)
"Foster"	5.35 (±2.8)	5.02 (±2.7)	4.43 (±2.8)
"ACSM2018"	5.67 (±2.9)	5.99 (±4.2)	6.33 (±4.6)
"ACSMLee"	5.71 (±4.5)	6.14 (±4.4)	6.41 (±3.8)
"Fitmate"	4.05 (±3.6)	5.30 (±3.1)	4.54 (±3.8)
Mean Difference	4.85 (±0.8)	5.04 (±0.9)	4.80 (±0.8)

Looking specifically at the ACSM equations, it was found that across all three tests, the "ACSM2018" equation under-predicted 69.8% of $\dot{V}O_{2max}$ values, while the "ACSMLee" equation under-predicted 60.4% of $\dot{V}O_{2max}$ results.

The HIGH30S measured $\dot{V}O_{2max}$ results and test durations were examined for each participant in the current study, to determine which stage of the Bruce protocol each participant achieved. Table 5.10 displays the coefficient of variation between the $\dot{V}O_{2max}$ results of participants who reached the same final stage of the Bruce protocol as each other, as well as the range in $\dot{V}O_{2max}$ values that were achieved. This analysis was done to demonstrate the wide range in actual $\dot{V}O_{2max}$ at specific stages of the Bruce protocol, versus what the ACSM running equation would have predicted as $\dot{V}O_{2max}$ for that stage.

The ACSM running equation predictions for $\dot{V}O_2$ achieved at different stages of the Bruce protocol are presented in table 5.11.

Table 5.10 Mean Coefficient of Variation (CV) between participants' $\dot{V}O_{2max}$ results at Stages 4, 5, and 6 of the Bruce protocol, and the ranges in $\dot{V}O_{2max}$ between participants at these stages

Test 1	**Stage 4**	**Stage 5**	**Stage 6**
Mean CV (%)	12.4	9.5	8.8
Range (ml·kg^{-1}·min^{-1})	39.58 – 51.51	44.94 – 62.24	58.93 – 69.81
Test 2	**Stage 4**	**Stage 5**	**Stage 6**
CV (%)	8.4	7.3	10.8
Range (ml·kg^{-1}·min^{-1})	39.91 – 47.22	47.40 – 59.20	53.19 – 71.37
Test 3	**Stage 4**	**Stage 5**	**Stage 6**
CV (%)	5.3	8.5	10.8
Range (ml·kg^{-1}·min^{-1})	40.23 – 43.35	46.92 – 57.98	53.41 – 71.92

Table 5.11 Predicted $\dot{V}O_2$ following ACSM Running equation ($ml \cdot kg^{-1} \cdot min^{-1}$)

	Stage 4	Stage 5	Stage 6
Predicted $\dot{V}O_2$ ($ml \cdot kg^{-1} \cdot min^{-1}$)	42.49	52.37	59.87

All raw data relating to this maximal Bruce protocol repeatability study can be found in Appendix 18 (pg. 150).

5.4 Discussion

5.4.1 Repeatability in measuring $\dot{V}O_{2max}$

The main aim of this study was to examine the repeatability of the Bruce protocol, by comparing the measured $\dot{V}O_{2max}$ between three repeated maximal Bruce treadmill tests. The results from the correlation coefficient analysis, as well as coefficients of variation and paired-samples t-tests, seemed to indicate that the $\dot{V}O_{2max}$ measurements were highly repeatable. Regardless of the method of $\dot{V}O_2$ data sampling, the correlation coefficients were high, the coefficient of variation percentages were low, and the paired-samples t-tests showed no significant difference between the repeated values. However, the limits of agreement plots showed less obvious results. Judging the repeatability of a test cannot be reliant on correlation coefficients or coefficients of variation, where similar results are expected because the same variable has been tested in the same participant under the same conditions[84]. The *differences* must be analysed, and in doing so, it can be seen in the results of this study that there are more differences between results than first meets the eye. The minimal clinically important difference (MCID) for $\dot{V}O_{2max}$ for healthy individuals is generally taken to be 3.5 $ml \cdot kg^{-1} \cdot min^{-1}$ [158,159]. In the comparison between the first and second tests, the limits of agreement spanned 4.49 to -4.71 $ml \cdot kg^{-1} \cdot min^{-1}$ (with HIGH30s data) or 5.78 to -5.45 $ml \cdot kg^{-1} \cdot min^{-1}$ (with LAST30s data) – either of which was clearly a much larger difference between repeated tests than could be accounted for by the MCID. Large increases in $\dot{V}O_{2max}$ may be seen in individuals completing high intensity exercise training between repeated $\dot{V}O_{2max}$ measurements, but not in controls who have had no change in their exercise levels[176-178], just as the current participants had no change in physical activity or exercise training which may have influenced their $\dot{V}O_{2max}$ between tests.

There was a smaller range in limits of agreement between the second and third tests repeated: 1.82 to -2.45 $ml \cdot kg^{-1} \cdot min^{-1}$ (HIGH30S) or 2.84 to -3.27 $ml \cdot kg^{-1} \cdot min^{-1}$ (LAST30S).

Examining these LOA plots (Figures 5.3 and 5.5) it can be seen that the majority of differences lie between 1 and -1 $ml \cdot kg^{-1} \cdot min^{-1}$ for Test 2 versus Test 3 results, which could be accounted for by the MCID. However, such a difference between the first and second repetition of the test – where no difference ought to have been present, as tests were carried out under the same conditions and with no intervention or training occurring between the two tests that might have influenced $\dot{V}O_{2max}$ in the participants – calls into question the reliability and repeatability of the Bruce protocol. Often, this protocol is utilised to determine a change in aerobic capacity in participants or patients post an intervention[179-182]. If such a variation can occur between identical repetitions of a test, how can researchers and clinicians be confident that any change in $\dot{V}O_{2max}$ is due to their intervention, and not merely caused by the test protocol itself?

Similarly, there was a larger variation in duration of maximal tests between the first and second tests, compared to between the second and third tests. As certain prediction equations for $\dot{V}O_{2max}$ depend on the duration of a maximal exercise test[1,113], the variation from one repetition to the other could influence the accuracy of the determined aerobic capacity.

The greater difference between the first and second repetition, compared to between the second and third, is likely due to a learning effect in the participants. Familiarity with what an exercise test entails can influence the outcome in results, whether that be mentally understanding what to expect in the test[81,82], or having previously physically experienced a full exercise test[183]. The current study was designed with no familiarisation session for participants to practice with the test equipment. Participants were given verbal explanations and instructions prior to testing, but the first experience with the treadmill, the testing equipment and the protocol was in completing Test 1 of the Bruce protocol. Effectively, this was their familiarisation session, and resulted in a change in test results and durations in their second repetitions. The participants were blinded throughout all three repetitions to the duration, HR_{max} and $\dot{V}O_{2max}$ that they achieved during testing, to control for self-competition from one test to the next. Testing conditions for each test repetition were carefully standardised, such as same levels of fasting and physical exertion prior to testing, the same time of day for each test, and the same level of physical training, if any, throughout the period of testing.

Therefore, the only difference from Test 1 to Tests 2 and 3 was experience in completing a maximal Bruce protocol. It is clear that learning effect influenced the participants in this study, and that familiarisation with the protocol and equipment should be done with each participant or patient prior to testing. Future research should examine whether completing a full maximal Bruce test is required to eliminate the effect of learning and experience, or if practice walking on the treadmill, and wearing the breath-analysis equipment would suffice.

It is interesting to note that in the current study, the learning effect did not necessarily result in an improvement in $\dot{V}O_{2max}$ and duration of the exercise test. Although there were no statistically significant differences between $\dot{V}O_{2max}$ or test durations between the repetitions, analysing the absolute values of these results can demonstrate certain trends. The number of participants whose $\dot{V}O_{2max}$ increased from Test 1 to Test 2 was almost equal to those whose $\dot{V}O_{2max}$ decreased (52.6% and 47.4%, respectively). Regarding duration, the majority of participants increased their duration from Test 1 to Test 2 (63.2%), but seven participants had a shorter second test compared to their first. The study by Baden et al.[82] found that $\dot{V}O_{2max}$ was lower for participants during a period of running for an unknown duration, compared to when running a known 20-minute duration. This could explain the increase in $\dot{V}O_{2max}$ from Test 1 to Test 2 for some of the current study's participants – as Test 1 was an unknown duration for these participants, and for Test 2 they had a better understanding of what to expect. In contrast, the memory of the strong leg fatigue and cramping or the heavy breathlessness post maximal exertion may have influenced some participants to hold back in their second test subconsciously, resulting in lower $\dot{V}O_{2max}$ values in their repeated test, as psychological factors can influence exercise performance[184].

5.4.2 $\dot{V}O_2$ data sampling methods and repeatability

Another factor that may influence the repeatability of the Bruce protocol is the lack of consensus or clear guidance on the correct sampling method to determine the $\dot{V}O_{2max}$ from the breath-by-breath measurement of $\dot{V}O_2$ throughout the test[165,166,170,175]. Due to "noise" in the data (inaccurate $\dot{V}O_2$ measurements) caused by coughing, sneezing, talking or interference with the measuring equipment during testing, the single, absolute highest value for $\dot{V}O_2$ is generally not taken as $\dot{V}O_{2max}$[169]. Instead, researchers

and clinicians use a variety of data averaging methods, sampling different durations or different numbers of breaths throughout the test or at the point of maximal effort, to determine the person's $\dot{V}O_{2max}$[169,170]. It can be seen from the five data sampling methods used in this study that there was variation in $\dot{V}O_{2max}$ results for each participant, depending on the sampling method used. Significant differences were noted between most sampling methods when paired-samples t-tests were used to compare methods, although the numerical difference between $\dot{V}O_{2max}$ values for the same test was lower than 3.5 $ml \cdot kg^{-1} \cdot min^{-1}$ (MCID for $\dot{V}O_{2max}$) for the majority of participants (Appendix 18.8, pg. 159). However, there *were* differences, and for some participants who achieved a plateau in $\dot{V}O_2$, or a higher point of $\dot{V}O_2$ during the testing than they had in their last thirty seconds, the type of sampling method used did affect their final $\dot{V}O_{2max}$ value. The repeatability analysis did not change between the two sampling methods chosen for comparison in this study, but the differences highlight the importance of consistency in whatever sampling method is used. Whether doing repeat testing, or exercise testing before and after an intervention, one definitive sampling method must be consistently used to analyse the breath-by-breath data.

5.4.3 Consistency in repeatability testing

Previous studies examining the Bruce protocol's repeatability have had varied or vague descriptions exercise testing conditions, instructions for participants and familiarisation with equipment prior to testing[1,17,56,58]. The integrity of a repeatability study depends on identical testing procedures and set up from one test to the next, as well as in an accurate choice of statistical analysis[29]. It is therefore imperative to have a strict testing procedure for clinical exercise testing, whether being used as a once-off test to determine aerobic capacity or to measure change post intervention. Results can be reliably compared once it is certain they have come from the same testing procedure. This is difficult to achieve for $\dot{V}O_{2max}$ testing, when there is such little guidance and consensus on pre-test instructions, familiarisation for participants with equipment and protocols, and data sampling methods to determine the highest $\dot{V}O_2$ level measured. The current study was designed to account for as many variables as possible, following practices from previously published repeatability studies, combined with guidance from the ACSM and experience from completing the submaximal testing study (Chapter 4).

Participants were male-only, to eliminate any influence of the menstrual cycle on the obtained $\dot{V}O_{2max}$ results. All participants fasted from midnight the night before testing and were requested to limit their physical activity levels prior to testing. To the researchers' best ability, participants completed their repeated tests on the same day and at the same time, each a week apart, and tests were conducted by the same researcher at each repetition. These measures controlled for many variables, to ensure that any found variation between $\dot{V}O_{2max}$ in repeated tests could be attributed to the Bruce protocol itself, and not to other reasons.

5.4.4 Reasons for Test Termination

Some researchers have found when comparing the Bruce protocol to other treadmill protocols that participants struggle with the protocol due to the steep, progressive incline and are limited more by leg cramping and discomfort rather than feeling they reached their maximal aerobic fatigue[24,28,185]. Participants reported similar difficulties in the current study, with 49.1% of participants quoting leg fatigue or cramping due to the incline as their reason for voluntary termination of the test, rather than breathlessness. A further 13.2% of tests were ended due to leg fatigue along with breathlessness. It may be hypothesised that if these participants had not suffered from leg pain due to the steepness of the treadmill, they may have been able to continue for a longer duration and potentially achieve a higher $\dot{V}O_2$ than they did in their current tests. It is another reason, although subjective, for querying the reliability and repeatability of the Bruce protocol – as this protocol is influenced by leg strength and endurance, as well as aerobic capacity.

5.4.5 Predicting $\dot{V}O_{2max}$

The importance of having the ability to predict $\dot{V}O_{2max}$ for people who cannot complete a maximal test, or for situations where the technical equipment is not available, has been well established[30,89,90]. Six equations were examined in the current study. The analysis results showed that many of the prediction equations differed significantly from the measured $\dot{V}O_{2max}$ values. The mean coefficient of variation between measured and predicted values ranged from 3.8% to 8.5%. Correlation coefficients comparing the "Bruce 1", "Bruce 2" and "Foster" predictions to the measured $\dot{V}O_{2max}$ showed high

correlation; however, as discussed previously, correlation coefficients are not always a reliable statistic when comparing two measurements of the same value or variable. The CC values for both ACSM equations against measured $\dot{V}O_{2max}$ were low, which is in line with previous research demonstrating the inaccuracy of the ACSM equations[31,35,102,103,120]. An interesting finding in the current study was that both ACSM equations tended to under-predict, rather than over-predict, the $\dot{V}O_{2max}$ value, which is contrast to other study findings[35,102,103]. Either way, it has been shown that the ACSM equations are not accurate or reliable in predicting $\dot{V}O_{2max}$.

Another issue with the method of predicting $\dot{V}O_{2max}$ using the ACSM running equation as explained in the ACSM's Guidelines for Exercise Testing and Prescription book[41] is that every participant reaching a specific stage receives the same $\dot{V}O_{2max}$ predicted value (for example, everyone reaching Stage 5 has a predicted $\dot{V}O_{2max}$ of 52.37 $ml \cdot kg^{-1} \cdot min^{-1}$) – when in reality, the variation in $\dot{V}O_{2max}$ for a specific stage among participants is large. This is demonstrated in Table 5.10. For example, in the current study for participants who reached Stage 5 of the Bruce protocol in their Test 1, the mean coefficient of variation between their results was 9.5%, with a range in values from 44.94 to 62.24 $ml \cdot kg^{-1} \cdot min^{-1}$. The same analysis for Test 2 and Test 3 between results of participants reaching Stage 5 showed mean coefficients of variation of 7.3% and 8.5% respectively, and ranges of 47.4 to 59.2 $ml \cdot kg^{-1} \cdot min^{-1}$, and 46.92 to 57.98 $ml \cdot kg^{-1} \cdot min^{-1}$, respectively. There is clearly a wide variation in $\dot{V}O_{2max}$ for one stage of the Bruce protocol, and therefore, an equation that generalises the prediction so broadly across a full three-minute stage is not reliable.

The "Fitmate" prediction equation was examined as it was the method used to predict $\dot{V}O_{2max}$ in the previous submaximal study (Chapter 4). Coefficients of variation and mean difference values were similar to those of the other equations examined, while correlation coefficient results lay between the high results found for the Bruce and Foster equations, and the lower correlation found for the ACSM equations. Regarding paired-samples t-tests, there was no significant different between predicted and measured $\dot{V}O_{2max}$ for any of the test repetitions. These results demonstrated that the "Fitmate" equation may be more reliable than the ACSM equations, but not as accurate as those predictions from the Bruce or Foster equations.

Looking at the absolute values of mean differences in $\dot{V}O_{2max}$ between measured and predicted methods, it can be seen that all equations excluding "Bruce 2" varied from the measured $\dot{V}O_{2max}$ value by more than the MCID for $\dot{V}O_{2max}$ (3.5 $ml \cdot kg^{-1} \cdot min^{-1}$). Therefore, by using any of these equations, it is not possible to ensure that the value predicted is accurate. If using the equations to determine a change in $\dot{V}O_{2max}$ post-intervention, for example, it will not be clear if the change is due to prediction error, or a physiological change in the person themselves, as the predicted value could vary a great deal from the true $\dot{V}O_{2max}$ for that person. The current study found that the "Bruce 2" equation showed the smallest percentage variation, indicating that it may be the only reliable prediction equation for $\dot{V}O_{2max}$ in healthy male populations, for use with the Bruce treadmill protocol.

5.4.6 Study Limitations and Future Recommendations

The current study was limited by the small number of participants included. Although the sample size calculation was for twenty-three people, the shut-down of the University due to the COVID-19 pandemic resulted in only fifteen participants completing all three tests. Similar to the submaximal study from Chapter 4, a limitation of this maximal repeatability study is the small sample size, which did not achieve the required number as per the sample size calculation. This may have impacted the statistical results of the study, as having a small sample size reduces the power of the study[62-64]. The main method used for determining repeatability in this study was limits of agreement, where a wide confidence interval for the mean difference may reflect a small sample size. Larger sample sizes would demonstrate a narrow 95% confidence interval (95% CI) if the measure was in fact repeatable[84, 137]. However, in this maximal study, the 95% CI for the mean difference in $\dot{V}O_{2max}$ testing between Test 1 and Test 2 ranges 2.26 $ml \cdot kg^{-1} \cdot min^{-1}$, and ranges 1.2 $ml \cdot kg^{-1} \cdot min^{-1}$ for results between Test 2 and Test 3. Both these ranges are within the MCID of 3.5 $ml \cdot kg^{-1} \cdot min^{-1}$, and therefore are deemed narrow confidence intervals. This indicates that the small sample size did not affect the limits of agreement analysis for this study, and the conclusions regarding the repeatability of the Bruce protocol based on limits of agreement may be deemed valid. Future studies would benefit from having larger sample sizes to improve the power of these findings.

Another limitation regarding participants was that the participant group was relatively homogenous, due in part to the inclusion/exclusion criteria, but also perhaps by the type of person interested in volunteering for fitness research. The age range for eligible volunteers was 18- to 45-year olds, however the majority of participants were under 25 years of age (63.2%). All participants were physically active at baseline, and all participants were at a low risk of any cardiovascular disease. Including participants with a wider risk level, a broader physical ability level, and looking at females as well as males would give a clearer picture of how repeatable the Bruce protocol is across a general population. This is an important consideration for future research, as the protocol is used in with a wide variety of populations[21,22,186-188]. The consistency in testing procedures and set up was discussed above; however, the limitations of timetabling and human decisions must be considered also. Although every attempt was made to ensure identical repeated tests, the seven-day duration between tests was sometimes unachievable due to participant illness or changes in their schedule. The influence of human error on human research may be inevitable and unavoidable, but future research into repeatability should continue to strive for consistency and control over as many variables as possible. The best method of data sampling and averaging still requires further research to achieve a consensus, but researchers should ensure whichever method chosen is used consistently throughout testing, to ensure results can be reliable compared to one another.

5.5 Conclusion

The results of the current study show that the Bruce protocol is not repeatable for measuring $\dot{V}O_{2max}$ in healthy male participants between the first and second repetition of a test. If, however, a session of familiarisation has been conducted with a full Bruce protocol maximal test, the results between a second and third repetition are likely repeatable. The influence of learning effect from the first to the second repetition of the test must be accounted for in future testing, so that all participants have a physical understanding of the testing experience. In completing Bruce protocol testing, a clearly defined protocol for familiarisation with testing equipment, verbal instructions, preparations for participants and facility set up must be followed, to ensure accurate

measurement of $\dot{V}O_{2max}$ during each test. The best method of breath-by-breath data analysis is currently unknown, and therefore it is recommended that one clear and consistent method is chosen and followed for all tests carried out with an individual in any specific research or clinical facility. If the use of prediction equations is required to determine $\dot{V}O_{2max}$, research must be done to choose an equation that best suits the population to be tested. The current study concludes that the most accurate prediction equation for use with healthy male adults is the “Bruce 2”, and the ACSM equations should be avoided for $\dot{V}O_{2max}$ prediction, when used with the Bruce treadmill protocol.

Chapter 6: Discussion

6.1 Introduction

The four studies in this thesis have each highlighted the importance of repeatability within the Bruce treadmill protocol, as well as that currently its repeatability is still questionable. Both literature reviews into the topics of repeatability and prediction equations provided clear areas for further research. With guidance from these reviews, two exercise studies were developed and conducted. Comparing the results of all four studies, the similarities and differences which were found between previous literature and the current studies' findings, emphasise the importance of clear, specific testing protocols, of accurate choice in statistical analysis of results, and of the strong influence of learning effect on repeated tests. These similarities and differences are discussed below.

6.2 Literature Reviews and their Influence

The literature search for previous Bruce protocol repeatability studies spanned several decades of research, with eleven studies found eligible for analysis. The quality of reporting in these studies was generally poor, as determined by the AXIS risk of bias assessment tool[52]. Scores for the assessment were generally low in areas regarding defining and justifying participants, and regarding statistical analysis. These findings assisted with the design of the submaximal and maximal studies that were to follow in this research masters, where a clear selection process for participants and calculation and justification of sample size was completed. The statistical methods chosen for analysing data, as well as the careful presentation of the results were also influenced by the poor quality of presentation in previous repeatability literature[1,17,53-55,58,60].

The submaximal and maximal research studies included only healthy male participants, although the repeatability literature review included males and females both healthy and with cardiac conditions. Although research has generally shown that different phases of the menstrual cycle do not have an effect on female $\dot{V}O_{2max}$[189-191], the use of oral contraceptives may affect $\dot{V}O_{2max}$[190,192,193], and the menstrual cycle may affect ventilation[194,195], lactate threshold or blood lactate response to exercise[196,197] and

submaximal $\dot{V}O_2$ values[198]. There is difficulty in generalising findings relating to variation during the menstrual cycle in physiological and performance factors across the female population[199]. As the current research was examining repeatability, and there is high importance in such research to control for as many variables as possible[29,200,201], the maximal and submaximal studies for the current thesis included only male participants, to eliminate any possible variation in $\dot{V}O_{2max}$ that could have been due to stage of menstrual cycle for female participants, rather than due to the Bruce protocol itself. For similar reasons, due to the strong relationship between $\dot{V}O_{2max}$ and cardiovascular disease[2], only healthy participants with low risk of cardiovascular disease (CVD) were included in the studies – again to control for any effect CVD may have on the $\dot{V}O_{2max}$ measurements. Future research into the repeatability of the Bruce protocol would do well to further investigate healthy females as well as those with cardiac or other medical conditions.

Previous repeatability studies and studies examining prediction equations showed wide variation with regards to pre-test instructions and conditions, and guidance during the testing itself, with several studies publishing no information regarding instructions given to participants[1,17,53,61,100,105]. The varying instructions across the literature guided the formation of the current studies' detailed protocols, which had strict instructions regarding fasting duration prior to testing, levels of physical activity allowed during the testing weeks, and verbal instruction during each maximal test.

6.3 Bruce Protocol Repeatability

6.3.1 Repeatability Statistical Analysis

As previously discussed in both the repeatability literature review and in the maximal Bruce protocol repeatability study (Chapters 2 and 5), correlation coefficients are not sufficient to independently determine the repeatability of the Bruce protocol. Correlation coefficients were included in both the submaximal and maximal repeatability studies during data analysis, to have as a comparison to previous study findings, but the conclusions drawn about the repeatability were based on analysis from multiple statistical tests conducted. It is interesting to compare the findings from

previous literature, to the submaximal and maximal findings from the current thesis. The literature review yielded the conclusion that further research into the repeatability of the Bruce protocol was required, with focus on limits of agreement analysis, and a more standardised procedure for conducting the Bruce protocol. The majority of previous studies only conducted two Bruce treadmill tests, and therefore based repeatability conclusions on comparisons between first and second $\dot{V}O_{2max}$ test results. The submaximal study (Chapter 4) was designed as an initial study to examine the repeatability of the Bruce protocol. Due to a learning effect identified in the literature review, three Bruce treadmill tests were conducted, with repeatability conclusions based on the comparison of second and third test $\dot{V}O_{2max}$ results. While correlation coefficients and coefficients of variation were similar to those from the literature review, the range in limits of agreement was much broader, when compared to the results from the study by Jakovljevic et al.[59]. Comparing Test 1 and Test 2 results, Jakovljevic et al.[59] found a range in LOA of 2.7 to -5.3 $ml \cdot kg^{-1} \cdot min^{-1}$. The submaximal study's LOA ranged from 4.86 to -10.35 $ml \cdot kg^{-1} \cdot min^{-1}$ between Tests 1 and 2, and between Tests 2 and 3 ranged from 6.67 to -7.96 $ml \cdot kg^{-1} \cdot min^{-1}$. These findings demonstrated that at a submaximal testing level, with $\dot{V}O_{2max}$ predicted through an equation, the Bruce protocol was not repeatable in determining the $\dot{V}O_{2max}$ for a healthy male population. Therefore, the research was continued following the maximal Bruce protocol with a new sample of participants from the same population, hypothesising that the repeatability analysis based on limits of agreement would show similarly poor results comparing repeated $\dot{V}O_{2max}$ measurements.

It was found, however, that the maximal exercise test study results more closely resembled the results from the repeatability literature review than they did the submaximal study results. This was especially in relation to the limits of agreement, which showed much narrower limits under maximal treadmill testing than with submaximal testing. This may have indicated that the maximal Bruce protocol was repeatable, where the submaximal version of the test was not. One similarity, however, between the submaximal and maximal testing studies was the learning effect noted in regard to repeated $\dot{V}O_{2max}$ measurements across the three repeated tests. This is an

important factor when considering the Bruce protocol's repeatability in measuring $\dot{V}O_{2max}$, and is discussed below.

6.3.2 Learning Effect

Learning effect is a topic that arose across all repeatability analysis in this thesis. Its influence over the accurate determination of aerobic fitness has been widely discussed. Several studies that found differences between their first and second repetitions of the Bruce protocol in $\dot{V}O_{2max}$ and test duration put these findings down to learning effect[54,55,59]. The submaximal Bruce protocol study identified a significant difference between first and second repeated tests (p = 0.013), but no significant difference between Test 2 and Test 3 results (p = 0.473). A larger range in LOA, and a larger mean difference in $\dot{V}O_{2max}$ values from Test 1 to Test 2, compared with differences from Test 2 to Test 3 was also observed. This was put down to the influence of a learning effect for participants. Similar results were found in the maximal Bruce protocol study. Although no statistically significant difference was found between any repeated tests, the LOA analysis identified a larger difference between the first and second $\dot{V}O_{2max}$ values, compared with differences between Test 2 and Test 3 results. As every measure was taken to ensure identical testing conditions for each test repetition, the difference in $\dot{V}O_{2max}$ results was attributed to learning effect across the three repetitions.

In conclusion for Bruce protocol repeatability, the maximal protocol appears to be repeatable in measuring $\dot{V}O_{2max}$, but only from a second to third repetition of the test. Due to the presence of a learning effect, familiarisation with the full maximal Bruce protocol test should be included with all participants prior to testing, because of the large range in $\dot{V}O_{2max}$ between the first and second test repetitions found in the current study.

Looking at the submaximal Bruce protocol study results, however, the prediction of $\dot{V}O_{2max}$ from submaximal Bruce protocol testing is not repeatable, with a large variation in differences between repeated $\dot{V}O_{2max}$ predicted results. This may be due to the prediction equation chosen, or due to the submaximal testing protocol itself, and further research is warranted to investigate both of these factors more closely.

6.4 Prediction Equations

6.4.1 Choice of prediction equations for current research

The Fitmate prediction equation described by Lee et al.[31] was used in the submaximal study, as it the most suitable equation identified for use following the literature review of Chapter 3. In the maximal Bruce protocol study, the "Fitmate" equation, along with five other equations also identified through the literature review were chosen for analysis. All the chosen equations utilised exercise-based variables to calculate $\dot{V}O_{2max}$. The literature review in Chapter 3 discussed prediction equations using non-exercise variables, finding mixed results regarding the benefit of these types of equations. It was concluded that as the gold standard for measuring $\dot{V}O_{2max}$ comes from completing a maximal exercise test, that equations based on exercise data would provide more accurate results than those based on non-exercise variables alone. Non-exercise equations are generally used due to their convenience as no exercise testing is required[125,131,132], but as the maximal Bruce protocol study would provide specific exercise data for participants, non-exercise equations were therefore not chosen for analysis.

Throughout the current research, the Tanaka et al.[154] equation was used to predict maximal heart rate based on age. Research has shown that the commonly used "220-age" equation is unreliable in predicting HR_{max} accurately in various populations[154,202,203]. It is difficult to attribute credit for the creation of the "22-age" formula[204] and it has been recommended by a number of researchers that other formulas, specific to the population, should be used to determine HR_{max} rather than "220-age"[204,205]. The Tanaka et al.[154] equation was chosen as its formulation was based on healthy males and females (similar to the current population of healthy males), and because it has been recommended as an accurate and suitable HR_{max} predictive equation by the ACSM[41] and other researchers[202,206].

6.4.2 Accuracy of equations used

When examining the predicted $\dot{V}O_{2max}$ values for participants in the submaximal Bruce protocol study, it was found that the majority of results were classed as "superior" or "excellent" (92.3%), compared to age-based normal $\dot{V}O_{2max}$ values[155]. The results

seemed unexpectedly high, but since no data was gathered regarding the physical activity levels of these participants (such as their average time spent in physical activity per week, or the level and intensity of training in which they participated) and as actual $\dot{V}O_{2max}$ was not measured in these participants, it was not possible to conclude that the prediction equation had over-estimated the participants' $\dot{V}O_{2max}$. Perhaps all those who volunteered for the study were very physically active and fit individuals, and did indeed have superior $\dot{V}O_{2max}$ levels compared to their peers. The maximal Bruce protocol study drew from the same population as the submaximal study of healthy male staff and students from Trinity College Dublin, and so the results from both studies could be accurately compared to each other. Participants in the maximal Bruce protocol study completed the International Physical Activity Questionnaire – Short Form (IPAQ-SF)[172] and also wore a physical activity monitor (ActiGraph) for seven days during the study, in an attempt to gain a clear picture of their normal physical activity levels. According to their self-reported physical activity levels, 79% of participants achieved at least 150 minutes of moderate-vigorous physical activity (MVPA) per week. Comparing this to their ActiGraph physical activity levels, all participants achieved more than 150 minutes of MVPA per week, but only 55.6% had MVPA in bouts of ten minutes or longer. The American College of Sports Medicine[41] recommends MVPA be conducted in bouts of ten minutes or longer to gain physiological benefits from exercise.

The $\dot{V}O_{2max}$ results for the maximal Bruce protocol study were compared to the normal values provided by Heyward and Gibson[155]. It was found that for these participants, the results were more evenly dispersed across the normal values (superior = 41.5%; excellent = 18.9%; good = 30.2%, fair/poor = 9.4%). The physical activity data gathered through the IPAQ-SF and ActiGraph monitoring showed that the participants in the maximal Bruce protocol study were generally physically active people, and therefore the majority of $\dot{V}O_{2max}$ values being classed as "good" or above reflected this. It can be assumed that participants from the submaximal study would have had similar physical activity levels to those in the maximal study, and so their predicted $\dot{V}O_{2max}$ values should have been dispersed similarly to that of the maximal Bruce study results. This finding indicates that the "Fitmate" equation may have over-predicted the $\dot{V}O_{2max}$ results in the submaximal study, although when comparing the measured $\dot{V}O_{2max}$ results with those

predicted by the various equations in the maximal Bruce protocol study, the "Fitmate" equation performed similarly to other equations ("Bruce 1", "Bruce 2" and "Foster" equations). Taking all these findings into consideration, it may be concluded that the "Fitmate" equation appears to incorrectly predict $\dot{V}O_{2max}$ in healthy male participants, but further research with a larger cohort should be conducted to confirm this.

Regarding the ACSM equations, the literature review in Chapter 3 concluded that these equations overestimated $\dot{V}O_{2max}$ for healthy participants. In contrast to this, it was found in the maximal Bruce protocol study that both ACSM equations examined ("ACSM2018" and "ACSMLee") under-estimated participants' $\dot{V}O_{2max}$. The "ACSM2018" equation underestimated 69.8% of all tests, while the "ACSMLee" equation underestimated 60.4% of all tests. The studies that examined the ACSM equations from the predictions equations literature review had larger participant numbers than the current study[31,102,103], and two of these included female participants[31,102], which could account for the differences in results. Other studies have found that $\dot{V}O_{2max}$ prediction equations tend to under-predict $\dot{V}O_{2max}$ for fitter populations[99,207,208]. As previously discussed, the participants from the maximal Bruce protocol study were fit individuals with generally high $\dot{V}O_{2max}$ results, and this could account for the underprediction of their $\dot{V}O_{2max}$ with the ACSM equations, compared to previous research. In either respect, it can be concluded that the ACSM equations have not been shown to accurately predict $\dot{V}O_{2max}$ for healthy individuals, and so should not be used in future research.

One finding from the prediction equations literature review (Chapter 3) was that equations based solely on exercise test duration failed to accurately predict $\dot{V}O_{2max}$. In the maximal Bruce protocol study, both the "Bruce 2" and the "Foster" equations used only duration data. It was found that both equations had statistically significant differences between all their predictions and the corresponding measured $\dot{V}O_{2max}$ (excluding predicted values by "Bruce" 2 in Test 3, $p = 0.057$). However, both equations had high correlation coefficient values ($r \geq 0.91$) with the measured $\dot{V}O_{2max}$, and "Bruce 2" had the lowest mean coefficient of variation percentages across all examined prediction equations (CV% ≤4.6%). Therefore, the results from the current research is in contrast with previous findings, and duration as a prediction variable may have strong influence over the predicted $\dot{V}O_{2max}$. Further research into duration as a prediction

variable, whether individually or as a combination with other variables such as treadmill grade and speed, or non-exercise data, is still required.

6.5 Limitations

The main limitations of the current research have been the small sample sizes in both studies. As the submaximal study was a pilot study and undertaken as part of an undergraduate research project, the sample size was small due to the time limitation on the project. For the maximal Bruce protocol study, the sample size calculation was for twenty-three participants. When the research began in October 2019, it was anticipated that there would be sufficient time for recruitment and testing. Due to the Covid-19 pandemic, however, this time was cut short in March 2020. There were a number of volunteers scheduled to begin testing at that time, whose participation was first postponed and subsequently cancelled due to continued University closure. There had also been a recruitment plan to display more posters and send the recruitment email to further eligible staff and students for the end of term and summer periods, which researchers were unable to carry out due to the pandemic. As discussed in both Chapters 4 and 5 regarding the impact of the small sample sizes, these may have affected the statistical results of the submaximal study. However, the maximal Bruce protocol study took into account the many factors affecting repeatability and $\dot{V}O_{2max}$ testing, and its limits of agreement results demonstrated that the small sample size did not affect its findings – therefore, the conclusions that the maximal Bruce protocol is not repeatable for measuring $\dot{V}O_{2max}$ between its first and second repetitions may be taken to be a valid result. Future studies should aim to have a larger sample size to strengthen the findings found from the current research.

Another limitation in both studies was the homogenous sample of participants regarding fitness level and age. Although people were eligible to participate if aged between eighteen and forty-five, the majority of participants were under 25 years of age. As recruitment was taken from the staff and students of Trinity College Dublin, the findings may not be universally applicable to the population of “healthy males”.

6.6 Future Recommendations

While the maximal Bruce protocol exercise test appears repeatable in healthy adult males, the main finding of the current research is the influence of learning effect on the $\dot{V}O_{2max}$ results between first and second rest repetition. Future studies into the area of Bruce protocol repeatability in measuring $\dot{V}O_{2max}$ should focus on exploring this learning effect further, with larger participant sample sizes. It should be determined whether a full maximal Bruce protocol is required as familiarisation, to eliminate the learning effect on $\dot{V}O_{2max}$ measurement, or whether a submaximal version of the protocol would suffice. The duration between the familiarisation session and the actual exercise test session to measure $\dot{V}O_{2max}$ should also be examined, to determine how long the learning effect has influence over the $\dot{V}O_{2max}$ measurement.

Expanding this repeatability research with broader population cohorts should also be done, such as female participants, younger and older cohorts, and those with cardiac conditions as well as other clinical conditions. The Bruce protocol is widely used across many populations both clinical and healthy, and although it appears repeatable in the currently examined healthy male cohort, its repeatability and in particular the effect of practice and learning on the measured results in other populations should be examined. As performed in the current research, repeatability studies should also use more detailed and specific statistical analysis in determining the Bruce protocol's repeatability in measuring $\dot{V}O_{2max}$, rather than simply looking at correlation between repeated measurements. The differences between repeated measures is highly important, and the use of limits of agreement, as described by Bland and Altman[84], will more clearly display how repeatable the measurement is.

Regarding prediction equations, further research is required into determining the best variable or combination of variables to accurately predict $\dot{V}O_{2max}$ in healthy males. Future research focusing on whether exercise-based equations are more accurate for individuals, compared to those using non-exercise variables alone would be beneficial in this field. Often the ability to complete a maximal treadmill test, or to measure $\dot{V}O_2$ during exercise testing, is restricted for researchers and clinicians, and having a reliable equation to use alongside the Bruce treadmill protocol or as a stand-alone equation, would be of great benefit. This again should be examined in a wide population base of

males and females, as well as in healthy and clinical cohorts, to make research findings more widely applicable.

Chapter 7: Conclusion

In conclusion, the maximal Bruce protocol is repeatable when used to directly measure $\dot{V}O_{2max}$ in healthy male participants, but only when a full exercise test is completed as familiarisation. Learning effect from the first to second repetition of the test may influence the $\dot{V}O_{2max}$ results, and should be taken into account with familiarisation sessions to ensure accurate testing results. The submaximal form of the Bruce protocol treadmill test was not found to be repeatable in predicting $\dot{V}O_{2max}$.

When measuring $\dot{V}O_{2max}$ is not possible directly, and prediction equations are required in conjunction with the Bruce protocol, the current research concludes that the ACSM equation – based on the grade and speed of treadmill tests – and the "Fitmate" equation as described by Lee et al.[31] – based on measured $\dot{V}O_2$ and heart rates – are not accurate methods of predicting $\dot{V}O_{2max}$. Prediction equations based on exercise data are preferable to those based on non-exercise variables. There are a number of variables that may be used, such as heart rate or test duration, but the findings of the present work are inconclusive as to which variables are the most accurate for $\dot{V}O_{2max}$ prediction. If at all possible, $\dot{V}O_{2max}$ should be measured through a maximal exercise test, and not predicted with equations, to ensure the most accurate result.

References

1. Bruce RA, Kusumi F, Hosmer D. Maximal oxygen intake and nomographic assessment of functional aerobic impairment in cardiovascular disease. Am Heart J. 1973;85(4):546-62.
2. Ross R, Blair SN, Arena R, Church TS, Després J-P, Franklin BA, et al. Importance of assessing cardiorespiratory fitness in clinical practice: a case for fitness as a clinical vital sign: a scientific statement from the American Heart Association. Circulation. 2016;134(24):e653-e99. http://doi.org/10.1161/CIR.0000000000000461.
3. Sallis JF, Patterson TL, Buono MJ, Nader PR. Relation of cardiovascular fitness and physical activity to cardiovascular disease risk factors in children and adults. Am J Epidemiol. 1988;127(5):933-41.
4. Beltz NM, Gibson AL, Janot JM, Kravitz L, Mermier CM, Dalleck LC. Graded exercise testing protocols for the determination of $\dot{V}O_{2max}$: historical perspectives, progress, and future considerations. J Sports Med (Hindawi Publ Corp). 2016;2016:3968393 http://doi.org/10.1155/2016/.
5. Beltrami FG, Froyd C, Mamen A, Noakes TD. The validity of the Moxus Modular metabolic system during incremental exercise tests: impacts on detection of small changes in oxygen consumption. Eur J Appl Physiol. 2014;114(5):941-50.
6. Garber CE, Blissmer B, Deschenes MR, Franklin BA, Lamonte MJ, Lee I-M, et al. Quantity and quality of exercise for developing and maintaining cardiorespiratory, musculoskeletal, and neuromotor fitness in apparently healthy adults: guidance for prescribing exercise. Medicine and science in sports and exercise. 2011;43(7):1334-59.
7. Bergen JL, Toole T, Elliott III RG, Wallace B, Robinson K, Maitland CG. Aerobic exercise intervention improves aerobic capacity and movement initiation in Parkinson's disease patients. NeuroRehabilitation. 2002;17(2):161-8.
8. Verges B, Patois-Verges B, Cohen M, Lucas B, Galland-Jos C, Casillas JM. Effects of cardiac rehabilitation on exercise capacity in Type 2 diabetic patients with coronary artery disease. Diabet Med. 2004;21(8):889-95.
9. Regensteiner JG, Bauer TA, Reusch JEB, Brandenburg SL, Sippel JM, Vogelsong AM, et al. Abnormal oxygen uptake kinetic responses in women with type II diabetes mellitus. J Appl Physiol. 1998;85(1):310-7.
10. Çimen ÖB, Deviren SD, Yorgancıoğlu ZR. Pulmonary function tests, aerobic capacity, respiratory muscle strength and endurance of patients with rheumatoid arthritis. Clin Rheumatol. 2001;20(3):168-73.
11. Antolini MR, Weston ZJ, Tiidus PM. Physical fitness characteristics of a front-line firefighter population. Acta Kinesiologiae Universitatis Tartuensis. 2015;21:61-74.
12. Myers J, Buchanan N, Walsh D, Kraemer M, McAuley P, Hamilton-Wessler M, et al. Comparison of the ramp versus standard exercise protocols. J Am Coll Cardiol. 1991;17(6):1334-42.
13. Basset FA, Boulay MR. Specificity of treadmill and cycle ergometer tests in triathletes, runners and cyclists. Eur J Appl Physiol. 2000;81(3):214-21.
14. American Thoracic Society, American College of Chest Physicians. ATS/ACCP Statement on cardiopulmonary exercise testing. Am J Respir Crit Care Med. 2003;167(2):211-77.
15. Lear SA, Brozic A, Myers JN, Ignaszewski A. Exercise stress testing. Sports Med. 1999;27(5):285-312.
16. Luong MW, Ignaszewski M, Taylor CM. Stress testing: A contribution from Dr Robert A. Bruce, father of exercise cardiology. B C Med J. 2016;58(2):70-6.
17. Bruce RA, Blackmon JR, Jones JW, Strait G. Exercising testing in adult normal subjects and cardiac patients. Pediatrics. 1963;32(4):742-56.

18. Venkata Ramana Y, Surya Kumari MVL, Sudhakar Rao S, Balakrishna N. Effect of changes in body composition profile on $\dot{V}O_{2max}$ and maximal work performance in athletes. J Exerc Physiol Online. 2004;7(1):34-9.
19. Dabney U, Butler M. Predictive ability of the YMCA test and Bruce test for triathletes with different training backgrounds. Emporia State Res Stud. 2006;43(1):38-44.
20. Rychlowska E, Fijalkowski B, Chlebna-Sokól D. Assessment of exercise capacity in children aged 9-13 years from Lodz. Przeglad Pediatryczny. 2012;42(3):141-8.
21. van der Cammen-van Zijp MH, van den Berg-Emons RJ, Willemsen SP, Stam HJ, Tibboel D, H. I. Exercise capacity in Dutch children: new reference values for the Bruce treadmill protocol. Scand J Med Sci Sports. 2010;20(1):e130-e6. http://doi.org/10.1111/j.600-0838.2009.00925.x.
22. Baldi JC, Aoina JL, Oxenham HC, Bagg W, Doughty RN. Reduced exercise arteriovenous O_2 difference in type 2 diabetes. J Appl Physiol. 2003;94(3):1033-8.
23. Ganju AA, Fuladi AB, Tayade BO, Ganju NA. Cardiopulmonary exercise testing in evaluation of patients of chronic obstructive pulmonary disease. Indian J Chest Dis Allied Sci. 2011;53(2):87-91.
24. Duff DK, De Souza AM, Human DG, Potts JE, Harris KC. A novel treadmill protocol for exercise testing in children: the British Columbia Children's Hospital protocol. BMJ Open Sport Exerc Med. 2017;3(1):e000197. http://doi.org/10.1136/bmjsem-2016.
25. Delisle AT, Piazza-Gardner AK, Cowen TL, Huq MBS, Delisle AD, Stopka CB, et al. Validation of a cardiorespiratory fitness assessment for firefighters. J Strength Cond Res. 2014;28(10):2717-23.
26. Kotte EMW, De Groot JF, Bongers BC, Winkler AMF, Takken TIM. Validity and reproducibility of a new treadmill protocol: the Fitkids Treadmill Test. Medicine and science in sports and exercise. 2015;47(10):2241-7.
27. Hanson NJ, Scheadler CM, Lee TL, Neuenfeldt NC, Michael TJ, Miller MG. Modality determines $\dot{V}O_{2max}$ achieved in self-paced exercise tests: validation with the Bruce protocol. Eur J Appl Physiol. 2016;116(7):1313-9.
28. Hamlin MJ, Draper N, Blackwell G, Shearman JP, Kimber NE. Determination of maximal oxygen uptake using the Bruce or a novel athlete-led protocol in a mixed population. J Hum Kinet. 2012;31:97-104.
29. Bartlett JW, Frost C. Reliability, repeatability and reproducibility: analysis of measurement errors in continuous variables. Ultrasound Obstet Gynecol. 2008;31(4):466-75.
30. Noonan V, Dean E. Submaximal exercise testing: clinical application and interpretation. Phys Ther. 2000;80(8):782-807.
31. Lee J-M, Bassett Jr DR, Thompson DL, Fitzhugh EC. Validation of the Cosmed Fitmate for prediction of maximal oxygen consumption. J Strength Cond Res. 2011;25(9):2573-9.
32. George J, Paul S, Hyde A, Bradshaw D, Vehrs P, Hager R, et al. Prediction of maximum oxygen uptake using both exercise and non-exercise data. Meas Phys Educ Exerc Sci. 2009;13(1):1-12.
33. Ekblom-Bak E, Björkman F, Hellenius M-L, Ekblom B. A new submaximal cycle ergometer test for prediction of $\dot{V}O_{2max}$. Scand J Med Sci Sports. 2014;24(2):319-26.
34. Shenoy S, Tyagi BS, Sandhu JS. Concurrent validity of the non-exercise based $\dot{V}O_{2max}$ prediction equation using percentage body fat as a variable in asian Indian adults. Sports Medicine, Arthroscopy, Rehabilitation, Therapy & Technology : SMARTT. 2012;4(1):34. http://doi.org/10.1186/758-2555-4-34.
35. Milani J, Fernhall B, Manfredi T. Estimating oxygen consumption during treadmill and arm ergometry activity in males with coronary artery disease. J Cardiopulm Rehabil. 1996;16(6):394-401.

36. Aguiar PF, Moriarty TA, Baracho WA, Paula FD, Sampaio PFM, Ottone VO, et al. The accuracy of two equations for predicting maximal oxygen uptake on individualized ramp protocol. Human Movement. 2018;19(4):42-8.
37. Taylor HL, Buskirk E, Henschel A. Maximal oxygen intake as an objective measure of cardio-respiratory performance. J Appl Physiol. 1955;8(1):73-80.
38. Astorino TA, Willey J, Kinnahan J, Larsson SM, Welch H, Dalleck LC. Elucidating determinants of the plateau in oxygen consumption at $\dot{V}O_{2max}$. Br J Sports Med. 2005;39(9):655-60.
39. Poole DC, Wilkerson DP, Jones AM. Validity of criteria for establishing maximal O_2 uptake during ramp exercise tests. Eur J Appl Physiol. 2008;102(4):403-10.
40. Mier CM, Alexander RP, Mageean AL. Achievement of $\dot{V}O_{2max}$ criteria during a continuous graded exercise test and a verification stage performed by college athletes. J Strength Cond Res. 2012;26(10):2648-54.
41. Riebe D, Ehrman JK, Liguori G, Magal M, American College of Sports Medicine. ACSM's Guidelines for Exercise Testing and Prescription. 10th ed. Riebe D, editor. Philadelphia: Wolters Kluwer; 2018. 472 p.
42. Midgley AW, McNaughton LR, Polman R, Marchant D. Criteria for determination of maximal oxygen uptake: a brief critique and recommendations for future research. Sports Med. 2007;37(12):1019-28.
43. Edvardsen E, Hem E, Anderssen SA. End criteria for reaching maximal oxygen uptake must be strict and adjusted to sex and age: a cross-sectional study. PLOS ONE. 2014;9(1):e85276. http://doi.org/10.1371/journal.pone.0085276.
44. Balady GJ, Arena R, Sietsema K, Myers J, Coke L, Fletcher GF, et al. Clinician's guide to cardiopulmonary exercise testing in adults: a scientific statement from the American Heart Association. Circulation. 2010;122(2):191-225.
45. Badawy MM, Muaidi QI. Cardio respiratory response: Validation of new modifications of Bruce protocol for exercise testing and training in elite Saudi triathlon and soccer players. Saudi J Biol Sci. 2019;26(1):105-11.
46. Marshall MR, Coe DP, Pivarnik JM. Development of a prediction model to predict $\dot{V}O_{2peak}$ in adolescent girls using the Bruce Protocol to exhaustion. Res Q Exerc Sport. 2014;85(2):251-6.
47. Pols MA, Peeters PH, Bueno-De-Mesquita HB, Ocké MC, Wentink CA, Kemper HC, et al. Validity and repeatability of a modified Baecke questionnaire on physical activity. Int J Epidemiol. 1995;24(2):381-8.
48. Reinstein DZ, Archer TJ, Silverman RH, Coleman DJ. Accuracy, repeatability, and reproducibility of Artemis very high-frequency digital ultrasound arc-scan lateral dimension measurements. Journal of cataract and refractive surgery. 2006;32(11):1799-802.
49. Pountney T, Mandy A, Gard P. Repeatability and limits of agreement in measurement of hip migration percentage in children with bilateral cerebral palsy. Physiotherapy. 2003;89(5):276-81.
50. Moher D, Liberati A, Tetzlaff J, Altman DG, Group TP. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med. 2009;6(7):e1000097. http://doi.org/10.1371/journal.pmed.
51. Fletcher GF, Ades PA, Kligfield P, Arena R, Balady GJ, Bittner VA, et al. Exercise standards for testing and training: A scientific statement from the American Heart Association. Circulation. 2013;128(8):873-934.
52. Downes MJ, Brennan ML, Williams HC, Dean RS. Development of a critical appraisal tool to assess the quality of cross-sectional studies (AXIS). BMJ Open. 2016;6(12):e011458. http://doi.org/10.1136/bmjopen-2016.

53. Profant GR, Early RG, Nilson KL, Kusumi F, Hofer V, Bruce RA. Responses to maximal exercise in healthy middle-aged women. J Appl Physiol. 1972;33(5):595-9.
54. Froelicher Jr VF, Brammell H, Davis G. A comparison of the reproducibility and physiologic response to three maximal treadmill exercise protocols. CHEST. 1974;65(5):512-7.
55. Ho BL. A study of treadmill exercise protocols for Chinese males. Aviat Space Environ Med. 1982;53(2):112-6.
56. Nordrehaug JE, Danielsen R, Stangeland L, Rosland GA, Vikmo H. Respiratory gas-exchange during treadmill exercise testing - reproducibility and comparison of different exercise protocols. Scand J Clin Lab Invest. 1991;51(7):655-8.
57. Fielding RA, Frontera WR, Hughes VA, Fisher EC, Evans WJ. The reproducibility of the Bruce protocol exercise test for the determination of aerobic capacity in older women. Medicine and science in sports and exercise. 1997;29(8):1109-13.
58. Cooke GA, Marshall P, al-Timman JK, Wright DJ, Riley R, Hainsworth R, et al. Physiological cardiac reserve: development of a non-invasive method and first estimates in man. Heart (British Cardiac Society). 1998;79(3):289-94.
59. Jakovljevic DG, Seferovic PM, Nunan D, Donovan G, Trenell MI, Grocott-Mason R, et al. Reproducibility of cardiac power output and other cardiopulmonary exercise indices in patients with chronic heart failure. Clin Sci (Lond). 2012;122(4):175-81.
60. Hall-López JA, Ochoa-Martínez PY, Moncada-Jiménez J, Ocampo Méndez MA, Martínez García I, Martínez García MA. Reliability of the maximal oxygen uptake following two consecutive trials by indirect calorimetry. Nutr Hosp. 2015;31(4):1726-32.
61. Harwood AE, Pymer S, Hitchman L, Totty J, Wallace T, Smith GE, et al. The intrarater and interrater reliability of measures derived from cardiopulmonary exercise testing in patients with abdominal aortic aneurysms. Ann Vasc Surg. 2019;56:175-82.
62. Faber J, Fonseca LM. How sample size influences research outcomes. Dental press journal of orthodontics. 2014;19(4):27-9.
63. Button KS, Ioannidis JPA, Mokrysz C, Nosek BA, Flint J, Robinson ESJ, et al. Power failure: why small sample size undermines the reliability of neuroscience. Nat Rev Neurosci. 2013;14(5):365-76.
64. Biau DJ, Kernéis S, Porcher R. Statistics in brief: the importance of sample size in the planning and interpretation of medical research. Clinical Orthopaedics and Related Research. 2008;466(9):2282-8.
65. Brandenburg SL, Reusch JE, Bauer TA, Jeffers BW, Hiatt WR, Regensteiner JG. Effects of exercise training on oxygen uptake kinetic responses in women with type 2 diabetes. Diabetes Care. 1999;22(10):1640-6.
66. Helgerud J, Høydal K, Wang E, Karlsen T, Berg P, Bjerkaas M, et al. Aerobic high-intensity intervals improve $\dot{V}O_{2max}$ more than moderate training. Medicine and science in sports and exercise. 2007;39(4):665-71.
67. Swank AM, Horton J, Fleg JL, Fonarow GC, Keteyian S, Goldberg L, et al. Modest increase in peak $\dot{V}O_2$ is related to better clinical outcomes in chronic heart failure patients: results from heart failure and a controlled trial to investigate outcomes of exercise training. Circ Heart Fail. 2012;5(5):579-85.
68. Bocalini DS, Serra AJ, Rica RL, Santos Ld. Repercussions of training and detraining by water-based exercise on functional fitness and quality of life: a short-term follow-up in healthy older women. Clinics. 2010;65(12):1305-9.
69. Neufer PD. The effect of detraining and reduced training on the physiological adaptations to aerobic exercise training. Sports Med. 1989;8(5):302-20.
70. Nolan PB, Keeling SM, Robitaille CA, Buchanan CA, Dalleck LC. The effect of detraining after a period of training on cardiometabolic health in previously sedentary individuals.

Int J Environ Res Public Health. 2018;15(10):2303. http://doi.org/10.3390/ijerph15102303.

71. Manfre MJ, Yu GH, Varma AA, Mallis GI, Kearney K, Karageorgis MA. The effect of limited handrail support on total treadmill time and the prediction of $\dot{V}O_{2max}$. Clinical Cardiology. 1994;17(8):445-50.
72. Berling J, Foster C, Gibson M, Doberstein S, Porcari J. The effect of handrail support on oxygen uptake during steady-state treadmill exercise. J Cardiopulm Rehabil Prev. 2006;26(6):391-4.
73. Smokler PE, MacAlpin RN, Alvaro A, Kattus AA. Reproducibility of a multi-stage near maximal treadmill test for exercise tolerance in angina pectoris. Circulation. 1973;48(2):346-51.
74. van der Cammen-van Zijp MHM, IJsselstijn H, Takken T, Willemsen SP, Tibboel D, Stam HJ, et al. Exercise testing of pre-school children using the Bruce treadmill protocol: new reference values. Eur J Appl Physiol. 2010;108(2):393-9.
75. Swinburn CR, Wakefield JM, Jones PW. Performance, ventilation, and oxygen consumption in three different types of exercise test in patients with chronic obstructive lung disease. Thorax. 1985;40(8):581-6.
76. Knox AJ, Morrison JF, Muers MF. Reproducibility of walking test results in chronic obstructive airways disease. Thorax. 1988;43(5):388-92.
77. Kraemer MD, Sullivan M, Atwood E, Forbes S, Myers J, Froelicher V. Reproducibility of treadmill exercise data in patients with atrial fibrillation. Cardiology. 1989;76(3):234-42.
78. Tonino RP, Driscoll PA. Reliability of maximal and submaximal parameters of treadmill testing for the measurement of physical training in older persons. J Gerontol. 1988;43(4): M101-M4.
79. Cox NJM, Hendriks JCM, Binkhorst RA, Folgering HTM, van Herwaarden CLA. Reproducibility of incremental maximal cycle ergometer tests in patients with mild to moderate obstructive lung diseases. Lung. 1989;167(1):129-33.
80. Barron A, Dhutia N, Mayet J, Hughes AD, Francis DP, Wensel R. Test–retest repeatability of cardiopulmonary exercise test variables in patients with cardiac or respiratory disease. Eur J Prev Cardiol. 2014;21(4):445-53.
81. Eston R, Stansfield R, Westoby P, Parfitt G. Effect of deception and expected exercise duration on psychological and physiological variables during treadmill running and cycling. Psychophysiology. 2012;49(4):462-9.
82. Baden DA, McLean TL, Tucker R, Noakes TD, Gibson ASC. Effect of anticipation during unknown or unexpected exercise duration on rating of perceived exertion, affect, and physiological function. Br J Sports Med. 2005;39(10):742-6.
83. Taylor R. Interpretation of the correlation coefficient: a basic review. J Diagn Med Sonogr. 1990;6(1):35-9.
84. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet. 1986;327(8476):307-10.
85. Riquelme M, Leiva V, Galea M, Sanhueza A. Influence diagnostics on the coefficient of variation of elliptically contoured distributions. J Appl Stat. 2011;38(3):513-32.
86. Shoukri MM, Colak D, Kaya N, Donner A. Comparison of two dependent within subject coefficients of variation to evaluate the reproducibility of measurement devices. BMC Med Res Methodol. 2008;8(1):24. http://doi.org/10.1186/471-2288-8-24.
87. Bland JM, Altman DG. Measurement error. Br Med J. 1996;312(7047):1654-4.
88. Santtila M, Häkkinen K, Pihlainen K, Kyröläinen H. Comparison between direct and predicted maximal oxygen uptake measurement during cycling. Mil Med. 2013;178(2):234-8.

89. Vehrs PR, George JD, Fellingham GW, Plowman SA, Dustman-Allen K. Submaximal treadmill exercise test to predict $\dot{V}O_{2max}$ in fit adults. Meas Phys Educ Exerc Sci. 2007;11(2):61-72.
90. Nordgren B, Fridén C, Jansson E, Österlund T, Grooten WJ, Opava CH, et al. Criterion validation of two submaximal aerobic fitness tests, the self-monitoring Fox-walk test and the Åstrand cycle test in people with rheumatoid arthritis. BMC Musculoskelet Disord. 2014;15(305):http://doi.org/10.1186/471-2474-15-305.
91. Keytel LR, Goedecke JH, Noakes TD, Hiiloskorpi H, Laukkanen R, van der Merwe L, et al. Prediction of energy expenditure from heart rate monitoring during submaximal exercise. J Sports Sci. 2005;23(3):289-97.
92. Sullivan K, Shikuma CM, Chow D, Cornelius E, Romine RK, Lindsey RA, et al. Aerobic fitness levels and validation of a non exercise $\dot{V}O_{2max}$ prediction equation for HIV-infected patients on HAART. HIV Clinical Trials. 2014;15(2):69-77.
93. Duque IL, Parra J-H, Duvallet A. A new non exercise-based $\dot{V}O_{2max}$ prediction equation for patients with chronic low back pain. J Occup Rehabil. 2009;19(3):293-9.
94. Jackson AS, Blair SN, Mahar MT, Wier LT, Ross RM, Stuteville JE. Prediction of functional aerobic capacity without exercise testing. Medicine and science in sports and exercise. 1990;22(6):863-70.
95. Foster C, Hare J, Taylor MM. Prediction of oxygen uptake during exercise testing in cardiac patients and healthy volunteers. Journal of Cardiac Rehabilitation. 1984;4(12):537-42.
96. Roy BA, Grove MA, Christie LG. A model for estimating oxygen uptake in patients who have undergone coronary artery bypass grafts. J Cardiopulm Rehabil. 1992;12(2):111-6.
97. Myers J, Kaminsky LA, Lima R, Christle JW, Ashley E, Arena R. A reference equation for normal standards for $\dot{V}O_{2max}$: Analysis from the Fitness Registry and the Importance of Exercise National Database (FRIEND Registry). Prog Cardiovasc Dis. 2017;60(1):21-9.
98. Evans HJ, Ferrar KE, Smith AE, Parfitt G, Eston RG. A systematic review of methods to predict maximal oxygen uptake from submaximal, open circuit spirometry in healthy adults. J Sci Med Sport. 2015;18(2):183-8.
99. Schembre SM, Riebe DA. Non-exercise estimation of $\dot{V}O_{2max}$ using the International Physical Activity Questionnaire. Meas Phys Educ Exerc Sci. 2011;15(3):168-81.
100. Jackson AS, Beard EF, Wier LT, Stuteville JE. Multivariate model for defining changes in maximal physical working capacity of men, ages 25 to 70 years. Proc Hum Factors Ergon Soc Annu Meet. 1992;36(2):171-4.
101. Grant JA, Joseph AN, Campagna PD. The prediction of $\dot{V}O_{2max}$: a comparison of 7 indirect tests of aerobic power. J Strength Cond Res. 1999;13(4):346-52.
102. Maeder M, Wolber T, Atefy R, Gadza M, Ammann P, Myers J, et al. A nomogram to select the optimal treadmill ramp protocol in subjects with high exercise capacity: Validation and comparison with the Bruce protocol. J Cardiopulm Rehabil. 2006;26(1):16-23.
103. Koutlianos N, Dimitros E, Metaxas T, Deligiannis AS, Kouidi E. Indirect estimation of $\dot{V}O_{2max}$ in athletes by ACSM's equation: valid or not? Hippokratia. 2013;17(2):136-40.
104. Nitin Y, Sucharita S, Madhura M, Thomas T, Sandhya TA. $\dot{V}O_{2max}$ in an Indian population: A study to understand the role of factors determining $\dot{V}O_{2max}$. Indian J Physiol Pharmacol. 2013;57(2):87-94.
105. Crouse SF, Tolson H, Lytle J, Johnson KA, Martin SE, Green JS, et al. Predicting $\dot{V}O_{2max}$ from treadmill performance in American-style football athletes. J Strength Cond Res. 2019;33(4):1028-34.
106. Nieman DC, Austin MD, Dew D, Utter AC. Validity of COSMED's quark CPET mixing chamber system in evaluating energy metabolism during aerobic exercise in healthy male adults. Res Sports Med. 2013;21(2):136-45.

107. Poole G. Professional Fitness and Lifestyle Consultant Resource Manual. Gloucester, Ontario: Canadian Society for Exercise Physiology; 1996.
108. Heyward VH. Designs for fitness: A guide to physical fitness appraisal and exercise prescription. Minneaapolis: Burgess Pub. Co.; 1984.
109. Whaley MH, Brubaker PH, Otto RM, Armstrong LE, American College of Sports Medicine. ACSM's Guidelines for Exercise Testing and Prescription. Baltimore: Lippincott Williams & Wilkins; 2005.
110. Bandyopadhyay A, Chatterjee S. Body composition, morphological characteristics and their relationship with cardiorespiratory fitness. Ergonomics SA. 2003;15:19-27.
111. Chatterjee S, Mitra SK, Samanta A. Aerobic capacity of the brick-field workers in eastern India. Ind Health. 1994;32(2):79-84.
112. Biswas R, Samanta A, Chatterjee S. Maximal aerobic capacity of Indian inland fishermen. Indian Journal of Physiology & Allied Sciences. 2004;58(3):70-9.
113. Foster C, Jackson AS, Pollock ML, Taylor MM, Hare J, Sennett SM, et al. Generalized equations for predicting functional capacity from treadmill performance. Am Heart J. 1984;107(6):1229-34.
114. Page MJ, McKenzie JE, Higgins JPT. Tools for assessing risk of reporting biases in studies and syntheses of studies: a systematic review. BMJ Open. 2018;8(3):e019703-e. http://doi.org/10.1136/bmjopen-2017.
115. Bird SR. Research Methods in Physical Activity and Health. 1st ed. Abingdon, Oxon.: Routledge 2019.
116. Martinson BC, Crain AL, Sherwood NE, Hayes MG, Pronk NP, O'Connor PJ. Population reach and recruitment bias in a maintenance RCT in physically active older adults. Journal of Physical Activity & Health. 2010;7(1):127-35.
117. Andreacci JL, LeMura LM, Cohen SL, Urbansky EA, Chelland SA, Von Duvillard SP. The effects of frequency of encouragement on performance during maximal exercise testing. J Sports Sci. 2002;20(4):345-52.
118. Akoglu H. User's guide to correlation coefficients. Turkish Journal of Emergency Medicine. 2018;18(3):91-3.
119. Ratner B. The correlation coefficient: Its values range between +1/−1, or do they? Journal of Targeting, Measurement and Analysis for Marketing. 2009;17(2):139-42.
120. Siconolfi SF, McConnell TR. Estimating $\dot{V}O_{2max}$ and $\dot{V}O_{2submax}$ in beta blocked patients during handrail-supported treadmill exercise. J Cardiopulm Rehabil. 1990;10(1):21-4.
121. Tapking C, Popp D, Herndon DN, Branski LK, Mlcak RP, Suman OE. Estimated versus achieved maximal oxygen consumption in severely burned children maximal oxygen consumption in burned children. Burns. 2018;44(8):2026-33.
122. de Souza e Silva CG, Franklin BA, de Araújo CGS. Influence of central obesity in estimating maximal oxygen uptake. Clinics (Sao Paulo, Brazil). 2016;71(11):629-34.
123. Scribbans TD, Vecsey S, Hankinson PB, Foster WS, Gurd BJ. The effect of training intensity on $\dot{V}O_{2max}$ in young healthy adults: A meta-regression and meta-analysis. International jjournal of eexercise sscience. 2016;9(2):230-47.
124. Mundwiler J, Schüpbach U, Dieterle T, Leuppi JD, Schmidt-Trucksäss A, Wolfer DP, et al. Association of occupational and leisure-time physical activity with aerobic capacity in a working population. PLOS ONE. 2017;12(1):e0168683. http://doi.org/10.1371/journal.pone.
125. Wang Y, Chen S, Lavie CJ, Zhang J, Sui X. An overview of non-exercise estimated cardiorespiratory fitness: estimation equations, cross-validation and application. Journal of Science in Sport and Exercise. 2019;1(1):38-53.
126. Bradshaw DI, George JD, Hyde A, LaMonte MJ, Vehrs PR, Hager RL, et al. An accurate $\dot{V}O_{2max}$ nonexercise regression model for 18–65-year-old adults. Res Q Exerc Sport. 2005;76(4):426-32.

127. Dyrstad S, Anderssen S, Edvardsen E, Hansen B. Cardiorespiratory fitness in groups with different physical activity levels. Scand J Med Sci Sports. 2015;26(3):291-8.
128. Malek MH, Housh TJ, Berger DE, Coburn JW, Beck TW. A new non-exercise-based $\dot{V}O_{2max}$ prediction equation for aerobically trained men. J Strength Cond Res. 2005;19(3):559-65.
129. Abut F, Akay MF, George J. Developing new $\dot{V}O_{2max}$ prediction models from maximal, submaximal and questionnaire variables using support vector machines combined with feature selection. Comput Biol Med. 2016;79:182-92.
130. Loe H, Nes BM, Wisløff U. Predicting $\dot{V}O_{2peak}$ from submaximal- and peak exercise models: the HUNT 3 Fitness Study, Norway. PLOS ONE. 2016;11(1):e0144873. http://doi.org/10.1371/journal.pone.
131. Akay MF, Zayid EIM, Aktürk E, George JD. Artificial neural network-based model for predicting $\dot{V}O_{2max}$ from a submaximal exercise test. Expert Syst Appl. 2011;38(3):2007-10.
132. Sloan RA, Haaland BA, Leung C, Padmanabhan U, Koh HC, Zee A. Cross-validation of a non-exercise measure for cardiorespiratory fitness in Singaporean adults. Singapore Med J. 2013;54(10):576-80.
133. Pollock RD, Carter S, Velloso CP, Duggal NA, Lord JM, Lazarus NR, et al. An investigation into the relationship between age and physiological function in highly active older adults. J Physiol. 2015;593(3):657-80.
134. Das B, Ghosh T, Gangopadhyay S. A comparative study of Physical Fitness Index (PFI) and predicted maximum aerobic capacity ($\dot{V}O_{2max}$) among the different groups of female students in West Bengal, India. International Journal of Sports and Health Sciences (Korea). 2010;22(1):13-23.
135. Silvestre R, West C, Maresh CM, Kraemer WJ. Body composition and physical performance in men's soccer: a study of a National Collegiate Athletic Association Division I team. J Strength Cond Res. 2006;20(1):177-83.
136. Wier LT, Jackson AS, Ayers GW, Arenare B. Nonexercise models for estimating $\dot{V}O_{2max}$ with waist girth, percent fat, or BMI. Medicine and science in sports and exercise. 2006;38(3):555-61.
137. Giavarina D. Understanding Bland Altman analysis. Biochemia Medica. 2015;25(2):141-51.
138. van Stralen KJ, Jager KJ, Zoccali C, Dekker FW. Agreement between methods. Kidney Int. 2008;74(9):1116-20.
139. Li J. Assessing the accuracy of predictive models for numerical data: Not r nor r^2, why not? Then what? PLOS ONE. 2017;12(8):e0183250-e. http://doi.org/10.1371/journal.pone.
140. Altman DG, Bland JM. Standard deviations and standard errors. British Medical Journal. 2005;331(7521):903-3.
141. Kelley K. Sample size planning for the coefficient of variation from the accuracy in parameter estimation approach. Behav Res Methods. 2007;39(4):755-66.
142. da Costa AV, da Cunha Costa M, de Oliveira SFM, de Albuquerque FL, de Sá Pereira Guimarães FJ, Barbosa TM. Validation of an equation for estimating maximal oxygen consumption of nonexpert adult swimmers. Open access journal of sports medicine. 2013;4:19-25.
143. Katch VL, McArdle WD, Katch FI. Essentials of Exercise Physiology. 4th Edition ed. Katch VL, editor. Baltimore: Lippincott Williams & Wilkins; 2011 Sat Jan 01 00:00:00 GMT 2011. p. 228 p.
144. Rankovic G, Mutavdzic V, Toskic D, Preljevic A, Kocic M, Nedin Rankovic G, et al. Aerobic capacity as an indicator in different kinds of sports. Bosnian Journal of Basic Medical Sciences. 2010;10(1):44-8.

145. Peterson MJ, Pieper CF, Morey MC. Accuracy of $\dot{V}O_{2max}$ prediction equations in older adults. Medicine and science in sports and exercise. 2003;35(1):145-9.
146. Storer TW, Davis JA, Caiozzo VJ. Accurate prediction of $\dot{V}O_{2max}$ in cycle ergometry. Medicine and science in sports and exercise. 1990;22(5):704-12.
147. Pinkstaff S, Peberdy MA, Kontos MC, Fabiato A, Finucane S, Arena R. Overestimation of aerobic capacity with the bruce treadmill protocol in patients being assessed for suspected myocardial Ischemia. J Cardiopulm Rehabil Prev. 2011;31(4):254-60.
148. Macsween A. The reliability and validity of the Åstrand nomogram and linear extrapolation for deriving $\dot{V}O_{2max}$ from submaximal exercise data. J Sports Med Phys Fitness. 2001;41(3):312-7.
149. Miller GS, Dougherty PJ, Green JS, Crouse SF. Comparison of cardiorespiratory responses of moderately trained men and women using two different treadmill protocols. J Strength Cond Res. 2007;21(4):1067-71.
150. Shah BN. On the 50th anniversary of the first description of a multistage exercise treadmill test: Re-visiting the birth of the 'Bruce protocol'. Heart (British Cardiac Society). 2013;99(24):1793-4.
151. Fletcher GF, Balady GJ, Amsterdam EA, Chaitman B, Eckel R, Fleg J, et al. Exercise standards for testing and training: A statement for healthcare professionals from the American Heart Association. Circulation. 2001;104(14):1694-740.
152. Gappmaier E. The submaximal clinical exercise tolerance test (SXTT) to establish safe exercise prescription parameters for patients with chronic disease and disability. Cardiopulmonary physical therapy journal. 2012;23(2):19-29.
153. Pescatello LS, Arena R, Riebe D, Thompson PD, American College of Sports Medicine. ACSM's Guidelines for Exercise Testing and Prescription. 9th ed. Pescatello LS, editor. Philadelphia: Wolters Kluwer/Lippincott Williams & Wilkins; 2014.
154. Tanaka H, Monahan KD, Seals DR. Age-predicted maximal heart rate revisited. J Am Coll Cardiol. 2001;37(1):153-6.
155. Heyward VH, Gibson AL. Advanced Fitness Assessment and Exercise Prescription. 7th ed. Champaign, IL: Human Kinetics; 2014.
156. Gisev N, Bell JS, Chen TF. Interrater agreement and interrater reliability: Key concepts, approaches, and applications. Res Social Adm Pharm. 2013;9(3):330-8.
157. Bunce C. Correlation, agreement, and Bland-Altman analysis: Statistical analysis of method comparison studies. Am J Ophthalmol. 2009;148(1):4-6.
158. Williams CJ, Gurd BJ, Bonafiglia JT, Voisin S, Li Z, Harvey N, et al. A multi-center comparison of O_{2peak} trainability between interval training and moderate intensity continuous training. Front Physiol. 2019;10(19):http://doi.org/10.3389/fphys.2019.00019.
159. Williamson PJ, Atkinson G, Batterham AM. Inter-individual responses of maximal oxygen uptake to exercise training: a critical review. Sports Med. 2017;47(8):1501-13.
160. Dias Neto JM, Silva FB, de Oliveira ALB, Couto NL, Dantas EHM, Nascimento MAdL. Effects of verbal encouragement on performance of the multistage 20 m shuttle run. Acta Scientiarum Health Science. 2015;37(1):25-30.
161. Halperin I, Pyne D, Martin D. Threats to internal validity in exercise science: A review of overlooked confounding variables. Int J Sports Physiol Perform. 2015;10(7):823-9.
162. Beekley MD, Brechue WF, deHoyos DV, Garzarella L, Werber-Zion G, Pollock ML. Croos-validation of the YMCA submaximal cycle ergometer test to predict $\dot{V}O_{2max}$. Res Q Exerc Sport. 2004;75(3):337-42.
163. Stokes M. Reliability and repeatability of methods for measuring muscle in physiotherapy. Physiotherapy Practice. 1985;1(2):71-6.

164. Vaz S, Falkmer T, Passmore AE, Parsons R, Andreou P. The case for using the repeatability coefficient when calculating test–retest reliability. PLOS ONE. 2013;8(9):e73990. http://doi.org/10.1371/journal.pone.0073990.
165. McNulty C, Robergs RA. Repeat trial and breath averaging: Recommendations for research of $\dot{V}O_2$ kinetics of exercise transitions to steady-state. Movement & Sport Sciences. 2019;106:37-44.
166. Robergs RA, Dwyer D, Astorino T. Recommendations for improved data processing from expired gas analysis indirect calorimetry. Sports Med. 2010;40(2):95-111.
167. Mezzani A. Cardiopulmonary eercise testing: Basics of methodology and measurements. Ann Am Thorac Soc. 2017;14(Supplement 1):S3-S11.
168. Hill DW, Stephens LP, Blumoff-Ross SA, Poole DC, Smith JC. Effect of sampling strategy on measures of $\dot{V}O_{2peak}$ obtained using commercial breath-by-breath systems. Eur J Appl Physiol. 2003;89(6):564-9.
169. Robergs R, Burnett A. Methods used to process data from indirect calorimetry and their application to $\dot{V}O_{2max}$. J Exerc Physiol Online. 2003;6(2):44-57.
170. Scheadler CM, Garver MJ, Hanson NJ. The gas sampling interval effect on $\dot{V}O_{2peak}$ is independent of exercise protocol. Medicine and science in sports and exercise. 2017;49(9):1911-6.
171. CSEP Expert Advisory Committee. PAR-Q and You: Physical Activity Readiness Questionnaire. Ottawa: Ontario: Canadian Society of Exercise Physiology; Revised 2002.
172. Craig CL, Marshall AL, Sjöström M, Bauman AE, Booth ML, Ainsworth BE, et al. International physical activity questionnaire: 12-country reliability and validity. Medicine and science in sports and exercise. 2003;35(8):1381-95.
173. Midgley A, Earle K, McNaughton L, Siegler J, Clough P, Earle F. Exercise tolerance during $\dot{V}O_{2max}$ testing is a multifactorial psychobiological phenomenon. Res Sports Med. 2017;25(4):480-94.
174. James DV, Sandals LE, Draper SB, Wood DM. Relationship between maximal oxygen uptake and oxygen uptake attained during treadmill middle-distance running. J Sports Sci. 2007;25(8):851-8.
175. Midgley AW, McNaughton LR, Carroll S. Effect of the $\dot{V}O_2$ time-averaging interval on the reproducibility of $\dot{V}O_{2max}$ in healthy athletic subjects. Clin Physiol Funct Imaging. 2007;27(2):122-5.
176. Astorino TA, Allen RP, Roberson DW, Jurancich M. Effect of high-intensity interval training on cardiovascular function, $\dot{V}O_{2max}$, and muscular force. J Strength Cond Res. 2012;26(1):138-45.
177. Astorino TA, Edmunds RM, Clark A, King L, Gallant RA, Namm S, et al. High-intensity interval training increases cardiac output and $\dot{V}O_{2max}$. Medicine and science in sports and exercise. 2017;49(2):265-73.
178. Wen D, Utesch T, Wu J, Robertson S, Liu J, Hu G, et al. Effects of different protocols of high intensity interval training for $\dot{V}O_{2max}$ improvements in adults: A meta-analysis of randomised controlled trials. J Sci Med Sport. 2019;22(8):941-7.
179. Esco M, Flatt A, Nakamura F. Initial weekly HRV response is related to the prospective change in $\dot{V}O_{2max}$ in female soccer players. Int J Sports Med. 2016;37(6):436-41.
180. Ferdowsi MH, Saiiari A, Valizadeh R, Gholamie A. The effect of eight week aerobic exercise on airway trachea indexes (FEV1, FVC, FEV1. FVC & FEF25-75) and $\dot{V}O_{2max}$ level in overweighed male students of Ahvaz Payam Noor University. Procedia Soc Behav Sci. 2011;15:2848-52.
181. Balaji P, Rajkumar D, Karthikeyan P. Impact of core stability exercises on vital capacity and $\dot{V}O_{2max}$ of sedentary men. International Journal of Physical Education, Sports and Health. 2016;3(6):122-4.

182. Joubert D, Oden G, Estes B. The effects of ellipical cross training on $\dot{V}O_{2max}$ in recently trained runners. International jjournal of eexercise sscience. 2011;4(1):4-12.
183. Elborn JS, Stanford CF, Nicholls DP. Reproducibility of cardiopulmonary parameters during exercise in patients with chronic cardiac failure. The need for a preliminary test. Eur Heart J. 1990;11(1):75-81.
184. Gosselink R, Troosters T, Decramer M. Exercise testing: why, which and how to interpret. Breathe. 2004;1(2):120-9.
185. Kang J, Chaloupka EC, Mastrangelo MA, Biren GB, Robertson RJ. Physiological comparisons among three maximal treadmill exercise protocols in trained and untrained individuals. Eur J Appl Physiol. 2001;84(4):291-5.
186. Hevey D, Brown A, Cahill A, Newton H, Kierns M, Horgan JH. Four-week multidisciplinary cardiac rehabilitation produces similar improvements in exercise capacity and quality of life to a 10-week program. J Cardiopulm Rehabil. 2003;23(1):17-21.
187. Adekunle AE, Akintomide AO. Gender differences in the variables of exercise treadmill test in type 2 diabetes mellitus. Ann Afr Med. 2012;11(2):96-102.
188. Gürgün A, Ekren P, Karapolat H, Tuncel Ş. Pulmonary rehabilitation response in elderly and younger patients with chronic obstructive pulmonary disease. Turkish Journal of Geriatrics. 2013;16(4):427-33.
189. Sunitha G, Ravi BN, Sudhir GK. Study of $\dot{V}O_{2max}$ during phases of menstruation in young female athletes. J Evol Med Dent Sci. 2013;2(23):4070-8.
190. Gordon D, Scruton A, Barnes R, Baker J, Prado L, Merzbach V. The effects of menstrual cycle phase on the incidence of plateau at $\dot{V}O_{2max}$ and associated cardiorespiratory dynamics. Clin Physiol Funct Imaging. 2018;38(4):689-98.
191. Janse de Jonge XAK. Effects of the menstrual cycle on exercise performance. Sports Med. 2003;33(11):833-51.
192. Casazza GA, Suh SH, Miller BF, Navazio FM, Brooks GA. Effects of oral contraceptives on peak exercise capacity. J Appl Physiol. 2002;93(5):1698-702.
193. Lebrun CM, Petit MA, McKenzie DC, Taunton JE, Prior JC. Decreased maximal aerobic capacity with use of a triphasic oral contraceptive in highly active women: a randomised controlled trial. Br J Sports Med. 2003;37(4):315-20.
194. Packard KA, Lenz TL, Elder B, Godfrey C, Holcomb R, Windle E. Oral contraceptive use may attenuate menstrual cycle-induced ventilatory changes in endurance trained runners. The Open Sports Medicine Journal. 2011;5:5-11.
195. Williams TJ, Krahenbuhl GS. Menstrual cycle phase and running economy. Medicine and science in sports and exercise. 1997;29(12):1609-18.
196. Forsyth JJ, Reilly T. The combined effect of time of day and menstrual cycle on lactate threshold. Medicine and science in sports and exercise. 2005;37(12):2046-53.
197. McCracken M, Ainsworth B, Hackney AC. Effects of the menstrual cycle phase on the blood lactate responses to exercise. Eur J Appl Physiol Occup Physiol. 1994;69(2):174-5.
198. Redman LM, Scroop GC, Norman RJ. Impact of menstrual cycle phase on the exercise status of young, sedentary women. Eur J Appl Physiol. 2003;90(5-6):505-13.
199. Constantini NW, Dubnov G, Lebrun CM. The menstrual cycle and sport performance. Clin Sports Med. 2005;24(2):e51-e82 http://doi.org/10.1016/j.csm.2005.01.003.
200. Hopkins WG. Measures of reliability in sports medicine and science. Sports Med. 2000;30(1):1-15.
201. McAlinden C, Khadka J, Pesudovs K. Precision (repeatability and reproducibility) studies and sample-size calculation. Journal of cataract and refractive surgery. 2015;41(12):2598-604.
202. Gellish RL, Goslin BR, Olson RE, McDonald A, Russi GD, Moudgil VK. Longitudinal modeling of the relationship between age and maximal heart rate. Medicine and science in sports and exercise. 2007;39(5):822-9.

203. Nes B, Janszky I, Wisloff U, Støylen A, Karlsen T. Age-predicted maximal heart rate in healthy subjects: The HUNT Fitness Study. Scand J Med Sci Sports. 2013;23(6):697-704.
204. Robergs R, Landwehr R. The surprising history of the "HR_{max}=220-age" equation. J Exerc Physiol Online. 2002;5(2):1-10.
205. Sporis G, Vucetic V, Jukic I, Omrcen D, Bok D, Custonja Z. How reliable are the equations for predicting maximal heart rate values in military personnel? Mil Med. 2011;176(3):347-51.
206. Roy S, McCrory J. Validation of maximal heart rate prediction equations based on sex and physical activity status. International jjournal of eexercise sscience. 2015;8(4):318-30.
207. Lunt H, Roiz De Sa D, Roiz De Sa J, Allsopp A. Validation of one-mile walk equations for the estimation of aerobic fitness in British military personnel under the age of 40 years. Mil Med. 2013;178(7):753-9.
208. Cureton KJ, Sloniger MA, O'Bannon JP, Black DM, McCormack WP. A generalized equation for prediction of $\dot{V}O_{2peak}$ from 1-mile run/walk performance. Medicine and science in sports and exercise. 1995;27(3):445-51.

Appendices

Appendix 1: Full List of Search Terms Input to the Embase, Medline, CINAHL and Web Of Science Databases

Search Number	**Search term**
1	'exercise test' OR 'treadmill' OR 'treadmill ergometry' OR treadmill exercise'
2	'bruce' OR 'bruces'
3	Search 1 AND Search 2
4	('bruce' OR 'bruces') WITHIN 3 WORDS OF ('protocol' OR 'test')
5	Search 3 OR Search 4
6	'diagnostic accuracy' OR 'comparative study' OR 'evaluation study' OR 'reproducibility' OR 'reliability' OR 'sensitivity and specificity' OR 'validity' OR 'validation study' OR 'prediction and forecasting' OR 'odds ratio' OR 'statistical model' OR 'receiver operating characteristic'
7	ABBREVIATIONS OF 'Reproducibil' OR 'sensitiv' OR 'specificity' OR 'reliab' OR 'valid' OR 'accura' OR 'repeatability' OR 'estimat'
8	'predict' WITHIN 2 WORDS OF 'value'
9	'utility' WITHIN 2 WORDS OF 'test'
10	'roc curve' or 'received operating characteristic' OR 'kappa coefficient' OR 'intra-rater' OR 'inter-rater' OR 'interrater' OR 'intrarater' OR 'rater' OR 'likelihood ratio' OR 'likelihood function' OR 'odds ratio' OR 'test-retest' OR 'responsive'
11	Search 6 OR Search 7 OR Search 8 OR Search 9 OR Search 10
12	Search 5 AND Search 11

Appendix 2: AXIS Risk of Bias Tool[52]

	Question	Yes	No	Don't know/ Comment
Introduction				
1	Were the aims/objectives of the study clear?			
Methods				
2	Was the study design appropriate for the stated aim(s)?			
3	Was the sample size justified?			
4	Was the target/reference population clearly defined? (Is it clear who the research was about?)			
5	Was the sample frame taken from an appropriate population base so that it closely represented the target/reference population under investigation?			
6	Was the selection process likely to select subjects/participants that were representative of the target/reference population under investigation?			
7	Were measures undertaken to address and categorise non-responders?			
8	Were the risk factor and outcome variables measured appropriate to the aims of the study?			
9	Were the risk factor and outcome variables measured correctly using instruments/measurements that had been trialled, piloted or published previously?			
10	Is it clear what was used to determined statistical significance and/or precision estimates? (e.g. p-values, confidence intervals)			
11	Were the methods (including statistical methods) sufficiently described to enable them to be repeated?			
Results				
12	Were the basic data adequately described?			
13	Does the response rate raise concerns about non-response bias?			
14	If appropriate, was information about non-responders described?			
15	Were the results internally consistent?			
16	Were the results presented for all the analyses described in the methods?			
Discussion				
17	Were the authors' discussions and conclusions justified by the results?			
18	Were the limitations of the study discussed?			
Other				
19	Were there any funding sources or conflicts of interest that may affect the authors' interpretation of the results?			
20	Was ethical approval or consent of participants attained?			

Appendix 3: Ethical Approval Confirmation Document for Submaximal Exercise Testing Study

Coláiste na Tríonóide, Baile Átha Cliath
Trinity College Dublin
Ollscoil Átha Cliath | The University of Dublin

Kate MacNamara
Discipline of Physiotherapy
Trinity Centre for Health Sciences,
St James's Hospital,
Dublin 8

5th February 2019

Ref: 181205

Title of Study: A study of the repeatability of the Submaximal Bruce Protocol graded treadmill test in measuring VO2 and predicting maximal VO2.

Dear Kate,

Further to a meeting of the Faculty of Health Sciences Ethics Committee held in December 2018. We are pleased to inform you that the above project has ethical approval to proceed.

As a researcher you must ensure that you comply with other relevant regulations, including DATA PROTECTION and HEALTH AND SAFETY.

Yours sincerely,

pp Brian Brown
Prof. Brian O'Connell
Chairperson
Faculty Research Ethics Committee

Appendix 4: Sample Size Calculation for Submaximal and Maximal Bruce Protocol Repeatability Studies

Sample size was calculated using the sample size table for two-sample t-tests, as shown below. A two-sided test was required, and the equation required was $D = \frac{\Delta}{\sigma}$. Previously published studies examining the repeatability of the Bruce protocol in measuring $\dot{V}O_{2max}$ were used to determine the appropriate values for α, β, Δ and σ (see key below).

Key:

Symbol	Definition	Value
α	The significance level deemed appropriate; in this case, 0.05 is the significance level used in many previous Bruce protocol repeatability studies[53, 55-57, 59, 61].	0.05
β	Level of risk deemed appropriate in case of failing to detect a difference in the sample size (probability of type II error).	0.2
Δ	The difference deemed important to detect; in this case, the minimal clinically important difference (MCID) in $\dot{V}O_{2max}$.	3.5
σ	Size of the standard deviation (SD) which describes the expected chance variation; in this case, calculated from mean SD across previous Bruce protocol repeatability studies[1, 53-55, 57, 59-61].	4.41

D = Δ / σ = 3.5 / 4.09 = 0.86 *(rounded to 0.85 on table below)*. With β = 0.2, sample size = 23.

	α = 0.025 (one-tail) α = 0.05 (two-tails)				α = 0.05 (one-tail) α = 0.10 (two-tails)			
β D	0.01	0.05	0.1	0.2	0.01	0.05	0.1	0.2
0.70	76	55	44	34	66	45	36	26
0.75	67	48	39	29	57	40	32	23
0.80	59	42	34	26	50	35	28	21
0.85	52	37	31	23	45	31	25	18
0.90	47	34	27	21	40	28	22	16
0.95	42	30	25	19	36	25	20	15
1.00	38	27	23	17	33	23	18	14
1.10	32	23	19	15	27	19	15	11
1.20	27	20	16	12	23	16	13	10
1.30	23	17	14	11	20	14	11	9
1.40	20	15	12	10	17	12	10	8
1.50	18	13	11	9	15	11	9	7
1.60	16	12	10	8	14	10	8	6
1.70	14	11	9	7	12	9	7	6
1.80	13	10	8	6	11	8	7	5
1.90	12	9	7	6	10	7	6	5
2.00	11	8	7	6	9	7	6	4
2.10	10	8	6	5	8	6	5	4
2.20	9	7	6	5	8	6	5	4
2.30	9	7	6	5	7	5	5	4
2.40	8	6	5	4	7	5	4	4
2.50	8	6	5	4	6	5	4	3
3.00	6	5	4	4	5	4	3	3

Appendix 5: Full List of Participant Exclusion Criteria

Exclusion Criteria:		**Yes**	**No**
Female gender			
Have a Physical or mental impairment leading to an inability to exercise adequately			
Personal History of a cardiac (acute cardiac event, unstable chest pain, irregular heartbeat, cardiac infection), respiratory or metabolic (e.g. diabetes, renal disease) condition, or show any major signs or symptoms suggestive of these conditions (such as those described by the American Heart Association's Guidelines for Exercise Testing (Fletcher et al.2003) and the American College of Sports Medicine Risk Stratifications)			
2 or more of the following:			
	Have family history of myocardial infarction, coronary revascularisation or sudden death before 55 years of age in father or first-degree male relative, or before 65 years of age in mother or other female first-degree relative *(first degree relative: parent/sibling/child)*		
	Sedentary lifestyle		
	BMI ≥ 30kg·m^{-2}		
	High levels of low-density cholesterol (LDL) or low levels of high-density cholesterol (HDL)		
	Pre-diabetes (HbA1C 5.7-6.4%)		
Answers "yes" to any PAR-Q questionnaire which could contraindicate exercise testing: screen each "yes" answer individually.			
Have Known blood pressure abnormalities (e.g. hypertension/hypotension)			
Have any neuromotor, musculoskeletal or rheumatic condition, any inflammatory, autoimmune or allergic conditions, any chronic infectious disease (e.g. Hepatitis C, HIV/AIDS)			
Use medications			
Have had any Musculoskeletal injury in the previous 3 months			
Current smoker, or quit smoking within the past 6 months			
Have Epilepsy			
Have a Fitted electronic device (e.g. pacemaker)			
Exclusion for any other reason deemed appropriate by the lead investigator			

Appendix 6: American College of Sports Medicine Cardiovascular Risk Classification Guidelines

Appendix 6.1 ACSM's Cardiovascular Risk Classification[153]

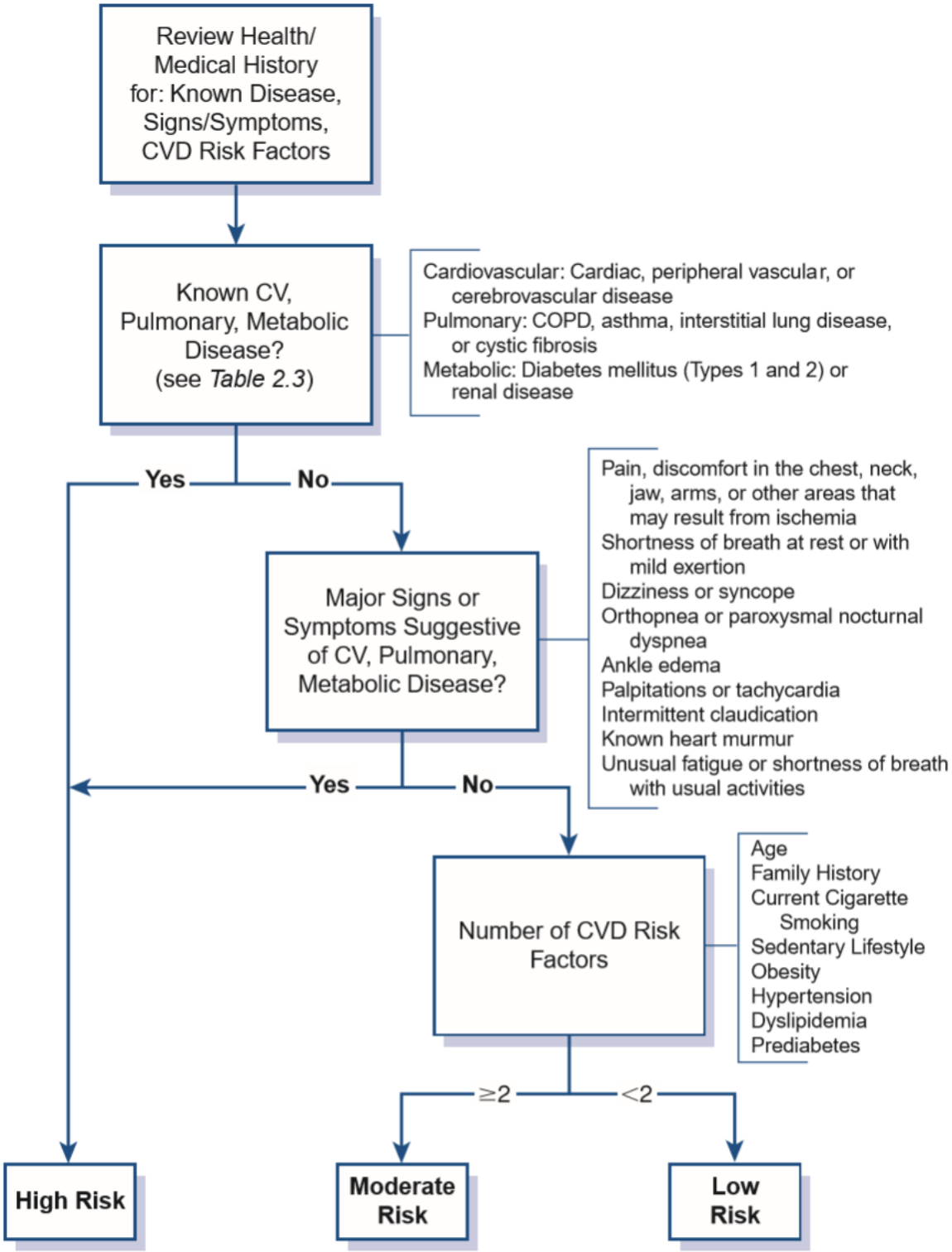


FIGURE 2.3. Logic model for classification of risk. CV, cardiovascular; CVD, cardiovascular disease.

Appendix 6.2 ACSM's Cardiovascular Disease Risk Factors[153]

TABLE 2.2. Atherosclerotic Cardiovascular Disease (CVD) Risk Factors and Defining Criteria (26,31)

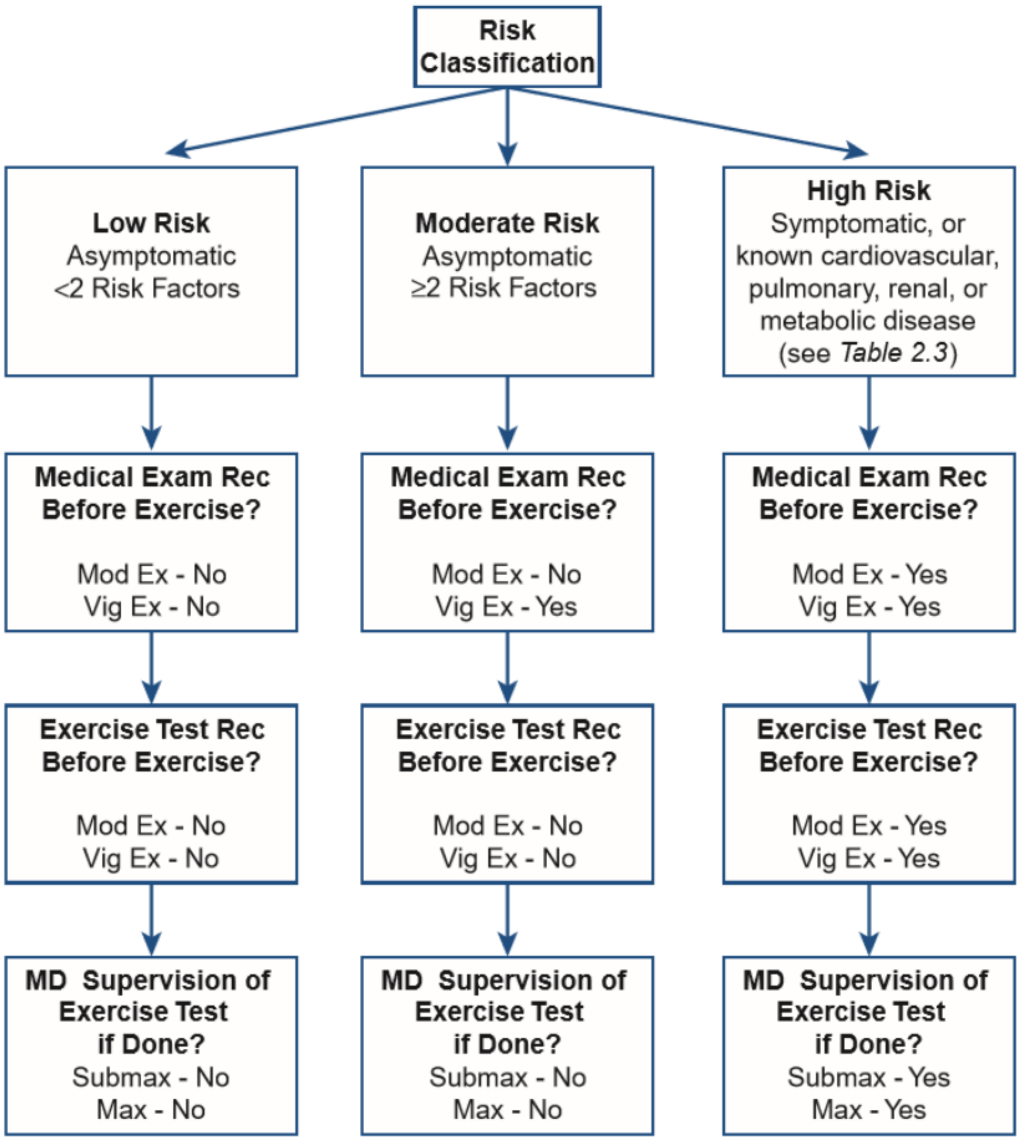
Risk Factors	Defining Criteria
Age	Men ≥45 yr; women ≥55 yr (12)
Family history	Myocardial infarction, coronary revascularization, or sudden death before 55 yr in father or other male first-degree relative or before 65 yr in mother or other female first-degree relative
Cigarette smoking	Current cigarette smoker or those who quit within the previous 6 mo or exposure to environmental tobacco smoke
Sedentary lifestyle	Not participating in at least 30 min of moderate intensity, physical activity (40%–<60% $\dot{V}O_2R$) on at least 3 d of the week for at least 3 mo (22,30)
Obesity	Body mass index ≥30 kg · m^{-2} *or* waist girth >102 cm (40 in) for men and >88 cm (35 in) for women (10)
Hypertension	Systolic blood pressure ≥140 mm Hg and/or diastolic ≥90 mm Hg, confirmed by measurements on at least two separate occasions, *or* on antihypertensive medication (9)
Dyslipidemia	Low-density lipoprotein (LDL) cholesterol ≥130 mg · dL^{-1} (3.37 mmol · L^{-1}) *or* high-density lipoprotein[b] (HDL) cholesterol <40 mg · dL^{-1} (1.04 mmol · L^{-1}) *or* on lipid-lowering medication. If total serum cholesterol is all that is available, use ≥200 mg · dL^{-1} (5.18 mmol · L^{-1}) (21)
Prediabetes[a]	Impaired fasting glucose (IFG) = fasting plasma glucose ≥100 mg · dL^{-1} (5.55 mmol · L^{-1}) and ≤125 mg · dL^{-1} (6.94 mmol · L^{-1}) *or* impaired glucose tolerance (IGT) = 2 h values in oral glucose tolerance test (OGTT) ≥140 mg · dL^{-1} (7.77 mmol · L^{-1}) and ≤199 mg · dL^{-1} (11.04 mmol · L^{-1}) confirmed by measurements on at least two separate occasions (5)
Negative Risk Factors	**Defining Criteria**
High-density lipoprotein (HDL) cholesterol	≥60 mg · dL^{-1} (1.55 mmol · L^{-1})

[a]If the presence or absence of a CVD risk factor is not disclosed or is not available, that CVD risk factor should be counted as a risk factor except for prediabetes. If the prediabetes criteria are missing or unknown, prediabetes should be counted as a risk factor for those ≥45 yr, especially for those with a body mass index (BMI) ≥25 kg · m^{-2}, and those <45 yr with a BMI ≥25 kg · m^{-2} and additional CVD risk factors for prediabetes. The number of positive risk factors is then summed.

[b]High HDL is considered a negative risk factor. For individuals having high HDL ≥60 mg · dL^{-1} (1.55 mmol · L^{-1}), for these individuals one positive risk factor is subtracted from the sum of positive risk factors.

$\dot{V}O_2R$, oxygen uptake reserve.

Appendix 6.3 ACSM's Recommendations based on Risk Classification[153]



Mod Ex: Moderate intensity exercise; 40%–<60% $\dot{V}O_2R$; 3–<6 METs
"An intensity that causes noticeable increases in HR and breathing."

Vig Ex: Vigorous intensity exercise; ≥60% $\dot{V}O_2R$; ≥6 METs
"An intensity that causes substantial increases in HR and breathing."

Not Rec: Reflects the notion a medical examination, exercise test, and physician supervision of exercise testing are not recommended in the preparticipation screening; however, they may be considered when there are concerns about risk, more information is needed for the Ex R_x, and/or are requested by the patient or client.

Rec: Reflects the notion a medical examination, exercise test, and physician supervision are recommended in the preparticipation health screening process.

■ **FIGURE 2.4.** Medical examination, exercise testing, and supervision of exercise testing preparticipation recommendations based on classification of risk. Ex R_x, exercise prescription; HR, heart rate; METs, metabolic equivalents; $\dot{V}O_2R$, oxygen uptake reserve.

Appendix 7: Participant Information Leaflet for Submaximal Exercise Testing Study

Trinity College Dublin
Coláiste na Tríonóide, Baile Átha Cliath
The University of Dublin

Participant Information Leaflet

A study of the repeatability of the Submaximal Bruce Protocol graded treadmill test in measuring $\dot{V}O_2$ and predicting maximal $\dot{V}O_2$

A research project at the Department of Physiotherapy in Trinity College, Dublin

Introduction

The maximum volume of oxygen used during exercise ($\dot{V}O_{2max}$) is considered the best way of measuring a person's level of fitness. $\dot{V}O_{2max}$ is also strongly related to a person's risk of cardiovascular disease. To assess a person's $\dot{V}O_{2max}$ most accurately, the person can complete a treadmill or cycle test which pushes them to their highest level of exertion, while their breath gases are analysed, giving a value for $\dot{V}O_{2max}$.

However, in some populations (e.g. the elderly, or persons with pre-existing medical conditions), maximal exercise testing is not recommended and may be associated with increased risk of adverse cardiac (or other health related) events during testing. In these cases, the person can exercise on the treadmill or cycle ergometer to a certain percentage of their estimated maximal heart rate (i.e. submaximal exertion) before stopping, while breath gases are analysed, and data is gathered. Calculations can then be done with the data gathered during the submaximal exercise, to estimate the person's $\dot{V}O_{2max}$.

Aim of current study

Our aim is to perform one particular submaximal treadmill exercise test, the Submaximal Bruce Protocol, in a healthy male population, to test the accuracy and repeatability of the protocol, and to study the accuracy of the equation used to calculate $\dot{V}O_{2max}$, by comparing results from three repetitions of the protocol in the same individual.

If you would like to take part in this study, please take time to read this document carefully. You should understand the risks and benefits of taking part in the study, so that you can make a decision that is right for you. This is known as "Informed Consent".

To take part you have to be available for three testing sessions, each one week apart. To be eligible to participate you must be male, aged between 18 and 35 years.

Procedures

If you take part in this study, you will visit the exercise laboratory in the Trinity Centre for Health Sciences in St James' Hospital, or the Department of Physiology on main college campus, on three occasions, each session one week apart. The first session will take approximately 40-60 minutes of your time, and the additional testing sessions will each last 30 minutes. Please see below for preparation guidelines for the day of testing.

There will be 2 main components on your first visit and only 1 component on your additional visits.

1. Body composition analysis

You will be asked to fast from midnight the night before each testing visit. Upon arrival, your standing height will be measured. The amount of fat, water and muscle in your body will be estimated using a machine (pictured, right) that analyses details of body weight, body mass index (BMI), percentage body fat, muscle mass and fat free mass. This machine is non-invasive and will not cause any pain. You will be asked to remove your shoes and socks only for this procedure. Waist circumference and body weight will also be measured manually. For each additional visit, only weight will be required.

2. Submaximal Bruce Protocol: Exercise Treadmill Test

Before you begin exercising on your first day of testing, you will be asked to complete a short questionnaire (the Physical Activity Readiness Questionnaire, or PAR-Q). The results of this questionnaire will indicate if you are safe to carry out an exercise test. Your cholesterol and blood sugar levels will also be checked my means of a finger-prick blood test. You will be familiarised with how the test will work, and be shown the equipment to be used in this study: a gas analysis system with a face-mask and heart rate monitor (pictured, right).

For each test you will be fitted with a face mask and asked to walk on the treadmill, to follow the submaximal Bruce exercise testing protocol. It begins at a slow pace of

1.7mph, at a 10% treadmill incline, and gradually increases in speed and incline every three minutes. Your heart rate, blood pressure and blood lactate will be monitored throughout the test. The researcher will end the test when you reach 85% of your calculated maximal Heart Rate. However, you can end the test at any point, if you feel pain or discomfort in your chest, neck, limbs or any other area, if you feel dizzy, too breathless, or as if you might faint, or if you feel too fatigued. The researcher will be monitoring for unexpected changes in blood pressure and heart rate, and will end the test if these occur, or if there are any technical difficulties with the equipment being used.

If there are no complications, there will be a 5-minute cool-down phase when the test is ended, at a slow pace while your heart returns to its normal pace.

Benefits

Full analysis of fitness levels and body composition measurements will be provided to each participant on completion of the testing procedures in the form of an individualised health report. These can be beneficial to understand your own fitness level and how you compare to fitness levels of the general population, and the data can be used to guide progression of your own exercise regime. Participating in this study will also benefit this field of research by adding to it.

Risks

Complications during submaximal exercise are rare, especially when you have been cleared for exercise and have no medical history that increases your risk during exercise. However, as you are exercising after a period of fasting, there is a small risk that you could experience:

- pain,
- fatigue,
- dizziness or ataxia (loss of control of voluntary muscle movements),
- difficulty breathing, or wheezing,
- nausea
- leg cramps (claudication), or
- signs of poor blood flow (cyanosis, or pallor)

If you experience any of the above, the test will be stopped immediately, and you will be monitored closely. If symptoms do not subside, you will be seen by a doctor and if necessary, brought to A&E of St. James's Hospital.

In the unlikely event of a data breach (i.e. your personal information is mislaid, lost or stolen), you will be alerted as soon as possible and the case will be reported to the Data Commissioner (please see contact details below).

Exclusion From Participation

You cannot be in this study if any of the following applies to you:

- Male gender less than 18 or over 35 years old
- Female gender
- Non-fluent in English
- Intellectual or cognitive disability that would affect your ability to give informed consent or exercise on a treadmill safely
- History of a cardiac (such as acute cardiac event, unstable chest pain, irregular heart beat, cardiac infection), respiratory or neurological condition
- Known blood pressure abnormalities (e.g. hypertension/hypotension)
- Answer "yes" to any one question on the PAR-Q, which would contra-indicate exercise testing
- Any neuromotor, musculoskeletal or rheumatic condition, any inflammatory, autoimmune or allergic conditions, any chronic infectious disease (e.g. Hepatitis C, HIV/AIDS), or metabolic condition such as diabetes
- Medication use
- Musculoskeletal injury in the previous 3 months
- Smoker
- Epilepsy
- BMI ≥ 30
- Fitted electronic device (e.g. pacemaker)
- Exclusion for any other reason deemed appropriate by the lead investigator.

Confidentiality and Data Protection

What data will be collected?

Personal data to be collected in this study will include your gender, age, any relevant past medical history, body composition, and aerobic fitness levels. This information is needed to analyse how effective and repeatable the Submaximal Bruce Protocol is. Only personal data which is relevant for the purpose of this study is collected and used (this is called "data minimisation").

Who has access to the collected data?

The data controllers for this study are the research team conducting the study (consisting of the Lead Investigator and the Research Supervisor as named below, as well as a number of final year undergraduate physiotherapy students) in conjunction with Trinity College Dublin. The Lead Investigator and Research Supervisor have undergone training in data protection law and practice, prior to starting this research.

How will your data be stored and protected?

Your identity will remain confidential. Your name and personal details will not be published and will not be disclosed to anyone outside of the research team. All of your details and

results will be coded with numerical ID to maintain your confidentiality (this is called "pseudonymisation"). Your data will be coded, rather than being kept completely anonymous, to allow for comparison of your results from one testing day to the next. All information relating to you in hard-copy form will be stored in a locked filing cabinet within a secure office only accessible by the research team, and information and records in electronic form will be stored on a password-protected PC at the Trinity Centre for Health Sciences. Your study information and results will be retained for 5 years in keeping with good research practice standards and data protection legislation. It will be destroyed after this time (electronic data will be erased, and hard copy forms will be shredded).

How will your data be used, now and in the future?

The information collected in this study will be analysed, and the overall findings of this study may be published in international peer reviewed journals and may be shared at research conferences. The results of the study may be used for comparative purposes in other studies of a similar nature examining the reliability of the Bruce treadmill protocol. However, your data will remain coded and your personal identifiers will never be published or disclosed to anyone outside of this research team.

Your data will only be used for comparative purposes in studies with have Research Ethics Committee approval. Your rights under GDPR and what will happen in the event of a data breach as outlined below will still apply for use of data in future studies. The data controllers and researchers in this project are bound by our Professional Code of Conduct to maintain confidentiality regarding all data gained during this research.

Is there any risk with processing and storing your data? What will happen if there is a data breach?

Considering that sensitive personal data relating to your health is involved, in the unlikely event of a data breach (i.e. data being mislaid, lost or stolen), you will be notified as soon as possible, and it will be reported immediately to the Data Protection Commissioner.

What are your rights under GDPR?

You have the right to:

- Access your data
- Rectify or correct any mistakes with your data
- Have your data erased or deleted
- Data portability (moving your data from one controller to another)
- Object to or stop the processing or profiling of your data

- Lodge a complaint to the Data Protection Commissioner (contact: +353 57 8684800 or +353 (0)761 104 800; https://dataprotection.ie/en/contact).

What is the lawful basis to using your personal data?

Your data will be processed under the lawful basis of Article 6(1)(e) and 9(2)(j) of the EU General Data Protection Act 2016. If you have any queries regarding your data, or the General Data Protection Rules (GDPR), you can contact the research team (details at end of document) or the Data Protection Officer of Trinity College Dublin, by email: dataprotection@tcd.ie .

Compensation

The research team is covered by standard medical malpractice insurance. Nothing in this document restricts or curtails your rights.

Voluntary Participation

Participation in this study is fully voluntary, and you will be asked to sign a consent form before taking prat, after you have read this document and have had the study procedures, risks and benefits explained to you. You do not have to take part in this study, and you may withdraw participation at any time, even if the study has already started. You do not have to give a reason for not taking part in or for leaving the study.

If you decide not to participate, or if you withdraw participation, you will not be penalized and will not give up any benefits which you had before entering the study. You should not feel in any way obliged to take part in this study. If you do not give consent, or withdraw your consent, no attempt will be made to access your data, and **no** application will be made by the research team for a consent exemption to the Health Research Consent Declaration Committee.

If you wish to seek more information about this research study, or if you wish to opt-out, please contact the research team on the details provided overleaf.

Stopping The Study

You understand that the research team may stop your participation in the study at any time without your consent.

Permission

This research project was given ethics committee approval from the Faculty of Health Sciences Research Ethics Committee, Trinity College Dublin (e-mail: ethicscommittee@tcd.ie) on

05/02/2029. There is no known personal connection between members of the research team and this ethics committee.

Future Contact

After completion of the study, you may be contacted again by the researchers in relation to your study results, if you express that you would like to receive feedback on them.

On The Day Of Testing

If you consent to taking part in this study, the following points explain how to prepare for your testing day.

- Please fast from midnight on the night before your assessment days. Please limit your liquid intake while fasting: you may drink water, but refrain from caffeine drinks, energy drinks, and alcohol. Fasting ensures more accurate results of body composition analysis. A snack and drink will be provided to you after each assessment.
- Please wear loose clothes and comfortable shoes that you will be able to exercise in.
- Bring a towel, shower gel, and change of clothes if you wish to shower after testing.
- Please try to drive or use public transport to get to the testing venue and avoid walking or cycling on the days of your visit. It will make for more accurate results of your fitness test if you have not done much physical activity prior to testing. Please refrain from strenuous physical activity for 24-hours prior to each assessment.
- We request that you continue your normal physical activity regime for the duration of the study, and not to increase the training intensity or frequency.

Further Information

For more information or answers to your questions about the study, your participation in the study, and your rights please see the contact details below.

Principal Investigator/Research Supervisor: Prof. John Gormley, Discipline of Physiotherapy, Trinity College Dublin. Contact: Tel (01) 8962121, E-mail: jgormley@tcd.ie

Lead Investigator: Ms. Kate Macnamara, Research Masters Student, Trinity College Dublin. Contact: Tel (01) 8963613, E-mail: macnamak@tcd.ie

Data Controller: Trinity College Dublin. **Data Protection Officer, Trinity College Dublin:** Contact E-mail: dataprotection@tcd.ie

Appendix 8: Participant Consent Form for Submaximal Exercise Testing Study

Trinity College Dublin
Coláiste na Tríonóide, Baile Átha Cliath
The University of Dublin

Participant Consent Form

Study Title: A study of the repeatability of the Submaximal Bruce Protocol graded treadmill test in measuring $\dot{V}O_2$ and predicting maximal $\dot{V}O_2$

Lead Investigator: Ms Kate Macnamara; **Principal Investigator:** Prof. John Gormley

Study Information

This study aims to test the accuracy and repeatability of the submaximal Bruce Protocol exercise treadmill test, and to study the accuracy of the equation used to calculate maximal oxygen consumption/aerobic capacity ($\dot{V}O_{2max}$), by comparing results from three repetitions of the protocol in the same individual.

As a participant, you are required to attend three testing sessions in the Trinity Centre for Health Sciences, St James's Hospital or at Trinity College Dublin. The first session will last 40-60 minutes, and the following two sessions will last approximately 30 minutes. Testing involves assessing body composition (using a non-invasive body-composition analyser) and cardiorespiratory fitness (by means of assessing breath gases during a graded exercise test on a treadmill). There will also be a minimally-invasive finger-prick blood test to measure cholesterol and blood glucose levels prior to exercise testing.

Once testing is completed, your fitness and body composition analysis results will be provided in the form of an individualised health report and you will be given an opportunity to discuss these results with the research team.

During the testing, there is a risk of experiencing pain (in chest, limbs, neck), dizziness, nausea, difficulty breathing, fatigue, or legs cramps while exercising intensely. If this occurs, you must inform the investigator. Testing will be stopped immediately and the investigators will assess your condition. If further care is required due to an adverse event, you will be transported to the nearest hospital facility for further care.

All data collected will be coded with numerical ID to maintain participant confidentiality. Data will be stored in a secure office and on a password protected computer for 5 years after the study is finished in accordance with good research standard guidelines and data protection legislation. It will be destroyed after this period of time by the relevant personnel. The information will not be

used for any other purpose other than this research study without your permission. Participation is entirely voluntary and you may withdraw your consent at any stage of testing.

Statement of Participant's Consent

Explicit, informed, voluntary consent to partake in the research study.

Please tick each item to give your consent.

I confirm that I have read and fully understood the relevant Participation Information Leaflet (dated 25th January 2019) provided to me.	YES ☐	NO ☐
I understand I will be asked to undertake a research assessment which collects data about my gender, age and relevant past medical history, as well as the assessment of body composition, and my aerobic fitness by use of a treadmill test (the Submaximal Bruce Protocol). I have been assured that information about me will be kept private and confidential.	YES ☐	NO ☐
I understand the risks explained to me and listed above, regarding pain, dizziness, nausea, fatigue and breathlessness during exercise testing. I am aware of the benefits and risks of this research study.	YES ☐	NO ☐
I have had the opportunity to discuss the study and ask questions about the study, and I have received satisfactory answers to all my questions.	YES ☐	NO ☐
I have received enough information about this study and understand what is involved if I agree to participate.	YES ☐	NO ☐
I understand that I am free to withdraw from the study at any time without giving a reason and with no consequence to myself.	YES ☐	NO ☐
I agree to be contacted by researchers as part of this study.	YES ☐	NO ☐
I freely and voluntarily consent to take part in this research study having been fully informed of the risks, benefits and alternatives.	YES ☐	NO ☐
Participant's Name (BLOCK CAPITALS)		
Participant's Signature		
Date		
Phone Number		

To be completed by the **Researcher:**

I have fully explained the purpose and nature (including benefits and risks) of this study to the participant in a way that he could understand. I have invited him to ask questions on any aspect of the study.	YES ☐	NO ☐
I confirm that I have given a copy of the information leaflet and consent form to the participant.	YES ☐	NO ☐
Researcher's Name (BLOCK CAPITALS)		
Researcher's Title and Qualifications		
Researcher's Signature		
Date		

You are now entering a separate part of this consent form, relating to data protection

Explicit, informed, voluntary consent regarding data protection.

I understand that all of my data will be pseudonymised and minimised for this study titled "A study of the repeatability of the	YES ☐	NO ☐

Submaximal Bruce Protocol graded treadmill test in measuring $\dot{V}O_2$ and predicting maximal $\dot{V}O_2$". The words 'pseudonymisation' and minimisation' have been explained to me.		
I understand my name or other personal identifiers will not be disclosed to anybody not involved with this study and my personal data will be kept strictly confidential.	YES ☐	NO ☐
I understand the research team (study investigators) will be processing my data and they, in conjunction with Trinity College Dublin, are in control of my data.	YES ☐	NO ☐
I understand that my data will be used for study analysis, published in peer reviewed journals, in presentations, and may be disseminated at conferences but my data will remain confidential and none of my personal identifiers will be disclosed in these circumstances.	YES ☐	NO ☐
I understand how my data will be stored (pseudonymised and minimised in secure locations) and that it will be stored for a total of 5 years and will then be destroyed by the study investigators.	YES ☐	NO ☐
I have been made aware of my rights under the General Data Protection Regulations, and contact details of the Data Protection Officer and Data Commissioner have been provided to me in the Participant Information Leaflet.	YES ☐	NO ☐
I have read and understood the personal data protection section of the Participant Information Leaflet (dated 25th January 2019).	YES ☐	NO ☐
I have had the opportunity to discuss data protection in this study and I have received satisfactory answers to all my questions.	YES ☐	NO ☐
I understand that I am free to withdraw from the study at any time without giving a reason with no consequence and my personal data will not be used and will be destroyed.	YES ☐	NO ☐
I freely and voluntarily consent to allow the researcher's use of my information (personal data) as part of this study as outlined in the information leaflet.	YES ☐	NO ☐
Participant's Name (BLOCK CAPITALS)		
Participant's Signature		
Date		

To be completed by the **Researcher:**

I have fully explained the purpose and nature of data protection in this study to the participant in a way that he could understand. I have invited him to ask questions on any aspect of the study.	YES ☐	NO ☐
I confirm that I have given a copy of the information leaflet and consent form to the participant.	YES ☐	NO ☐
Researcher's Name (BLOCK CAPITALS)		
Researcher's Title and Qualifications		
Researcher's Signature		
Date		

You are now entering a separate part of this consent form relating to use of data for future studies

Explicit, informed, voluntary consent of storage and future use of data/information.

I understand that my data collected from the research study, "A study of the repeatability of the Submaximal Bruce Protocol graded treadmill test in measuring $\dot{V}O_2$ and predicting maximal $\dot{V}O_2$" may be used for comparative purposes in similar studies assessing the reliability and repeatability of the Bruce Protocol.	YES ☐	NO ☐
I understand my name or other personal identifiers will not be disclosed to anybody not involved with this study and my personal data will be kept strictly confidential in these circumstances.	YES ☐	NO ☐
I understand that my data will be used for **comparative** study analysis which may be published in peer reviewed journals, in presentations and may be disseminated at conferences but my data will remain confidential and none of my personal identifiers will be disclosed in these circumstances.	YES ☐	NO ☐
I understand how my data will be stored (the same as for the original study – pseudonymised and minimalised in secure locations), that it will be sotred for a total of 5 years for these purposes, and will then be destroyed by the study investigators.	YES ☐	NO ☐
I have read and understood the use of personal data in future studies section of the Participant Information Leaflet (dated 25th January 2019).	YES ☐	NO ☐
I have had the opportunity to discuss future use of personal data in this study and I have received satisfactory answers to all my questions.	YES ☐	NO ☐
I understand that I am free to withdraw my consent to use my personal data in future studies at any time without giving a reason, with no consequence, and my personal data will not be used and will be destroyed.	YES ☐	NO ☐
I freely and voluntarily consent for my data to be stored **for comparative purposes in similar studies assessing the repeatability of the Bruce Protocol** which is **unrelated** to the current study **without further consent** being required but only if the research is approved by a Research Ethics Committee.	YES ☐	NO ☐

Participant's Name (BLOCK CAPITALS)	
Participant's Signature	
Date	

To be completed by the **Researcher:**

I have fully explained the purpose and nature of storage and future use of personal data of this study, "A study of the repeatability of the Submaximal Bruce Protocol graded treadmill test in measuring $\dot{V}O_2$ and predicting maximal $\dot{V}O_2$" to the participant in a way that he could understand. I have invited him to ask questions on any aspect of the study.	YES ☐	NO ☐
I confirm that I have given a copy of the information leaflet and consent form to the participant.	YES ☐	NO ☐

Researcher's Name (BLOCK CAPITALS)	
Researcher's Title and Qualifications	
Researcher's Signature	
Date	

Appendix 9: Physical Activity Readiness Questionnaire (PAR-Q)[171]

Physical Activity Readiness Questionnaire - PAR-Q (revised 2002)

PAR-Q & YOU

(A Questionnaire for People Aged 15 to 69)

Regular physical activity is fun and healthy, and increasingly more people are starting to become more active every day. Being more active is very safe for most people. However, some people should check with their doctor before they start becoming much more physically active.

If you are planning to become much more physically active than you are now, start by answering the seven questions in the box below. If you are between the ages of 15 and 69, the PAR-Q will tell you if you should check with your doctor before you start. If you are over 69 years of age, and you are not used to being very active, check with your doctor.

Common sense is your best guide when you answer these questions. Please read the questions carefully and answer each one honestly: check YES or NO.

YES	NO		
☐	☐	1.	**Has your doctor ever said that you have a heart condition and that you should only do physical activity recommended by a doctor?**
☐	☐	2.	**Do you feel pain in your chest when you do physical activity?**
☐	☐	3.	**In the past month, have you had chest pain when you were not doing physical activity?**
☐	☐	4.	**Do you lose your balance because of dizziness or do you ever lose consciousness?**
☐	☐	5.	**Do you have a bone or joint problem (for example, back, knee or hip) that could be made worse by a change in your physical activity?**
☐	☐	6.	**Is your doctor currently prescribing drugs (for example, water pills) for your blood pressure or heart condition?**
☐	☐	7.	**Do you know of any other reason why you should not do physical activity?**

If you answered

YES to one or more questions

Talk with your doctor by phone or in person BEFORE you start becoming much more physically active or BEFORE you have a fitness appraisal. Tell your doctor about the PAR-Q and which questions you answered YES.

- You may be able to do any activity you want — as long as you start slowly and build up gradually. Or, you may need to restrict your activities to those which are safe for you. Talk with your doctor about the kinds of activities you wish to participate in and follow his/her advice.
- Find out which community programs are safe and helpful for you.

NO to all questions

If you answered NO honestly to all PAR-Q questions, you can be reasonably sure that you can:

- start becoming much more physically active – begin slowly and build up gradually. This is the safest and easiest way to go.
- take part in a fitness appraisal – this is an excellent way to determine your basic fitness so that you can plan the best way for you to live actively. It is also highly recommended that you have your blood pressure evaluated. If your reading is over 144/94, talk with your doctor before you start becoming much more physically active.

DELAY BECOMING MUCH MORE ACTIVE:

- if you are not feeling well because of a temporary illness such as a cold or a fever – wait until you feel better; or
- if you are or may be pregnant – talk to your doctor before you start becoming more active.

PLEASE NOTE: If your health changes so that you then answer YES to any of the above questions, tell your fitness or health professional. Ask whether you should change your physical activity plan.

Informed Use of the PAR-Q: The Canadian Society for Exercise Physiology, Health Canada, and their agents assume no liability for persons who undertake physical activity, and if in doubt after completing this questionnaire, consult your doctor prior to physical activity.

No changes permitted. You are encouraged to photocopy the PAR-Q but only if you use the entire form.

NOTE: If the PAR-Q is being given to a person before he or she participates in a physical activity program or a fitness appraisal, this section may be used for legal or administrative purposes.

"I have read, understood and completed this questionnaire. Any questions I had were answered to my full satisfaction."

NAME _____

SIGNATURE _____ DATE _____

SIGNATURE OF PARENT or GUARDIAN (for participants under the age of majority) _____ WITNESS _____

Note: This physical activity clearance is valid for a maximum of 12 months from the date it is completed and becomes invalid if your condition changes so that you would answer YES to any of the seven questions.

Supported by: Health Canada Santé Canada

continued on other side...

Appendix 10: Raw data for Submaximal Bruce Protocol Repeatability Study (Chapter 4)

Appendix 10.1 Baseline characteristics and body composition analysis results

Participant	Age (years)	Height (cm)	Weight (kg) Test 1	Weight (kg) Test 2	Weight (kg) Test 3	BMI ($kg \cdot m^{-2}$)	Fat mass (kg)	Muscle mass (kg)	Fat (%)	Free fat mass (kg)	Basal metabolic rate (kcal)	Bone mass (kg)
1	27	171.8	74.35	72.95	72.65	25.2	14.30	57.05	missing	missing	missing	missing
2	20	188.4	75.20	75.85	76.45	21.2	7.90	63.95	10.5	67.3	1989	3.35
3	21	178.0	77.00	76.20	76.85	24.3	9.95	63.75	12.9	67.05	1983	3.35
4	25	182.0	72.05	71.25	71.00	21.8	8.85	60.05	12.3	63.2	1847	3.15
5	23	198.4	89.30	88.50	88.25	22.7	13.05	72.5	14.6	76.25	2261	3.75
6	24	180.4	96.20	93.45	93.70	29.6	22.00	70.5	22.9	74.15	2226	3.65
7	20	176.0	66.10	65.30	65.50	21.3	10.25	53.05	15.5	55.85	1670	2.8
8	23	171.2	55.85	56.50	56.10	19.1	6.15	47.15	11	49.7	1741	2.55
9	21	185.7	75.10	75.20	74.80	21.8	7.50	64.25	10	67.6	1990	3.35
10	22	181.6	99.10	97.15	96.20	30	20.25	75.05	20.4	78.9	2376	3.85
11	20	175.3	75.95	76.25	77.15	24.7	14.05	58.8	18.5	61.9	1855	3.1
12	20	182.2	75.20	70.15	70.65	22.7	13.00	59.1	17.3	62.2	1859	3.1
13	21	180.3	68.85	69.30	69.20	21.2	10.55	55.4	15.3	58.3	1734	2.95
14	19	185.6	73.70	74.35	73.40	21.4	8.50	61.95	11.5	65.2	1936	3.25
15	21	182.9	71.75	72.10	70.85	21.4	9.15	59.45	12.8	62.55	1849	3.15
16	18	182.5	70.70	70.85	71.00	21.2	10.30	57.4	14.5	60.4	1812	3.05
17	33	189.6	89.00	90.30	90.10	24.8	21.65	64	24.3	67.35	1990	3.35
18	34	171.5	74.55	75.15	74.80	25.3	11.15	60.2	15	63.35	1832	3.15

BMI = body mass index

Appendix 10.2 Number of days between repeated tests

Participant	Test 1 to Test 2	Test 2 to Test 3
1	7 days	7 days
2	7 days	7 days
3	7 days	7 days
4	7 days	7 days
5	7 days	7 days
6	12 days	7 days
7	7 days	7 days
8	7 days	7 days
9	7 days	14 days
10	11 days	7 days
11	12 days	7 days
12	8 days	6 days
13	7 days	7 days
14	7 days	7 days
15	7 days	7 days
16	7 days	7 days
17	7 days	7 days
18	11 days	3 days

Appendix 10.3 $\dot{V}O_{2max}$ ($ml \cdot kg^{-1} \cdot min^{-1}$) predicted from "Fitmate" equation from submaximal Bruce protocol data

Participant	Test 1	Test 2	Test 3
1	*	59.58	64.02
2	57.62	58.30	59.03
3	60.22	61.84	63.29
4	*	58.52	59.26
5	55.29	58.04	55.56
6	42.35	52.83	51.66
7	62.52	60.10	61.72
8	55.19	64.95	60.58
9	63.57	60.33	63.58
10	57.50	57.06	61.91
11	64.93	69.48	67.18
12	51.86	56.70	51.74
13	60.09	59.64	68.05
14	54.59	56.39	57.54
15	66.70	70.05	64.45
16	46.40	48.23	47.64
17	49.08	50.95	55.52
18	54.69	61.63	63.51

Appendix 10.4 Submaximal Bruce protocol test durations (seconds)

Participant	Test 1	Test 2	Test 3
1	*	632	660
2	568	562	584
3	620	569	595
4	*	554	591
5	610	599	622
6	412	431	490
7	562	566	552
8	580	661	530
9	605	601	606
10	636	618	647
11	646	693	701
12	416	550	420
13	588	584	601
14	560	496	463
15	742	702	650
16	388	377	403
17	543	561	579
18	606	597	607

* = test was not completed

Appendix 11: Ethical Approval Confirmation Document for Maximal Exercise Testing Study

Coláiste na Tríonóide, Baile Átha Cliath
Trinity College Dublin
Ollscoil Átha Cliath | The University of Dublin

Kate Macnamara
Discipline of Physiotherapy,
Trinity Centre for Health Sciences,
St James's Hospital,
Dublin 8

18th June 2019

Ref: 190503

Title of Study: A study of the repeatability of the Bruce Protocol graded treadmill test in measuring maximal VO2.

Dear Kate,

Further to a meeting of the Faculty of Health Sciences Ethics Committee held in May 2019. We are pleased to inform you that the above project has ethical approval to proceed.

As a researcher you must ensure that you comply with other relevant regulations, including DATA PROTECTION and HEALTH AND SAFETY.

Yours sincerely,

pp Gloria Brown
Prof. Brian O'Connell
Chairperson
Faculty Research Ethics Committee

Appendix 12: Participant Information Leaflet for Maximal Exercise Testing Study

Trinity College Dublin
Coláiste na Tríonóide, Baile Átha Cliath
The University of Dublin

Participant Information Leaflet

A study of the repeatability of the Bruce Protocol graded treadmill test in measuring maximal $\dot{V}O_2$

A research project at the Department of Physiotherapy in Trinity College, Dublin

Introduction

Cardiopulmonary exercise testing is the best way of measuring a person's level of fitness. This is done by measuring the maximum volume of oxygen used during exercise ($\dot{V}O_{2max}$). $\dot{V}O_{2max}$ is also strongly related to a person's risk of cardiovascular disease. To assess $\dot{V}O_{2max}$ most accurately, subjects can complete a treadmill or cycle test which pushes them to their highest level of exertion, while their breath gases are analysed to get a value for $\dot{V}O_{2max}$. In some populations (e.g. the elderly; people with pre-existing medical conditions), maximal exercise testing is not recommended and may be associated with increased risk of adverse cardiac (or other health related) events during testing. In these cases, subjects can complete a "submaximal" exercise test, to a lesser intensity, while breath gases are analysed, and data is gathered. Calculations are then done with the data to estimate $\dot{V}O_{2max}$.

The most commonly used maximal test protocol is the Bruce Treadmill Protocol, in which the participant starts walking on a treadmill at 1.7mph at an incline of 10%, and the speed and incline of the treadmill are increased every 3 minutes, while $\dot{V}O_2$ and heart rate are measured. The test is stopped when the person is too fatigued to walk or run any further. The highest measured $\dot{V}O_2$ is that person's $\dot{V}O_{2max}$.

Aim of current study

Our aim is to perform three repeated maximal treadmill tests following the Bruce Protocol in a healthy male population, to test the accuracy and repeatability of the protocol, and to study the accuracy of equations used to calculate $\dot{V}O_{2max.}$ To take part, please read this document carefully. You should understand the risks and benefits of taking part, so that you can make a decision that is right for you. This is known as "Informed Consent".

Procedures

To participate, you must be available to visit the exercise laboratory in the Trinity Centre for Health Sciences at St James's Hospital on three occasions, each session one week apart. The first session will take approximately 40-60 minutes, and the additional testing sessions will each last 30 minutes. There will be 4 main components on your first visit and 1 component on your additional visits.

1. Physical Activity Questionnaires

Before you begin the exercise test, you will complete a short questionnaire (the Physical Activity Readiness Questionnaire, or PAR-Q). The result will indicate if you can safely complete an exercise test. You will be asked to complete a second questionnaire about your physical activity levels, to help with analysing your exercise test results by relating your fitness level to your current activity levels.

2. Body composition analysis

At your first appointment, the amount of fat, water and muscle in your body will be estimated using a machine (pictured, right) that analyses details of body weight, body mass index (BMI), percentage body fat, muscle mass and fat free mass. This machine is non-invasive and will not cause any pain. You must remove your shoes and socks for this procedure. Your standing height will also be measured at the first session. For each additional visit, only body weight will be re-measured. Please fast from midnight the night before each testing visit, to allow for the most accurate test results.

3. Bruce Protocol: Maximal Exercise Treadmill Test

Before exercise testing begins, your cholesterol and blood sugar levels will be checked with a finger-prick blood test, and your resting blood pressure will be measured. The exercise test process will be explained, and the equipment to be used will be shown to you: a gas analysis system with a face-mask and heart rate monitor (pictured, right).

For each test you will be fitted with a face mask and asked to walk on the treadmill, to follow the Bruce Protocol. It begins at 1.7mph, at a 10% treadmill incline, and gradually increases in speed and incline every three minutes. Your heart rate, blood pressure, and expired breath gases will be monitored throughout the test. Your

blood lactate levels (via a finger prick blood test) will be measured once during each stage of the treadmill test from Stages 3 onwards, until you complete the test. You should continue the exercise test until you feel exhausted and can go no further (i.e. to your maximal capacity). However, you can end the test at any point, if you feel pain or discomfort in your chest, neck, limbs or any other area, if you feel dizzy, too breathless, or as if you might faint. The researcher will monitor for unexpected changes in heart rate and blood lactate throughout testing, and will end the test if these occur, or if there are any technical difficulties with the equipment. If there are no complications, there will be a 5-minute cool-down phase when the test is ended, at a slow pace, while your heart rate returns to its normal pace.

4. Measurement of day-to-day physical activity

You will be asked to wear a small device called an accelerometer (pictured right) for 7 days between your first and second exercise assessment sessions. The accelerometer is about the size of a matchbox and will sit on your belt. You will wear it during waking hours and it will record all your movements throughout the day, such as walking, running, cycling, housework etc. We will provide you with an information leaflet and an activity log sheet to record any activity you took part in while not wearing the accelerometer (e.g. swimming). You will return the accelerometer to us at your following exercise testing session. The accelerometer is not water-resistant and should not be worn in the shower/bath or while swimming.

Benefits

Full analysis of fitness levels and body composition measurements will be provided to you on completion of the testing procedures in the form of an individualised health report. These can help with understanding your own fitness level and how you compare to the general population. The results can be used to guide progression of your own exercise regime. Participating in this study will also benefit this field of research by adding to it.

Risks

Complications during maximal exercise are rare, especially when you have been cleared for exercise and have no medical history that increases your risk during exercise. However, there is a small risk that you could experience:

- pain,
- fatigue,
- dizziness or ataxia (loss of control of voluntary muscle movements),

- difficulty breathing, or wheezing,
- nausea
- leg cramps (claudication), or
- signs of poor blood flow (cyanosis, or pallor)

If any of these occur, the test will be stopped immediately and you will be monitored closely. If symptoms do not subside, you will be seen by a doctor and if necessary, brought to A&E at St. James's Hospital.

Exclusion From Participation

You cannot be in this study if any of the following applies to you:

- Aged less than 18 or over 45 years old;
- Female gender;
- Non-fluent in English;
- Current smoker, or quit within the past 6 months;
- Intellectual or cognitive disability that affects your ability to give informed consent or exercise on a treadmill safely;
- History of a cardiac (such as acute cardiac event, unstable chest pain, irregular heartbeat, cardiac infection), pulmonary (e.g. COPD, asthma) or metabolic (e.g. diabetes; renal disease) condition, or show any major signs or symptoms suggestive of these conditions;
- 2 or more of the following:
    - Family history of cardiac conditions in first-degree relatives;
    - Sedentary lifestyle;
    - BMI $\geq 30 kg \cdot m^{-2}$;
    - High levels of low-density cholesterol (LDL) or low levels of high-density cholesterol (HDL);
    - Pre-diabetes
- Known blood pressure abnormalities (e.g. hypertension/hypotension)
- Any neuromotor, musculoskeletal or rheumatic condition, any inflammatory, autoimmune or allergic conditions, or any chronic infectious disease (e.g. Hepatitis C, HIV/AIDS)
- Medication use
- Musculoskeletal injury in the previous 3 months
- Epilepsy
- Fitted electronic device (e.g. pacemaker)
- Exclusion for any other reason deemed appropriate by the lead investigator.

On the day of testing

If you consent to taking part, the following points explain how to prepare for your testing day.

- Please fast from midnight on the night before your assessment days. Limit your liquid intake while fasting: you may drink water, but refrain from caffeine drinks, energy drinks, and alcohol. Fasting ensures more accurate results of body composition analysis. Please bring a snack and drink to have after each assessment.
- Please wear loose clothes and comfortable shoes that you can exercise in.

- Bring a towel, shower gel, and change of clothes if you wish to shower after testing.
- Please drive or use public transport to get to the testing venue and avoid walking or cycling on the days of your assessments. It will make for more accurate results of your fitness test if you have not done much physical activity prior to testing. Please refrain from strenuous physical activity for 24-hours prior to each assessment.
- We request that you continue your normal physical activity regime for the duration of the study, and not to increase the training intensity or frequency.

Confidentiality and Data Protection

What information about you (personal data) will be collected?

Personal data to be collected in this study will include your gender, age, any relevant past medical history, physical activity levels, body composition, and aerobic fitness levels. This information is needed to analyse how effective and repeatable the Bruce Protocol is. Only the minimal amount of personal data which is relevant for the purpose of this study is collected and used.

Who has access to the collected data?

The data controller (the organisation responsible for keeping your information safe) for this study is Trinity College Dublin. The Lead Investigator and Research Supervisor have undergone training in data protection law and practice, prior to starting this research.

How will your data be stored and protected?

Your name and personal details will not be published and will not be disclosed to anyone outside the research team. All your details and results will be coded with number ID to maintain your confidentiality. Your data will be coded, rather than kept completely anonymous, to allow for comparison of your results from one testing day to the next. All information relating to you in hard-copy form will be stored in a locked filing cabinet in a secure office only accessible by the research team. Information and records in electronic form will be stored on a password-protected PC at the Trinity Centre for Health Sciences. Your study information and results will be retained for 7 years in keeping with good research practice standards and data protection legislation. It will be destroyed after this time (electronic data will be erased, and hard copy forms will be shredded).

Is there any risk with processing and storing your data? What will happen if there is a data breach?

Considering that sensitive personal data relating to your health is involved, in the unlikely event of a data breach which poses a high risk to your privacy rights, it will be reported immediately to the Data Protection Officer, and you will be notified as soon as possible.

How will your data be used, now and in the future?

Your data will be used for health research, which is in the public interest. The information collected in this study will be analysed, and the overall findings of this study may be published in international peer reviewed journals and may be shared at research conferences. The results of the study may be used with your consent for comparative purposes in other studies of a similar nature examining the reliability of the Bruce treadmill protocol. However, your data will remain coded and your personal identifiers will never be published or disclosed to anyone outside of this research team. Your data will only be used for comparative purposes in other studies which have Research Ethics Committee approval.The researchers in this project are bound by our Professional Code of Conduct to maintain confidentiality regarding all data gained during this research.

What are your rights under GDPR?

You have the right to:

- Access your personal data
- Rectify or correct any mistakes with your personal data
- Have your personal data erased or deleted. However, it will not be possible to remove anonymised data
- Data portability (move your personal data from one controller to another)
- Object to the use of your personal data (except where it has already been analysed, or anonymised)

You can exercise these rights by contacting any member of the research team. If you are not satisfied with how your data is being used, you can also lodge a complaint to the Data Protection Commissioner (contact: +353 57 8684800 or +353 (0)761 104 800; https://dataprotection.ie/en/contact).

Compensation

The research team is covered by standard medical malpractice insurance. You will not be paid to take part in this research. This research project is self-funded by Trinity College Dublin.

Voluntary Participation

Participation in this study is fully voluntary. Once you have read this document and have had the study procedures, risks and benefits explained, you will be asked to sign a consent form

before the testing begins. You do not have to take part in this study and should not feel obliged to do so. You may withdraw participation at any time without giving a reason, even if the study has already begun. You will not be penalized or give up any benefits which you had before entering the study. If you do not give consent, or withdraw your consent, no attempt will be made to access your data. If you wish to opt-out at any stage, contact the research team (details below).

Stopping The Study

The research team may stop your participation in the study at any time without your consent.

Permission

This research project was given ethics committee approval from the Faculty of Health Sciences Research Ethics Committee, Trinity College Dublin (e-mail: ethicscommittee@tcd.ie) on 18th June 2019. There is no known personal connection between members of the research team and this ethics committee.

Future Contact

After completion of the study, you may be contacted again by the researchers in relation to your study results, if you express that you would like to receive feedback on them. You may also be contacted again if you have given permission to be contacted for future research.

Further Information

For more information or answers to your questions about the study, your participation in the study, and your rights, please contact the research team:

Principal Investigator/Research Supervisor: Prof. John Gormley, Discipline of Physiotherapy, Trinity College Dublin. Contact: Tel (01) 8962121, E-mail: jgormley@tcd.ie

Lead Investigator: Ms. Kate Macnamara, Research Masters Student, Trinity College Dublin. Contact: Tel (01) 8963613, E-mail: macnamak@tcd.ie

Data Controller: Trinity College Dublin.

For information regarding your rights under data protection law, please contact:

Data Protection Officer, Trinity College Dublin: Contact E-mail: dataprotection@tcd.ie

Appendix 13: Participant Consent Form for Maximal Exercise Testing Study

Trinity College Dublin
Coláiste na Tríonóide, Baile Átha Cliath
The University of Dublin

Participant Consent Form

Study Title: A study of the repeatability of the Bruce Protocol graded treadmill test in measuring maximal $\dot{V}O_2$

Lead Investigator: Ms. Kate Macnamara

Principal Investigator: Prof. John Gormley

Study Information

This study aims to test the repeatability of the Bruce Protocol exercise treadmill test, which is used to measure a person's maximal oxygen consumption ($\dot{V}O_{2max}$), indicating their level of fitness. This study also aims to examine the accuracy of equations used on collected data to calculate $\dot{V}O_{2max}$. Participants will attend the Trinity Centre for Health Sciences on three occasions to complete 3 repetitions of the test Results will be compared to examine the test's repeatability. The first session will last 40-60 minutes, and the following sessions will last approximately 30 minutes.

During the testing, there is a risk of experiencing pain (in chest, limbs, neck), dizziness, nausea, difficulty breathing, fatigue, or legs cramps while exercising intensely. If this occurs, you must inform the investigator. Testing will be stopped immediately and the investigators will assess your condition. If further care is required, you will be accompanied to A&E of St. James's Hospital.

Statement of Participant's Consent

Explicit, informed, voluntary consent to partake in the research study.

Please tick yes or no to each of the following items:

I have read the Participation Information Leaflet (dated 6th September 2019) provided to me, and I have had the opportunity to discuss the study and ask questions. I have received satisfactory answers to all my questions.	YES ☐	NO ☐
I understand I will undertake a research assessment which collects data about my gender, age, relevant past medical history and physical activity levels, as well as the assessment of body composition, blood glucose, cholesterol and lactate levels, and my aerobic fitness by use of a treadmill test (Bruce Protocol). I understand that information about me will be kept private and confidential.	YES ☐	NO ☐
I understand the risks as listed above, regarding pain, dizziness, nausea, fatigue and breathlessness during exercise testing. I am aware of the benefits and risks of this research study.	YES ☐	NO ☐
I understand that I can withdraw from the study at any time, and that I can request at any time that my personal data will be deleted and not be used (except where the data has already been analysed/published, or has been anonymised).	YES ☐	NO ☐
I agree to be contacted by researchers as part of this study.	YES ☐	NO ☐
I freely and voluntarily give my consent to take part in this research study having been fully informed of the risks, benefits and alternatives.	YES ☐	NO ☐

You are now entering a separate part of this consent form, relating to data protection

Explicit, informed, voluntary consent regarding data protection.

I understand that all my personal data will be made non-identifiable (pseudonymised) for this study titled "A study of the repeatability of the Bruce Protocol graded treadmill test in measuring maximal $\dot{V}O_2$", and my personal identifiers will not be shared with anyone outside the research team.	YES ☐	NO ☐
I understand that my personal data will be used for this research study, and non-identifiable data may be published in peer reviewed journals, in presentations, and may be disseminated at conferences.	YES ☐	NO ☐

You are now entering a separate part of this consent form relating to use of your information (personal data) for future studies

Explicit, informed, voluntary consent of storage and future use of personal data/.

I consent that my non-identifiable personal data collected in this study may be used without further consent from me, for comparative purposes in similar studies conducted by Trinity	YES ☐	NO ☐

College Dublin, assessing the reliability of exercise treadmill tests, only if the research is approved by a research ethics committee.		
I understand that my personal data used for comparative study analysis (as above) may be published in peer reviewed journals, in presentations and may be disseminated at conferences but my data will remain confidential and none of my personal identifiers will be disclosed in these circumstances.	YES □	NO □
I understand that I am free to withdraw my consent to the use of my personal data in future similar studies at any time, and my personal data will not be used and will be destroyed (unless the data has been analysed/published, or anonymised).	YES □	NO □
I understand that my data will be stored for a total of 7 years post completion of any future study in compliance with legal and regulatory obligations.	YES □	NO □
Participant's Name (BLOCK CAPITALS)		
Participant's Signature		
Participant's Phone Number		
Date		

To be completed by the **Researcher:**

I have fully explained the purpose and nature of this study, "A study of the repeatability of the Bruce Protocol graded treadmill test in measuring maximal $\dot{V}O_2$" to the participant in a way that he could understand. I have invited him to ask questions on any aspect of the study.	YES □	NO □
I confirm that I have given a copy of the information leaflet and consent form to the participant.	YES □	NO □
Researcher's Name (BLOCK CAPITALS)		
Researcher's Title and Qualifications		
Researcher's Signature		
Date		

Appendix 14: International Physical Activity Questionnaire – Short Form[172]

INTERNATIONAL PHYSICAL ACTIVITY QUESTIONNAIRE

We are interested in finding out about the kinds of physical activities that people do as part of their everyday lives. The questions will ask you about the time you spent being physically active in the **last 7 days**. Please answer each question even if you do not consider yourself to be an active person. Please think about the activities you do at work, as part of your house and yard work, to get from place to place, and in your spare time for recreation, exercise or sport.

Think about all the **vigorous** activities that you did in the **last 7 days**. **Vigorous** physical activities refer to activities that take hard physical effort and make you breathe much harder than normal. Think *only* about those physical activities that you did for at least 10 minutes at a time.

1. During the **last 7 days**, on how many days did you do **vigorous** physical activities like heavy lifting, digging, aerobics, or fast bicycling?

   _____ **days per week**

   ☐ No vigorous physical activities → ***Skip to question 3***

2. How much time did you usually spend doing **vigorous** physical activities on one of those days?

   _____ **hours per day**

   _____ **minutes per day**

   ☐ Don't know/Not sure

Think about all the **moderate** activities that you did in the **last 7 days**. **Moderate** activities refer to activities that take moderate physical effort and make you breathe somewhat harder than normal. Think only about those physical activities that you did for at least 10 minutes at a time.

3. During the **last 7 days**, on how many days did you do **moderate** physical activities like carrying light loads, bicycling at a regular pace, or doubles tennis? Do not include walking.

   _____ **days per week**

   ☐ No moderate physical activities → ***Skip to question 5***

4. How much time did you usually spend doing **moderate** physical activities on one of those days?

_____ **hours per day**

_____ **minutes per day**

☐ Don't know/Not sure

Think about the time you spent **walking** in the **last 7 days**. This includes at work and at home, walking to travel from place to place, and any other walking that you have done solely for recreation, sport, exercise, or leisure.

5. During the **last 7 days**, on how many days did you **walk** for at least 10 minutes at a time?

_____ **days per week**

☐ No walking → ***Skip to question 7***

6. How much time did you usually spend **walking** on one of those days?

_____ **hours per day**

_____ **minutes per day**

☐ Don't know/Not sure

The last question is about the time you spent **sitting** on weekdays during the **last 7 days**. Include time spent at work, at home, while doing course work and during leisure time. This may include time spent sitting at a desk, visiting friends, reading, or sitting or lying down to watch television.

7. During the **last 7 days**, how much time did you spend **sitting** on a **week day**?

_____ **hours per day**

_____ **minutes per day**

☐ Don't know/Not sure

This is the end of the questionnaire, thank you for participating.

Appendix 15: Standardised Verbal Encouragement for Maximal Exercise Testing Study

Time during the test	**Statement Given**
02:30 minutes	"In 30s the treadmill will get a bit faster and steeper."
05:30 minutes	"In 30s the treadmill will get a bit faster and steeper."
08:30 minutes	"In 30s the treadmill will get a bit faster and steeper. Can you show me your effort of breathing on this chart?" *(for RPE measurement)*
09:00 minutes	"Ok place your hand on the rail." *(for blood lactate measurement)*
09:20 minutes	"Way to go [name]!"
09:40 minutes	"Good job [name], keep it up!"
10:00 minutes	"Excellent work [name]."
10:20 minutes	"Keep on going."
10:40 minutes	"Keep pushing to your max, [name]!"
11:00 minutes	"You're doing great!"
11:20 minutes	"Go for as long as you can, [name]!"
11:40 minutes	"In 20s the treadmill will get a bit faster and steeper again. Can you show me your effort of breathing on this chart?"
12:00 minutes	"Ok place your hand on the rail."
12:20	"Way to go [name]!"
12:40	"Good job [name], keep it up!"
13:00	"Excellent work [name]."
13:20	"Keep on going."
13:40	"Keep pushing to your max, [name]!"
14:00	"You're doing great!"
14:20	"Go as long as you can, [name]!"
14:40	"In 20s the treadmill will get a bit faster and steeper again. Can you show me your effort of breathing on this chart?"
15:00	"Ok place your hand on the rail."
Continue as above for each subsequent stage, until the participant indicates they want to stop, or they reach $\dot{V}O_{2max}$.	

Appendix 16: ActiGraph Physical Activity Monitor Instruction Booklet and Activity Diary

Trinity College Dublin
Coláiste na Tríonóide, Baile Átha Cliath
The University of Dublin

Participant Information Leaflet – ActiGraph Activity Monitor

Participant Information Leaflet

Thank you for agreeing to wear the ActiGraph Activity Monitor. The ActiGraph measures your physical activity levels and provides us with information on the about of time you spend engaging in different intensities of activity. The following information leaflet addresses some frequently asked questions. Should you have any queries please contact the Physiotherapy Postgraduate and Research Room at the Trinity Centre for Health Sciences, St. James's Hospital on 01-8963613.

1. How many days do I wear the monitor?

You are requested to wear the activity monitor for **one week (7 days)** during all waking hours of each day.

2. Do I wear the monitor to bed?

No. While this monitor can detect night-time activity if worn, we do not require that information for this study. You can remove the monitor when you are going to sleep at night. Please do record the time at which you took it off, and when you put it back on again, in the diary provided.

3. Do I wear the monitor in the shower?

No. You should remove the monitor during any **water-based activity** such as showering, bathing or swimming. You are requested to record these activities, including the times your take the monitor on and off, in the activity diary provided.

4. Do I need to press any button to start / finish the monitor?

No. The monitor is set-up by the researcher leading your study. You do not have to press any button to activate or stop the monitor.

5. Where on my body is the monitor worn?

The monitor is connected to a flexible strap with a clip. The strap should be worn like a belt around your waist with the monitor sitting at hip level on the right side of your body (see picture). Ensure **the black disk on the side of the monitor is pointing towards your head.** The

strap should not be too tight or too loose. You can adjust the strap size if necessary. You may wear the monitor under or over your clothes.

6. Do I need to charge the monitor during the week?

No. Do not plug the monitor into any power source or connect to any USB cable during the week and this may wipe the data collected.

7. I forgot to wear the monitor – what should I do?

If you forget to wear the activity monitor on a particular day don't worry. Please write down clearly in the activity diary which day you forgot to wear the monitor and just carry on wearing it as normal the following day.

8. What should I do when I finish wearing the activity monitor?

When you finish wearing the monitor, please bring it to your next testing session at the Trinity Centre for Health Sciences to return it to the researcher leading your study. If you cannot attend your next appointment, please contact the research team and arrange return of the monitor with them.

Try not to change your activity levels while wearing the monitor as the aim is to get an idea of <u>normal</u> activity patterns!

Thank you very much for recording your physical activity.

Trinity College Dublin
Coláiste na Tríonóide, Baile Átha Cliath
The University of Dublin

Physical Activity Diary

You are requested to wear your ActiGraph Activity Monitor during **all waking hours.** You will have to remove the activity monitor when you are going to bed or during water-based activities such as showering or swimming. Please record the time you put the activity monitor and the time you take it off in the following activity diary. This record will help us analyse your physical activity data as accurately as possible.

Should you have any further queries please contact Ms. Kate Macnamara at the Physiotherapy Postgraduate and Research Room at the Trinity Centre for Health Sciences, St. James's Hospital on 01-8963613.

Example activity diary:

On Date	On Time	Physical activities completed while wearing monitor	Off Date	Off Time	Activity completed while not wearing monitor
04.10.2017	8.20am	Walk 30 mins to work and 30 mins home	04.10.2017	7.10pm	Shower
04.10.2017	7.30pm	Soccer training x 2 hours	04.10.2017	10.30pm	Sleeping in bed
05.10.2017	8.10am	Run x 5km, walk x 30 mins to work	05.10.2017	10.50pm	Sleeping in bed

Trinity College Dublin
Coláiste na Tríonóide, Baile Átha Cliath
The University of Dublin

Physical Activity Diary

Participant's Study ID: _____

On Date	On Time	Physical activities completed while wearing monitor	Off Date	Off Time	Activity completed while not wearing monitor

Appendix 17: Personalised Health Report for Maximal Exercise Testing Study

Trinity College Dublin
Coláiste na Tríonóide, Baile Átha Cliath
The University of Dublin

Individualised Health & Fitness Report

A study of the repeatability of the Bruce Protocol graded treadmill test in measuring maximal $\dot{V}O_2$

Dates of Testing: _____
_____
_____

Thank you for your recent participation in the study "*A study of the repeatability of the Bruce Protocol graded treadmill test in measuring maximal $\dot{V}O_2$".* The research team has analysed your data and compiled your results. As well as reporting your individual results, **the normal reference ranges for the general population are also given.** These are specific to your age and gender. You can find where your results fall within these ranges to establish your health status for each component of the assessment.

If you have further queries about these results, please feel free to consult with your physiotherapist or a healthcare professional.

General Data

Age	Height	Weight
_____ years	_____ cm	_____ kg

Blood Pressure

Normal resting blood pressure (BP) is below 120/80 mmHg, as defined by new guidelines in 2017 by the American College of Cardiology and the American Heart Association. The first, higher number is the systolic blood pressure. It is the pressure in the blood vessels when the heart is contracting. The second, lower number is the diastolic blood pressure, and is the pressure in-between beats, when the heart is relaxed.

An 'elevated' blood pressure is a systolic pressure 120-129mmHg *and* a diastolic pressure less than 80mmHg. High blood pressure ('hypertension') Stage 1 is systolic pressure 130-139mmHg

or diastolic pressure 80-89mmHg. Hypertension Stage 2 is systolic pressure greater than 140mmHg or diastolic pressure greater than 90mmHg.

Your average blood pressure during your sessions for this study was: _____.

Resting Heart Rate

A normal resting heart rate for a healthy adult is between 60-100 beats per minute (bpm). Generally, a lower resting heart rate can indicate that you have better heart function and cardiorespiratory fitness.

Your average resting heart rate during your sessions for this study was: _____.

Body Composition

Body fat is vital to daily body function. However, it is not the amount you weigh but your percentage body fat that potentially influences your health. Weight alone does not distinguish between fat and lean body tissue (muscle and bone). If you start exercising and don't appear to be losing weight you may in fact be reducing your body fat and replacing it with newly developed muscle mass, which is denser and heavier than fat.

Excessive body fat can increase your risk of developing serious health problems such as high blood pressure, high cholesterol, heart disease, diabetes and cancer. Maintaining a healthy body fat percentage can reduce your risk and help prevent the onset of these conditions. Too little body fat can also be unhealthy. Women's bodies require a higher percentage of body fat to be healthy compared to men.

Body composition can be measured using your height and weight to calculate your Body Mass Index (BMI). It can also be measured using bioelectrical impedance analysis (BIA), which we did with you in the lab on your first session. This method gives values for your fat mass (kg) and body fat percentage, as well as your basal metabolic rate (BMR: the number of calories your body burns while at rest), and your fat free mass (FFM), muscle mass and bone mass, all measured in kg.

Normal Values (Males in your age group)	
BMI ($kg \cdot m^{-2}$)	Underweight <18.5 Normal 18.5-24.9 Overweight 25.0-29.9 Obese ≥ 30
Fat %	Very lean 4.2-7.1 Excellent 7.1-11 Good 11-15.3 Fair 15.3-19.1 Poor 19.1-24.1 Very poor 24.1-33.4
BMR	*No universally accepted norms*
Muscle Mass	
Bone Mass	
FFM	

Your Results	
BMI:	_____ $kg \cdot m^{-2}$
Fat % (Fat mass):	_____% (_____ kg)
BMR:	_____ kcal
Muscle Mass (%):	_____ kg (_____%)
Bone Mass (%):	_____ kg (_____%)
FFM:	_____ kg

Blood Glucose and Cholesterol Levels

Blood glucose is a type of sugar that is in your blood, which comes from the food you eat and is needed by your body to produce energy. Keeping a normal range of blood glucose levels in your blood is important to prevent the development of diabetes.

Cholesterol is a type of fat found in your blood. Normal levels are required to produce hormones – however too high a level can cause the cholesterol to build up in your blood vessels and form a plaque, which can restrict or stop blood flow in the blood vessels. This could lead to a heart attack or stroke if left at high levels for too long. There are different types of cholesterol, and to measure these fully you can get a full blood test done. The finger prick blood test done at your first session measured the Total Cholesterol level.

	Your Results	Normal Values
Blood Glucose	_____ $mmol \cdot L^{-1}$	4-7 $mmol \cdot L^{-1}$
Blood Total Cholesterol	_____ $mmol \cdot L^{-1}$	<5 $mmol \cdot L^{-1}$

Cardiorespiratory Fitness

Cardiorespiratory fitness (CRF) is related to your body's ability to provide oxygen to your skeletal muscles during prolonged physical activity. Higher CRF is related to a lower risk of heart disease, lung cancer, type 2 diabetes, stroke and other diseases. It is also a strong indicator of mortality. The best way to measure is CRF is by determining the maximal volume of oxygen your body can use during exercise ($\dot{V}O_{2max}$). This is measured in $ml \cdot kg^{-1} \cdot min^{-1}$.

$\dot{V}O_{2max}$ can be measured during a maximal exercise test, just as you have completed in your sessions during this study. In some cases, where the gas measuring equipment is not available,

or a person cannot exercise to their maximum, the $\dot{V}O_{2max}$ can be estimated through various equations instead.

Your results from the three exercise tests are as follows:

	$\dot{V}O_{2max}$ (mean $\dot{V}O_2$ in last 30s)	Maximum Heart Rate	Test Duration	Lactate at end of test	$\dot{V}O_{2max}$ Criteria* Achieved?
Test 1	____ ml·kg^{-1}·min^{-1}	____ bpm	____ minutes	____ mmol·L^{-1}	Yes / No
Test 2	____ ml·kg^{-1}·min^{-1}	____ bpm	____ minutes	____ mmol·L^{-1}	Yes / No
Test 3	____ ml·kg^{-1}·min^{-1}	____ bpm	____ minutes	____ mmol·L^{-1}	Yes / No

* $\dot{V}O_{2max}$ criteria: 3 of the following:

- Respiratory exchange ratio ≥1.10
- Post-exercise venous lactate concentration ≥ 9mmol·L^{-1}
- Rating of Perceived Exertion ≥7 on the 0-10 scale
- Plateau in $\dot{V}O_2$ (failure to increase in $\dot{V}O_2$ with rise in exercise intensity)

Normal $\dot{V}O_{2max}$ Values for males in your age group			
	Age: 20-29 years	**Age: 30-39 years**	**Age: 40-49 years**
Superior	≥56 ml·kg^{-1}·min^{-1}	≥54 ml·kg^{-1}·min^{-1}	≥53 ml·kg^{-1}·min^{-1}
Excellent	51-55 ml·kg^{-1}·min^{-1}	48-53 ml·kg^{-1}·min^{-1}	46-52 ml·kg^{-1}·min^{-1}
Good	46-49 ml·kg^{-1}·min^{-1}	44-47 ml·kg^{-1}·min^{-1}	42-45 ml·kg^{-1}·min^{-1}
Fair	42-45 ml·kg^{-1}·min^{-1}	41-43 ml·kg^{-1}·min^{-1}	38-41 ml·kg^{-1}·min^{-1}
Poor	≤41 ml·kg^{-1}·min^{-1}	≤40 ml·kg^{-1}·min^{-1}	≤37 ml·kg^{-1}·min^{-1}

Physical Activity Levels

The recommended physical activity levels for healthy individuals is a minimum of 150 minutes of moderate-intensity aerobic exercise over at least 5 days of the week. Aerobic exercise includes activities such as jogging or running, swimming, ball sports, and brisk walking. Increased physical activity can have many health benefits, such as reduced risk of cardiovascular disease, diabetes and high blood pressure, reduced risk of certain cancers, improved sleep quality, reduced symptoms of anxiety and depression, sharper memory and brain function, and improved quality of life.

You wore an activity monitor for a week during this study. The data showed that over 7 days, you spent:

- _____ minutes in moderate-intensity physical activity
- _____ minutes in vigorous-intensity physical activity.

Thank you for your participation in this research study. If you have any further questions regarding any of your results, or the study itself, please contact the research team, as below.

Principal Investigator/Research Supervisor: Prof. John Gormley, Discipline of Physiotherapy, Trinity College Dublin
Contact: Tel (01) 8962121, E-mail: jgormley@tcd.ie
Lead Investigator: Ms. Kate Macnamara, Research Masters Student, Trinity College Dublin
Contact: Tel (01) 8963613, E-mail: macnamak@tcd.ie

Appendix 18: Raw data for Maximal Bruce Protocol Repeatability Study (Chapter 5)

Appendix 18.1 Baseline characteristics and body composition analysis results

Participant	Age (years)	Height (cm)	Weight (kg) Test 1	Weight (kg) Test 2	Weight (kg) Test 3	BMI ($kg \cdot m^{-2}$)	Basal metabolic rate (kcal)	Fat mass (kg)	fat (%)	Free fat mass (kg)	muscle mass (kg)	bone mass (kg)
1	19	182.0	72.50	72.75	73.35	21.9	1855	10.40	14.30	62.15	59.05	3.10
2	18	183.5	75.20	73.95	74.90	22.3	1985	8.60	11.40	66.60	63.30	3.30
3	23	181.0	79.10	78.70	77.55	24.1	2048	9.40	11.90	69.70	66.25	3.45
4	23	174.0	59.80	59.60	59.40	19.8	1613	4.65	7.80	55.14	52.35	2.75
5	22	190.0	99.70	100.55	99.95	27.6	2488	16.85	16.90	82.85	78.80	4.05
6	19	188.5	80.55	80.20	79.85	22.7	1983	14.50	18.00	66.05	62.75	3.30
7	22	190.0	87.70	87.40	87.30	24.3	2189	14.15	16.10	73.60	69.95	3.60
8	21	178.5	71.60	72.00	72.10	22.5	1822	10.10	14.10	61.50	58.45	3.05
9	21	180.0	70.35	70.15	70.75	21.7	1818	8.85	12.60	61.50	58.45	3.05
10	20	177.0	76.40	78.30	77.25	24.4	1873	13.80	18.10	62.55	59.45	3.15
11	32	175.5	75.75	76.70	75.50	24.6	1773	15.15	20.00	60.60	57.55	3.05
12	20	180.0	66.00	66.90	*	20.4	1729	7.70	11.70	58.30	55.40	2.90
13	28	175.5	69.30	70.20	69.95	22.5	1745	9.35	13.50	59.95	56.95	3.00
14	36	172.5	66.50	67.85	68.10	22.3	1547	13.50	20.30	53.00	50.30	2.70
15	41	170.5	75.25	75.75	*	25.9	1740	15.20	20.20	60.05	57.05	3.00
16	24	174.0	68.35	67.70	68.95	22.6	1877	3.55	5.10	64.80	61.60	3.25
17	38	183.5	89.05	89.60	89.80	26.4	1958	22.50	25.30	66.50	63.20	3.30
18	39	179.5	84.00	84.05	*	26.1	1962	16.55	19.70	67.45	64.10	3.35
19	33	182.0	75.55	75.60	*	22.8	1837	12.25	16.20	63.30	60.15	3.15

BMI = body mass index.

* = test was not completed

Appendix 18.2 Results of the International Physical Activity Questionnaire – Short Form

Participant	Q1	Q2	Q3	Q4	Q5	Q6	Q7
1	3 days/week	1.5 hours/day	2 days/week	1.5 hours/day	7 days/week	1 hour/day	6 hours/day
2	2 days/week	2 hours/day	2 days/week	2 hours/day	7 days/week	Unsure	Unsure
3	5 days/week	1.5 hours/day	7 days/week	30 mins/day	7 days/week	20 mins/day	2 hours/day
4	3 days/week	90 mins/day	no moderate PA	Skip	7 days/week	30 mins/day	3 hours/day
5	3 days/week	1 hour/day	6 days/week	30 mins/day	5 days/week	15 mins/day	3 hours/day
6	No vigorous PA	Skip	1 day/week	30 mins/day	Skip	Skip	10 hours/day
7	2 days/week	1 hour/day	1 day/week	30 mins/day	No walking	Skip	5 hours/day
8	1 day/week	30 mins/day	4 days/week	1 hour/day	7 days/week	1 hour/day	4-5 hours/day
9	1 day/week	1.5 hours/day	no moderate PA	Skip	7 days/week	2 hours/day	8 hours/day
10	4 days/week	1 hour/day	no moderate PA	Skip	7 days/week	45 mins/day	9 hours/day
11	4 days/week	1 hour/day	1 day/week	10 mins/day	3 days/week	20 mins/day	10 hours/day
12	No vigorous PA	Skip	1 day/week	Unsure	7 days/week	3 hours/day	4 hours/day
13	5 days/week	1.5 hours/day	No moderate PA	Skip	3 days/week	20 mins/day	5 hours/day
14	No vigorous PA	Skip	No moderate PA	Skip	7 days/week	1.5 hours/day	12 hours/day
15	2 days/week	1.5 hours/day	3 days/week	1.5 hours/day	5 days/week	1 hour/day	10 hours/day
16	5 days/week	45 mins/day	5 days/week	15 mins/day	7 days/week	50 mins/day	8 hours/day
17	3 days/week	50 mins/day	1 day/week	1.5 hours/day	No walking	Skip	8 hours/day
18	5 days/week	1.5 hours/day	5 days/week	45 mins/day	7 days/week	1 hour/day	4 hours/day
19	7 days/week	1 hour/day	6 days/week	15 mins/day	7 days/week	40 mins/day	12 hours/day

Key: PA = physical activity

Appendix 18.3 ActiGraph activity monitor data relating to time spent in moderate-vigorous physical activity per week

Participants	No. of days/week in Freedson bouts†	No. of minutes/week in Freedson bouts	Did participant meet ≥150 mins MVPA per week?
1	4	175	Yes
2	5	327	Yes
3	4	111	No
4	2	27	No
5	3	169	Yes
6	5	282	Yes
7	5	205	Yes
8	3	145	No
9	5	107	No
10	5	218	Yes
11	4	136	No
12	6	368.5	Yes
13	4	97	No
14	6	160	Yes
15	2	46	No
16	7	521	Yes
17	4	172	Yes
18	2	37	No
19	*	*	*

Key: No. = number; MVPA = moderate to vigorous physical activity

† Freedson bouts relate to MVPA occurring in greater than or equal to 10 minute bouts.

* = data was not collected

Appendix 18.4 Number of days between repeated tests

Participant	Test 1 to Test 2	Test 2 to Test 3
1	14 days	7 days
2	14 days	7 days
3	7 days	21 days
4	7 days	7 days
5	7 days	23 days
6	7 days	7 days
7	7 days	7 days
8	7 days	7 days
9	7 days	7 days
10	10 days	6 days
11	7 days	7 days
12	7 days	*Test 3 cancelled*
13	6 days	7 days
14	7 days	7 days
15	7 days	*Test 3 cancelled*
16	7 days	7 days
17	7 days	7 days
18	7 days	*Test 3 cancelled*
19	7 days	*Test 3 cancelled*

Appendix 18.5 Maximal Bruce protocol test durations (seconds)

Participant	Test 1	Test 2	Test 3
1	835	909	904
2	926	939	946
3	743	738	736
4	759	748	738
5	799	785	770
6	848	825	850
7	908	904	914
8	819	824	797
9	858	880	915
10	646	615	631
11	857	870	861
12	899	910	*
13	743	747	740
14	637	712	715
15	703	713	*
16	1019	1000	999
17	761	773	779
18	793	814	*
19	715	742	*

* = test was not completed

Appendix 18.6 Measured $\dot{V}O_{2max}$ from each of the five data sampling methods

Test 1 Measured $\dot{V}O_{2max}$ ($ml \cdot kg^{-1} \cdot min^{-1}$)					
Participant	**LAST30S**	**HIGH30S**	**BLOCK15S**	**BLOCK30S**	**ROLL15BR**
1	57.42	57.45	57.38	57.01	57.90
2	61.37	62.14	61.58	61.47	62.39
3	51.04	51.24	51.40	50.99	51.28
4	48.75	48.75	48.95	48.76	49.08
5	47.43	49.85	50.63	48.95	50.03
6	55.64	56.34	56.63	56.13	56.56
7	58.82	58.93	60.12	59.13	59.43
8	56.56	57.24	57.23	56.85	57.90
9	55.74	55.86	56.53	55.74	56.82
10	38.16	39.58	39.52	39.23	39.67
11	54.04	54.46	54.65	53.91	55.07
12	61.52	62.24	61.92	61.54	62.33
13	46.95	46.95	47.78	46.95	47.96
14	43.69	44.40	44.81	44.16	44.45
15	51.08	51.38	52.08	51.08	51.86
16	69.81	69.81	71.20	69.94	71.38
17	43.66	44.94	45.77	44.59	45.10
18	55.71	55.71	56.40	55.71	56.23
19	51.03	51.51	51.52	51.02	52.38

- LAST30S = average of last 30-seconds of $\dot{V}O_2$ data at maximal exertion;
- HIGH30S = highest 30-second rolling average during the test;
- BLOCK15S = highest 15-second block average during the test;
- BLOCK30S = highest 30-second block average during the test;
- ROLL15BR = highest 15-breath rolling average during the test.

Test 2 Measured $\dot{V}O_{2max}$ ($ml \cdot kg^{-1} \cdot min^{-1}$)					
Participant	**LAST30S**	**HIGH30S**	**BLOCK15S**	**BLOCK30S**	**ROLL15BR**
1	50.55	53.19	52.88	52.85	53.27
2	60.31	60.69	61.20	60.31	61.21
3	49.06	49.23	49.53	49.02	49.34
4	50.62	50.73	50.72	50.62	50.87
5	44.90	48.22	48.96	48.26	48.26
6	52.55	55.39	54.70	54.67	55.78
7	60.78	60.96	60.85	60.78	60.93
8	54.67	54.98	55.39	54.28	55.17
9	58.71	59.20	59.01	58.71	60.05
10	39.31	39.91	40.04	39.68	40.08
11	53.39	54.32	54.92	54.01	54.96
12	65.37	65.59	66.35	65.37	65.83
13	49.59	49.74	49.83	49.59	50.78
14	41.44	43.71	43.17	42.94	43.83
15	47.04	47.22	47.53	47.04	47.77
16	71.02	71.37	71.38	70.93	72.36
17	47.04	47.40	47.84	47.04	47.65
18	56.78	56.78	56.73	56.52	56.98
19	52.17	52.28	52.61	52.28	52.65

Test 3 Measured $\dot{V}O_{2max}$ ($ml \cdot kg^{-1} \cdot min^{-1}$)					
Participant	**LAST30S**	**HIGH30S**	**BLOCK15S**	**BLOCK30S**	**ROLL15BR**
1	50.81	53.41	53.60	53.16	53.59
2	60.10	60.54	60.72	60.06	61.05
3	46.53	46.92	46.86	46.62	47.04
4	50.63	50.87	50.94	50.63	50.83
5	45.99	48.13	47.14	46.42	47.73
6	56.45	57.98	58.10	57.67	58.24
7	61.42	61.50	61.31	61.16	61.77
8	54.41	55.61	56.28	55.25	56.40
9	58.00	60.52	61.29	59.69	61.97
10	39.65	40.23	40.07	39.47	40.70
11	54.66	55.57	55.25	54.94	56.17
12	*	*	*	*	*
13	50.59	50.59	50.60	50.59	50.87
14	41.68	43.35	43.25	42.96	43.29
15	*	*	*	*	*
16	71.91	71.92	72.39	71.91	72.52
17	44.34	46.63	46.68	46.12	46.76
18	*	*	*	*	*
19	*	*	*	*	*

* = test was not completed

Appendix 18.7 Paired-samples t-test results comparing each data sampling method

Test 1 *p* values (* = significant)									
LAST30S v HIGH30S	**LAST30S v BLOCK15S**	**LAST30S v BLOCK30S**	**LAST30S v ROLL15BR**	**HIGH30S v BLOCK15S**	**HIGH30S v BLOCK30S**	**HIGH30S v ROLL15BR**	**BLOCK15S v BLOCK30S**	**BLOCK15S v ROLL15BR**	**BLOCK30S v ROLL15BR**
0.00127*	0.000039*	0.03305*	0.0000002*	0.00413*	0.000345*	0.000067*	0.00000084*	0.413722	0.00000005*

Test 2 *p* values (* = significant)									
LAST30S v HIGH30S	**LAST30S v BLOCK15S**	**LAST30S v BLOCK30S**	**LAST30S v ROLL15BR**	**HIGH30S v BLOCK15S**	**HIGH30S v BLOCK30S**	**HIGH30S v ROLL15BR**	**BLOCK15S v BLOCK30S**	**BLOCK15S v ROLL15BR**	**BLOCK30S v ROLL15BR**
0.00354*	0.00061*	0.04718*	0.000088*	0.139212	0.0000101*	0.000116*	0.000016*	0.084173	0.0000011*

Test 3 *p* values (* = significant)									
LAST30S v HIGH30S	**LAST30S v BLOCK15S**	**LAST30S v BLOCK30S**	**LAST30S v ROLL15BR**	**HIGH30S v BLOCK15S**	**HIGH30S v BLOCK30S**	**HIGH30S v ROLL15BR**	**BLOCK15S v BLOCK30S**	**BLOCK15S v ROLL15BR**	**BLOCK30S v ROLL15BR**
0.0005003*	0.000806*	0.01094*	0.000112*	0.66962	0.000562*	0.007968*	0.000138*	0.001681*	0.00005*

Appendix 18.8 Mean differences across different sampling methods for each participant

Test 1: Absolute differences between $\dot{V}O_{2max}$ measured through different sampling methods ($ml \cdot kg^{-1} \cdot min^{-1}$)											
Participant	**HIGH30S - LAST30S**	**BLOCK15S - LAST30S**	**BLOCK30S - LAST30S**	**ROLL15BR - LAST30S**	**BLOCK15S - HIGH30S**	**BLOCK30S - HIGH30S**	**ROLL15BR - HIGH30S**	**BLOCK30S - BLOCK15S**	**ROLL15BR - BLOCK15S**	**ROLL15BR - BLOCK30S**	**Mean difference**
1	0.03	0.04	0.41	0.48	0.07	0.44	0.45	0.37	0.51	0.89	0.37
2	0.77	0.22	0.10	1.02	0.56	0.67	0.25	0.12	0.80	0.92	0.54
3	0.20	0.36	0.05	0.24	0.16	0.25	0.04	0.42	0.12	0.30	0.22
4	0.00	0.20	0.02	0.33	0.20	0.02	0.33	0.18	0.13	0.31	0.17
5	2.42	3.19	1.52	2.60	0.77	0.90	0.18	1.68	0.60	1.08	1.49
6	0.70	0.98	0.49	0.91	0.29	0.21	0.22	0.50	0.07	0.43	0.48
7	0.11	1.30	0.31	0.61	1.19	0.21	0.51	0.98	0.68	0.30	0.62
8	0.68	0.67	0.29	1.34	0.00	0.39	0.66	0.39	0.66	1.05	0.61
9	0.13	0.79	0.00	1.09	0.67	0.13	0.96	0.79	0.29	1.09	0.59
10	1.42	1.36	1.07	1.51	0.06	0.35	0.09	0.29	0.16	0.44	0.68
11	0.42	0.61	0.13	1.03	0.19	0.55	0.61	0.74	0.42	1.16	0.59
12	0.72	0.40	0.02	0.81	0.32	0.70	0.09	0.37	0.41	0.79	0.46
13	0.00	0.83	0.00	1.01	0.83	0.00	1.01	0.83	0.18	1.01	0.57
14	0.71	1.12	0.47	0.76	0.41	0.24	0.05	0.65	0.37	0.28	0.51
15	0.30	1.00	0.00	0.78	0.70	0.31	0.48	1.01	0.22	0.78	0.56
16	0.00	1.39	0.13	1.57	1.39	0.13	1.57	1.27	0.17	1.44	0.90
17	1.28	2.11	0.93	1.44	0.83	0.35	0.16	1.18	0.67	0.50	0.94
18	0.00	0.69	0.00	0.52	0.69	0.00	0.52	0.68	0.16	0.52	0.38
19	0.48	0.50	0.01	1.35	0.01	0.49	0.87	0.51	0.85	1.36	0.64
									Mean of all differences		**0.60**

Test 2: Absolute differences between $\dot{V}O_{2max}$ measured through different sampling methods ($ml \cdot kg^{-1} \cdot min^{-1}$)

Participant	HIGH30S - LAST30S	BLOCK15S - LAST30S	BLOCK30S - LAST30S	ROLL15BR - LAST30S	BLOCK15S - HIGH30S	BLOCK30S - HIGH30S	ROLL15BR - HIGH30S	BLOCK30S - BLOCK15S	ROLL15BR - BLOCK15S	ROLL15BR - BLOCK30S	Mean difference
1	2.64	2.33	2.30	2.72	0.31	0.34	0.08	0.03	0.39	0.42	1.16
2	0.38	0.89	0.00	0.90	0.51	0.38	0.52	0.89	0.02	0.91	0.54
3	0.17	0.47	0.04	0.28	0.30	0.21	0.11	0.51	0.20	0.31	0.26
4	0.11	0.10	0.00	0.25	0.01	0.11	0.14	0.10	0.15	0.25	0.12
5	3.32	4.06	3.36	3.36	0.74	0.04	0.04	0.70	0.70	0.00	1.63
6	2.84	2.15	2.12	3.23	0.68	0.71	0.40	0.03	1.08	1.11	1.44
7	0.18	0.07	0.00	0.15	0.12	0.18	0.03	0.07	0.09	0.15	0.10
8	0.31	0.72	0.39	0.50	0.41	0.70	0.20	1.11	0.21	0.90	0.54
9	0.49	0.30	0.00	1.34	0.19	0.49	0.85	0.30	1.04	1.34	0.63
10	0.60	0.73	0.37	0.77	0.13	0.23	0.17	0.36	0.03	0.40	0.38
11	0.93	1.53	0.62	1.57	0.60	0.31	0.64	0.91	0.05	0.96	0.81
12	0.22	0.98	0.00	0.46	0.76	0.22	0.24	0.98	0.52	0.46	0.48
13	0.16	0.25	0.00	1.19	0.09	0.16	1.03	0.25	0.94	1.19	0.52
14	2.27	1.73	1.50	2.39	0.54	0.77	0.12	0.23	0.67	0.90	1.11
15	0.18	0.48	0.00	0.72	0.31	0.18	0.55	0.48	0.24	0.72	0.39
16	0.35	0.36	0.09	1.34	0.01	0.44	0.99	0.44	0.98	1.43	0.64
17	0.36	0.80	0.00	0.60	0.44	0.36	0.25	0.80	0.20	0.60	0.44
18	0.00	0.05	0.27	0.19	0.05	0.26	0.20	0.21	0.25	0.46	0.19
19	0.11	0.44	0.11	0.48	0.33	0.00	0.37	0.33	0.03	0.37	0.26
									Mean of all differences		**0.61**

Test 3: Absolute differences between $\dot{V}O_{2max}$ measured through different sampling methods ($ml \cdot kg^{-1} \cdot min^{-1}$)											
Participant	**HIGH30S - LAST30S**	**BLOCK15S - LAST30S**	**BLOCK30S - LAST30S**	**ROLL15BR - LAST30S**	**BLOCK15S - HIGH30S**	**BLOCK30S - HIGH30S**	**ROLL15BR - HIGH30S**	**BLOCK30S - BLOCK15S**	**ROLL15BR - BLOCK15S**	**ROLL15BR - BLOCK30S**	**Mean difference**
1	2.60	2.79	2.35	2.78	0.19	0.25	0.18	0.43	0.01	0.42	1.20
2	0.44	0.62	0.04	0.95	0.18	0.48	0.51	0.65	0.33	0.99	0.52
3	0.39	0.33	0.09	0.51	0.05	0.30	0.13	0.24	0.18	0.42	0.26
4	0.24	0.31	0.00	0.20	0.07	0.24	0.04	0.30	0.11	0.20	0.17
5	2.14	1.15	0.43	1.74	0.99	1.71	0.40	0.72	0.59	1.31	1.12
6	1.53	1.66	1.22	1.80	0.12	0.31	0.26	0.43	0.14	0.57	0.80
7	0.08	0.12	0.26	0.35	0.19	0.33	0.27	0.14	0.46	0.60	0.28
8	1.20	1.87	0.84	1.99	0.67	0.36	0.79	1.03	0.12	1.15	1.00
9	2.52	3.29	1.69	3.97	0.76	0.83	1.45	1.60	0.68	2.28	1.91
10	0.58	0.42	0.18	1.05	0.16	0.76	0.47	0.60	0.64	1.23	0.61
11	0.91	0.59	0.28	1.51	0.32	0.63	0.60	0.31	0.92	1.23	0.73
12	*	*	*	*	*	*	*	*	*	*	*
13	0.00	0.01	0.00	0.28	0.01	0.00	0.28	0.01	0.27	0.28	0.12
14	1.67	1.57	1.28	1.61	0.10	0.39	0.06	0.29	0.04	0.33	0.73
15	*	*	*	*	*	*	*	*	*	*	*
16	0.01	0.48	0.00	0.61	0.47	0.01	0.60	0.48	0.13	0.60	0.34
17	2.29	2.34	1.78	2.42	0.05	0.50	0.14	0.55	0.09	0.64	1.08
18	*	*	*	*	*	*	*	*	*	*	*
19	*	*	*	*	*	*	*	*	*	*	*
										Mean of all differences	**0.72**

* = test was not completed

Appendix 18.9 Predicted $\dot{V}O_{2max}$ from each of the six prediction equations

Test 1 Predicted $\dot{V}O_{2max}$ ($ml \cdot kg^{-1} \cdot min^{-1}$)						
Participant	**"Bruce 1"**	**"Bruce 2"**	**"Foster"**	**"ACSM2018"**	**"ACSMLee"**	**"Fitmate"**
1	50.64	52.77	50.61	52.37	51.35	50.41
2	55.74	58.50	56.83	52.37	47.71	53.94
3	45.49	46.97	44.10	52.37	50.26	51.27
4	46.38	47.98	45.23	42.49	52.86	46.39
5	48.62	50.50	48.08	52.37	48.94	54.46
6	51.37	53.59	51.52	52.37	45.84	49.00
7	54.73	57.36	55.63	52.37	55.56	58.18
8	49.74	51.76	49.49	52.37	46.04	54.85
9	51.93	54.22	52.21	52.37	49.70	52.74
10	40.06	40.87	37.26	33.38	45.37	40.74
11	51.87	54.15	52.15	52.37	55.99	61.26
12	54.22	56.80	55.02	52.37	47.36	47.47
13	45.49	46.97	44.10	52.37	49.21	45.10
14	39.55	40.30	36.63	42.49	52.45	42.58
15	43.25	44.46	41.25	42.49	44.10	48.14
16	60.94	64.35	62.68	59.87	63.59	69.15
17	46.50	48.11	45.38	52.37	48.56	45.73
18	48.29	50.12	47.65	52.37	55.79	61.12
19	43.92	45.21	42.11	42.49	50.47	46.94

"Bruce 1" = $\dot{V}O_{2max}$ = 6.70 - (2.82 x sex weighting) + (0.056 x Duration) *[men = 1; women = 0; duration in seconds]*

"Bruce 2" = $\dot{V}O_{2max}$ = (3.778 x Duration) + 0.19 *[Duration in minutes]*

"Foster" = $\dot{V}O_{2max}$ = 14.8 - (1.379 x Duration) + (0.451 x $Duration^2$) - (0.012 x $Duration^3$) *[Duration in minutes]*

"ACSM2018" = $\dot{V}O_2$ = 3.5 + (0.2 x speed) + (0.9 x speed x grade) *[speed in $m \cdot min^{-1}$; grade in percentage as decimal; test conducted to 85% HR_{max}, calculate $\dot{V}O_2$ for last completed stage, and extrapolate to age-predicted HR_{max}]*

"ACSMLee" = $\dot{V}O_2$ = 3.5 + (0.1 x speed) + (1.8 x speed x grade) *[speed in $m \cdot min^{-1}$; grade in percentage as decimal; test conducted to 85% HR_{max}, calculate $\dot{V}O_2$ for Stages 1, 2, 3, average the measured HR from last 30s of each stage, extrapolate to age-predicted HR_{max} with line of best fit through $\dot{V}O_2$ and HRs from stages 1, 2, 3]*

"Fitmate" = Linear equation of the line, plotting measured $\dot{V}O_2$ against heart rate in test conducted to 85% HR_{max}, extrapolated to age-predicted HR_{max}

Test 2 Predicted $\dot{V}O_{2max}$ ($ml \cdot kg^{-1} \cdot min^{-1}$)						
Participant	**"Bruce 1"**	**"Bruce 2"**	**"Foster"**	**"ACSM2018"**	**"ACSMLee"**	**"Fitmate"**
1	54.78	57.43	55.70	52.37	48.45	46.21
2	56.46	59.32	57.68	52.37	50.60	50.58
3	45.21	46.66	43.74	52.37	51.61	56.47
4	45.77	47.29	44.45	42.49	47.50	55.52
5	47.84	49.62	47.08	52.37	59.23	57.33
6	50.08	52.14	49.91	52.37	51.48	53.41
7	54.50	57.11	55.36	52.37	56.34	59.52
8	50.02	52.07	49.84	52.37	53.83	52.32
9	53.16	55.60	53.73	52.37	49.83	55.38
10	38.32	38.91	35.13	33.38	47.39	47.32
11	52.60	54.97	53.04	52.37	56.89	59.43
12	54.84	57.49	55.76	52.37	52.36	52.96
13	45.71	47.23	44.38	52.37	48.05	45.97
14	43.75	45.02	41.89	33.38	59.99	38.15
15	43.81	45.09	41.96	33.38	48.57	50.25
16	59.88	63.16	61.54	59.87	62.28	66.87
17	47.17	48.86	46.23	52.37	50.42	44.52
18	49.46	51.44	49.14	59.87	59.54	63.21
19	45.43	46.91	44.02	52.37	60.92	51.12

Test 3 Predicted $\dot{V}O_{2max}$ ($ml \cdot kg^{-1} \cdot min^{-1}$)						
Participant	**"Bruce 1"**	**"Bruce 2"**	**"Foster"**	**"ACSM2018"**	**"ACSMLee"**	**"Fitmate"**
1	54.50	57.11	55.36	52.37	49.65	49.91
2	56.86	59.76	58.14	52.37	47.67	47.45
3	45.10	46.53	43.60	52.37	48.17	45.34
4	45.21	46.66	43.74	42.49	46.51	52.70
5	47.00	48.67	46.02	52.37	56.87	54.83
6	51.48	53.71	51.66	52.37	49.89	53.80
7	55.06	57.74	56.03	52.37	55.79	59.93
8	48.51	50.37	47.93	52.37	52.32	55.27
9	55.12	57.80	56.10	52.37	53.48	54.10
10	39.22	39.92	36.22	42.49	45.92	43.86
11	52.10	54.40	52.42	52.37	46.26	63.06
12	*	*	*	*	*	*
13	45.32	46.79	43.88	52.37	46.56	49.15
14	43.92	45.21	42.11	52.37	49.69	39.21
15	*	*	*	*	*	*
16	59.82	63.09	61.48	52.37	57.66	60.88
17	47.50	49.24	46.66	52.37	48.03	47.79
18	*	*	*	*	*	*
19	*	*	*	*	*	*

* = test was not completed