Field Exercise Tests Used to Measure Physical Fitness and Functional Capacity in Cardiac Rehabilitation

Submitted by

Lisa Chontalle Hanson

B. Physiotherapy (Hons)

A thesis submitted in total fulfilment

Of the requirements for the degree of

Doctor of Philosophy

Department of Rehabilitation, Nutrition and Sport

School of Allied Health

College of Science, Health and Engineering, La Trobe University, Bundoora

La Trobe University

Victoria, Australia

April 2017

Table of Contents

List of Tables

List of Figures

List of Abbreviations

10 m ISWT	10 m incremental shuttle walk test
6MWT	6 Minute Walk Test
ANOVA	Analysis of variance
BMI	Body mass index = $\left({Weight(kg)}/{Height(m)^2}\right)$
CI	Confidence intervals
ES	Effect size
ICC	Intraclass correlation coefficient
LL	Lower limit (as of a confidence interval)
m	metre
M	Mean
M_{diff}	Mean difference
MIC	Minimum important change
mmHG	millimetres of mercury
MOS SF-36	Medical Outcomes Study Short Form 36
ROC	Receiver operating curve
SD	Standard deviation
SDC	smallest detectable change
SD_{diff}	Standard deviation of the difference
SEM	Standard error measurement
UL	Upper limit (as of a confidence interval)

Acknowledgements

I am most indebted to my supervisors, *Dr Nicholas Taylor* and *Dr Helen McBurney,* without your guidance, and generosity of time in the supervision of my candidature, this achievement would not have been possible. *Helen*, I will be always grateful for your encouraging words that led to me enrolling in a Ph.D., and your continued support, expert advice and encouragement. *Nick*, I am incredibly privileged to have had your supervision throughout my candidature, your mentorship has delivered more from me than I thought possible. I thank you both, for the research, professional and personal skills that I have developed from undertaking this task.

I would like to thank the professional staff at the testing sites, in particular the cardiac rehabilitation coordinators who assisted in participant recruitment. I thank the participants who took the time to complete the exercise tests and the questionnaires.

Finally, I would like to thank my friends and family for their support. My parents, *Brian* and *Cheryl*, and my husband, *Ash*, for their unconditional and unwavering support. To my Bendigo friends who enthusiastically celebrated achievements with me, no matter how small. A special thank you to *Holly* who so generously helped with the care of my children during the busy times.

I dedicate this thesis to my three young children, *Ben*, *Charlotte* and *George*, who were born during this candidature.

This work was supported by an Australian Government Research Training Program Scholarship.

Summary

The aim of this thesis was to determine the best field exercise test for use in cardiac rehabilitation for the measurement of physical fitness and functional capacity. To do this, the studies and methods followed the framework for assessing measurement properties proposed by the COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) group. The most common field tests used in cardiac rehabilitation were the six-minute walk test (6MWT) and the 10 m incremental shuttle walk test (10 m ISWT). A series of three empirical studies with a total of 107 participants and a systematic review were completed.

In a population of patients commencing cardiac rehabilitation, the 6MWT demonstrated a high level of systematic error and insufficient retest reliability when up to three tests were performed. The 10 m ISWT was highly reliable when both relative reliability and measurement error were assessed for both consistency and absolute agreement. Evidence supported validity for the 10 m ISWT when compared with other measures of function and the criterion, the symptom-limited exercise test. The 10 m ISWT was responsive to change over an eight-week cardiac rehabilitation program. The test showed acceptable interpretability, the minimal important change across a cardiac rehabilitation program was greater than the smallest detectable change. Support for the measurement properties and interpretability did not change when one or two tests were performed. Results were consistent with the systematic review (n = 78 studies), which synthesised measurement properties of field exercise tests in cardiac rehabilitation.

In conclusion, the 10 m ISWT was found to be the preferred field test for measurement of physical fitness and functional capacity in cardiac rehabilitation, with acceptable measurement properties when a single test is performed. A single test is suitable for ambulant low-intermediate risk patients entering low to moderate intensity cardiac rehabilitation programs.

Statement of Authorship

Except where reference is made in the text of the thesis, this thesis contains no material published elsewhere or extracted in whole or in part from a thesis submitted for the award of any other degree or diploma. No other person's work has been used without due acknowledgment in the main text of the thesis. This thesis has not been submitted for the award of any degree or diploma in any other tertiary institution. All research procedures reported in the thesis were approved by the relevant Ethics Committee or Safety Committee or authorised officer.

Signed ....

Date 28-04-17

List of Publications during Candidature

Hanson, L. C., McBurney, H., & Taylor, N. F. (2012). The retest reliability of the six minute walk test in patients referred to a cardiac rehabilitation programme. *Physiotherapy Research International, 17*, 55-61. doi:10.1002/pri.513

Hanson, L. C., Taylor, N. F., & McBurney, H. (2016). The 10 m incremental shuttle walk test is a highly reliable field exercise test for patients referred to cardiac rehabilitation: a retest reliability study. *Physiotherapy, 102*(3), 243-248. doi:10.1016/j.physio.2015.08.004

Hanson, L. C., McBurney, H. & Taylor, N. F. (2017). Is the 10 m incremental shuttle walk test a useful test of exercise capacity for patients referred to cardiac rehabilitation? *European Journal of Cardiovascular Nursing,* 1474515117721129

Publications have been reprinted, with permission, in Appendix 1.

List of Conference Presentations during Candidature

Hanson, L., McBurney, H. & Taylor, N. (2013), Determining minimal important change in incremental shuttle walk test results in a cardiac rehabilitation population. 23rd Annual Australian Cardiovascular Health and Rehabilitation Association Conference, August 12-14, 2013, MELBOURNE, Victoria.

Hanson, L., McBurney, H. & Taylor, N. (2011), A comparison of the retest reliability of two walk tests in cardiac rehabilitation. 21st Annual Australian Cardiovascular Health and Rehabilitation Association Conference, August 8-10, 2011, PERTH, WA.

Award: Best exercise research paper presented at the 21st Annual Australian Cardiovascular Health and Rehabilitation Conference.

Invitation to present this paper at the Victorian Association of Cardiac Rehabilitation conference, October, 2011, MELBOURNE, Victoria.

Chapter 1: Introduction

1.1 Problem Statement

Cardiac rehabilitation is a secondary prevention service for people with cardiovascular disease, which reduces cardiovascular mortality and is a cost effective use of health care resources. One aim of cardiac rehabilitation is to increase physical activity and exercise and thereby improve physical fitness and functional capacity in the patients who attend. The results of exercise tests are used to make inferences about the health outcomes of physical fitness and functional capacity. In addition, performance during the tests can provide clinicians with important information on the individual response to the test, and exercise tolerance and intolerance. Field exercise testing has been suggested as a feasible alternative to laboratory exercise testing in cardiac rehabilitation, requiring only simple equipment and less intensive supervision. However, for inferences to be made about the usefulness of field exercise testing in cardiac rehabilitation, it is important that the measurement properties of these tests are established. This thesis will identify the most appropriate field exercise test in cardiac rehabilitation based on an evaluation of the measurement properties, interpretability and clinical utility of the tests. This chapter will discuss the role of field exercise tests, and the importance of establishing their measurement properties, and provide the measurement framework to be used in the thesis. In order to provide context, brief descriptions of cardiac rehabilitation programs and evidence of their effectiveness are presented.

1.2 Overview of Cardiac Rehabilitation

Cardiac rehabilitation is an integrated secondary prevention service for people with cardiovascular disease and is available worldwide (Babu et al., 2016; Price, Gordon, Bird, & Benson, 2016). Cardiovascular disease is an umbrella term for coronary heart disease, cerebral vascular accidents and hypertensive disease (Australian Institute for Health and Welfare, 2017). In 1958, the World Health Organisation identified the need to develop cardiac rehabilitation services, and later defined them as programs consisting of any activities required to improve or restore the physical and psychosocial health of cardiac patients so they could return to a meaningful and satisfying role in their community (World Health Organization, 1964; World Health Organization Expert Committee, 1993).

The delivery of cardiac rehabilitation is a priority in countries with a high prevalence of cardiovascular disease (Heran et al., 2011). In Australia in 2014-15, 22% of the national population reported cardiovascular disease, with coronary heart disease being the most common presentation (Australian Institute for Health and Welfare, 2017). Globally, the burden of coronary heart disease is high, and was the leading single cause of global disability-adjusted life years in 2010, accounting for 5% of the 2.490 billon disability-adjusted life years measured in 2010 (Murray et al., 2012). Disability-adjusted life years are a measure of both years of life lost and years lived with a disability; for coronary heart disease 93% of the burden is due to years of life lost and the remaining 7% from living with a disability (Murray et al., 2012).

Cardiovascular disease is the leading cause of death in the world, and in 2008, accounted for over 17.3 million deaths (Mendis, Puska, & Norrving, 2011). Of these

deaths, coronary heart disease was the most common cardiac disease diagnosis (46% of cases for men, and 38% of cases for women) (Mendis et al., 2011). Cardiovascular disease mortality rates are similar across Australia, Canada, continental Europe Ireland, Japan, New Zealand, Scandinavia, United Kingdom (UK), and United States of America (USA), ranging from 76 to 180 deaths per 100,000 people (Mendis et al., 2011). However, mortality rates for coronary heart disease are higher in the Finland, Germany, Ireland, New Zealand, and USA (75-108 per 100,000), than Australia, Japan, UK, and the remaining continental European and Scandinavian countries (12-74 per 100,000) (Mendis et al., 2011). In 2014, in Australia, 29% of deaths were attributable to cardiovascular disease (Australian Institute for Health and Welfare, 2017). The incidence was higher for those living in regional locations compared with major cities (Alston, Allender, Peterson, Jacob, & Nichols, 2017); Aboriginal and Torres Strait Islander peoples compared with non-Indigenous Australians; and the lowest socioeconomic group compared with the highest socioeconomic group (Australian Institute for Health and Welfare, 2017).

The World Health Organisation recommend all patients with cardiovascular disease be referred to and attend cardiac rehabilitation (World Health Organization Expert Committee, 1993). Eligibility criteria have traditionally included those with stable angina, acute coronary syndrome, coronary revascularisation procedures, chronic ischaemic heart disease, cardiomyopathy, and adults with congenital heart disease (World Health Organization Expert Committee 1993). Acute coronary syndrome is an umbrella term for a range of cardiac clinical presentations including ST-segment elevation myocardial infarction (STEMI), non-ST-segment elevation myocardial infarction (NSTEMI) and unstable angina pectoris (Acute Coronary Syndrome

Guidelines Working Group, 2006; Chew, Allan, Aroney, & Sheerin, 2005). The eligibility for cardiac rehabilitation commonly extends to chronic heart failure, valve surgery, cardiac transplant, arrhythmias, and conduction disturbances (Balady et al., 2007; British Association for Cardiovascular Prevention and Rehabilitation, 2012; Clark et al., 2012; National Heart Foundation of Australia & Australian Cardiac Rehabilitation Association, 2004; National Institute for Health and Care Excellence, 2013). Less commonly, patients with diabetes mellitus, peripheral vascular disease and cardiovascular disease risk factors may be offered cardiac rehabilitation (British Association for Cardiovascular Prevention and Rehabilitation, 2012; National Heart Foundation of Australia & Australian Cardiac Rehabilitation Association, 2004; Piepoli et al., 2010). Modern cardiac rehabilitation programs are now required to accommodate a heterogeneous group of people of all ages with a range of cardiovascular diseases along the disease continuum, and with a range of impairments, limitations and restrictions, as well as differing social and other contextual factors.

Cardiac rehabilitation service capacity is unlikely to meet demand for the services. The potential for unmet needs of cardiac patients was demonstrated in Ontario, Canada where there was a large under-supply of cardiac rehabilitation program places (Candido et al., 2011). There were 53,270 cardiac hospital admissions that met the eligibility criteria for cardiac rehabilitation and this number increased to 128,869 when patients with newly diagnosed diabetes mellitus were included. The number of eligible people for cardiac rehabilitation may have further increased if the study had considered non-hospital admission referrals. The capacity for cardiac rehabilitation services in Ontario was 18,087, limiting the service to 34% of eligible patients

admitted to hospital. It is likely that cardiac rehabilitation programs around the world face this problem and it emphasises the need to demonstrate value in cardiac rehabilitation and reduce costs that are not linked to improved patient outcomes (Porter, 2010). In addition, these limits on capacity mean that patients with coronary heart disease are more likely to be offered cardiac rehabilitation than patients with cardiovascular disease with a neurological deficit.

The cost of delivery of cardiac rehabilitation programs and services varies worldwide. Costs are dependent on the program duration, number of sessions per week as well as staff costs. Programs that offer moderate to high intensity exercise incur additional costs associated with risk assessment and stratification, and staff supervision. Programs with doctors as core members of the cardiac rehabilitation team incur additional costs. The cost of cardiac rehabilitation in Australia is similar to that of the UK and New Zealand; but less than the costs of cardiac rehabilitation in Canada and the USA where there is an increase in frequency of sessions, laboratory exercise testing and additional monitoring to accommodate moderate to high intensity exercise (Brodie, Bethell, & Breen, 2006; Goble & Worcester, 1999, p.183). In Australia, in 1999, the estimated cost per patient per cardiac rehabilitation session was estimated to be $34.56 in a publically funded program and $43.35 in a privately funded cardiac rehabilitation program (Goble & Worcester, 1999, p. 185). The cost of laboratory exercise testing was not included in these costings as they are not routinely used to guide exercise prescription in Australian cardiac rehabilitation programs (Globle & Worcester, 1999, p. 183). More recent data from the UK show the average cost per patient treated was £288 and ranged from £82 to £927 (Brodie et al., 2006), with higher costs associated with specialist medical services available within the program.

In an effort to manage the high demands for services, Australian guidelines specifically address methods to minimise program costs without reducing the quality of patient care (Goble & Worcester, 1999, p. *xx*).

1.2.1 Cardiac rehabilitation content.

Globally, the recognised core components of cardiac rehabilitation include an individual baseline assessment, exercise training, and counselling and education for risk factor modification including nutritional, psychosocial and physical activity (Babu et al., 2016; Balady et al., 2007; Buckley et al., 2013; Pavy et al., 2012; Piepoli et al., 2014; Woodruffe et al., 2015). Exercise training is an important intervention in cardiac rehabilitation (Anderson et al., 2016; Balady et al., 2011; Perk et al., 2012; Smith et al., 2011). When exercise is the predominant intervention for cardiac rehabilitation it may be referred to as exercise-based cardiac rehabilitation. However, if exercise is combined with counselling and education, it is referred to as comprehensive cardiac rehabilitation (Heran et al., 2011). Supervised exercise therapy in cardiac rehabilitation most commonly occurs in centre-based settings, such as an acute hospital or rehabilitation facility (Anderson et al., 2016) but can also occur in the home (National Institute for Health and Care Excellence, 2013). If the patient received standard medical and pharmacological care but no additional exercise therapy, counselling or education it is commonly referred to as usual care (Anderson et al., 2016; Taylor, Brown, et al., 2004).

The staffing requirements of a cardiac rehabilitation team vary with the size and level of service offered. The World Health Organisation recommendation for the most basic service is a community health worker, preferably with medical or health training, and

for advanced services, specialist medical, specialist nursing and allied health professionals (World Health Organization Expert Committee, 1993). The allied health professionals that may support cardiac rehabilitation programs in varying capacities include physiotherapists, exercise physiologists, dieticians, psychologists, occupational therapists, social workers, and pharmacists (National Heart Foundation of Australia & Australian Cardiac Rehabilitation Association, 2004; National Institute for Health and Care Excellence, 2013; World Health Organization Expert Committee, 1993). In Australia, outpatient cardiac rehabilitation programs have an appointed program coordinator. The coordinator is responsible for the management of referrals into the program, liaising with the cardiologist, and general practitioner as well the clients and their families, the successful running of the program, coordination of support staff, program evaluation, quality improvement, and program promotion (National Heart Foundation of Australia & Australian Cardiac Rehabilitation Association, 2004). The coordinator is often a specialised nurse, but it can be any health professional. In Australia and the UK, cardiologists are not often involved in cardiac rehabilitation programs, in a national UK survey, only two of the 28 programs coordinators sampled reported direct involvement of a cardiologist (Brodie et al., 2006).

Despite recommendations, participation in cardiac rehabilitation remains low, and the barriers are multifactorial, such as patient factors, provider factors and health care system factors (Thomas et al., 2010). Referral sources for cardiac rehabilitation programs are generally medical: such as general practitioner or cardiologist; another department within the hospital or other hospital or health care services; or in some countries, such as Australia, self-referral. A recent systematic review found that staff

knowledge of cardiac rehabilitation and attitudes towards the patient's likelihood of attendance as well as insufficient resources within health care services meant that referrals to cardiac rehabilitation for all eligible patients did not always occur (Clark et al., 2012). Referral rates were lower for patients who lived greater distances from cardiac rehabilitation programs, were older, female, had more comorbidities, or were from indigenous or ethnic minority groups (Brown et al., 2009; Clark et al., 2012; Redfern et al., 2014; Suaya et al., 2007). Referral rates for patients admitted with acute coronary syndrome without a revascularisation procedure were lower than for patients following a revascularisation procedure such as coronary artery bypass graft surgery, ranging from 53% to 74% (Brown et al., 2009). In Australia and New Zealand, a recent study found that 46% of eligible patients with diagnosed acute coronary syndrome were referred to cardiac rehabilitation (Redfern et al., 2014). This is consistent with other reports of 56% of all eligible patients with acute coronary syndrome with or without a revascularisation procedure (Brown et al., 2009). Suggestions have been made to address this issue, such as the implementation of an automatic referral process with support from a cardiac rehabilitation liaison officer to provide early cardiac rehabilitation education (Ades et al., 2006; Grace et al., 2011; Hutchinson, Meyer, & Marshall, 2015). In Australia, cardiac rehabilitation coordinators for outpatient programs in a hospital setting often hold a dual role of assessing for outpatient programs and providing education to cardiac patients during their inpatient stay.

There is variation in the international literature on the optimum start time, program length, duration and frequency and the intensity of the exercise training of exercise-based or comprehensive cardiac rehabilitation (Price et al., 2016). There is no

international consensus on the wait time for commencement of cardiac rehabilitation (Price et al., 2016), with reported optimal wait times ranging from as soon as possible after discharge from hospital (Goble & Worcester, 1999; National Guideline Clearinghouse, 2013; Van de Werf et al., 2003) to four weeks (United Kingdom Department of Health, 2000). In Australia, there is no limitation on referral to cardiac rehabilitation with respect to days following treatment or diagnosis, meaning that a patient could attend cardiac rehabilitation at any point after his or her diagnosis or medical intervention.

The variation in cardiac rehabilitation program length ranges from as little as three weeks in Germany and some Australian programs to up to 12 months in Austria (Price et al., 2016). The number of sessions per week and the length of exercise sessions also varies. Some guidelines, such as in Australia and the UK, recommend up to three exercise sessions per week, however, the USA, European and Canadian guidelines recommended more than three sessions per week (Price et al., 2016). Aerobic exercise training was included in all international cardiac rehabilitation guidelines, but the intensity of training varied from low to moderate intensity exercise training to moderate to vigorous intensity exercise (Price et al., 2016).

1.2.2 Effectiveness of cardiac rehabilitation.

Cardiac rehabilitation is reported to be a cost effective use of health care resources (Anderson et al., 2016; Jolliffe et al., 2001) and a Class I recommendation supporting attendance at cardiac rehabilitation has been assigned by the American Heart Association, the American College of Cardiology, and the European Society of Cardiology (Anderson et al., 2016; Balady et al., 2011; Perk et al., 2012; Smith et al.,

2011). Recent systematic reviews support the effectiveness of cardiac rehabilitation attendance for those with stable angina pectoris (Anderson et al., 2016; Oldridge, 2012), acute coronary syndrome (Anderson et al., 2016; Lawlor, Filion, & Eisenberg, 2011; Oldridge, 2012), coronary revascularisation interventions (Anderson et al., 2016; Oldridge, 2012), chronic heart failure (Lewinter et al., 2014), heart transplant (Rosenbaum et al., 2016) and combined revascularisation and valve surgery (Goel et al., 2015).

It has been demonstrated consistently that exercise-based cardiac rehabilitation reduces cardiovascular mortality (Anderson et al., 2016; Heran et al., 2011; Oldridge, 2012; Oldridge, Guyatt, Fischer, & Rim, 1988). A recent systematic review of studies published between 1974 and 2014, showed that when compared with no exercise, participants in exercise-based cardiac rehabilitation programs had a reduction in pooled cardiovascular mortality from 10.4 to 7.6% (number needed to treat 37) and a reduction in hospital admissions from 30.7 to 26.1% (number needed to treat 22) (Anderson et al., 2016).

1.3 Outcome Measures of Cardiac Rehabilitation

Outcomes in cardiac rehabilitation are distinct from the care processes or interventions in place to achieve these, and from the biological predictors of future events (Porter, 2010). The tools used to measure health outcomes in cardiac rehabilitation need to have adequate scale width to measure across the phases of recovery of the heterogeneous group of patients eligible for cardiac rehabilitation. Commonly used clinical outcomes relate to risk factors for cardiovascular disease, such as; tobacco use, blood pressure control, lipid control, physical activity habits and

sedentary time, weight, blood sugar levels, depression or depressive symptoms, medication adherence, physical fitness and functional capacity, and health-related quality of life (Thomas et al., 2007). Outcomes either describe an individual or group at a single period in time or describe how the individual or group change over time. The measures used to record these outcomes vary between cardiac rehabilitation programs and the decision of which test to use may be influenced by the resources available to the program, and the practicality of implementing the tests within the program without disruption to patient care (Bergner & Rothman, 1987; Sanderson, Southard, & Oldridge, 2004).

1.3.1 Measurement of physical fitness and functional capacity.

The purpose of measuring physical fitness and functional capacity is to discriminate and describe an individual or group at a point in time, such as prior to commencing cardiac rehabilitation or at the completion of cardiac rehabilitation (Guyatt, Kirshner, & Jaeschke, 1992a, 1992b; Kirshner & Guyatt, 1985; Williams & Naylor, 1992). Exercise tests completed prior to cardiac rehabilitation may be used to screen for decreases in physical fitness and functional capacity, monitor response to exercise, set and implement individualised exercise programs, predict likelihood of future events and prognosis. Repeating the test at the end of a cardiac rehabilitation program provides an opportunity for longitudinal evaluation of performance at an individual or group level (Guyatt et al., 1992a, 1992b; Kirshner & Guyatt, 1985; Williams & Naylor, 1992).

Estimation of physical fitness and functional capacity can occur through the performance of an exercise test, or standardised questionnaires. Selection of the

measurement tool may be dependent on both the resources available to the cardiac rehabilitation program, and the tool having adequate measurement properties in the context that it is being used (Fitzpatrick, Davey, Buxton, & Jones, 1998; Scott, 2011). Specifically the measurement tool should be reliable, valid and responsive for patients attending cardiac rehabilitation, for the purpose of the measurement, for example to measure a treatment effect. Furthermore, in order to make inferences about the improvement of a patient, the minimal important change (MIC), a component of interpretability, must be established (Scott, 2011). Assessment of self-reported physical activity is of low burden to the patient and clinician, but there may be inaccuracies in self-reporting physical activity (Strath et al., 2013).

The outcomes of an exercise test provide a symbolic representation of the physical fitness and functional capacity construct in cardiac rehabilitation (Dybkaer, 2011), which may inform decision-making on the implementation and adjustment to an exercise intervention, tolerance or intolerance to exercise.

1.4 Exercise Testing and the Cardiac Rehabilitation Patient

Laboratory-based exercise tests, such as the symptom-limited exercise test, often use advanced or specialised monitoring techniques and equipment that are not always readily available in a clinical or field environment. Field exercise tests such as the six minute walk test (6MWT) and the 10 m incremental shuttle walk test (10 m ISWT) are performed in the clinical environment with minimal equipment. The use of exercise tests prior to the commencement of a cardiac rehabilitation program varies. In Australia, an exercise test is not considered necessary for patients who are classified as low-moderate risk, on entry to a low to moderate intensity exercise

program (Goble & Worcester, 1999, pp. 64-65; National Heart Foundation of Australia & Australian Cardiac Rehabilitation Association, 2004). For safety of patients classified as high risk or patients attending a high intensity exercise program, a symptom-limited exercise test is recommended but not mandated (Goble & Worcester, 1999; NSW Department of Health, 2006). For the assessment of physical fitness or functional capacity Australian guidelines recommend either a laboratory-based exercise test, a field exercise test such as the 6MWT or 10 m ISWT (Goble & Worcester, 1999, p. 162; NSW Department of Health, 2006) or a validated self-reported physical activity assessment tool (Woodruffe et al., 2015). No data were found on the type of assessment used among Australian cardiac rehabilitation programs. In the UK, of 28 cardiac rehabilitation programs surveyed, 10 performed treadmill laboratory-based exercise tests, five performed the 10 m ISWT, three performed a step test, two performed a 6MWT and eight programs did not use an exercise test to measure physical fitness and functional capacity (Brodie et al., 2006). Eleven of the 28 cardiac rehabilitation programs repeated the same format of exercise test at the end of the program (Brodie et al., 2006).

1.4.1 Laboratory-based exercise tests.

Laboratory-based exercise tests in cardiac patients may be either maximum or submaximum exercise tests. Maximum exercise tests may directly measure oxygen uptake and when this occurs, the test is known as a cardiopulmonary exercise test. In some circumstances, peak oxygen uptake may be estimated using a standard formula rather than directly sampled. The endpoints of a maximum exercise test include attainment of maximum heart rate or oxygen uptake, indicated when the heart rate or oxygen uptake no longer increase with increased increments in workload. A

maximum exercise test becomes a symptom-limited exercise test when the onset of symptoms such as chest pain, undue fatigue or breathlessness necessitates cessation of the test. When the test supervisor introduces an imposed endpoint such as achievement of 70% of the theoretical maximum heart rate the test becomes a submaximal exercise test.

The symptom-limited exercise test and cardiopulmonary exercise test may be a diagnostic or prognostic tool or a tool to estimate physical fitness and functional capacity. When used for diagnosis or prognosis, patients usually undertake the test without the benefit of their usual medications (American Association of Cardiovascular and Pulmonary Rehabilitation, 2013, p. 62). As a diagnostic test, the purpose is to detect exercise-induced myocardial ischemia or other arrhythmias (Bruce & Hornsten, 1969; Jelinek & Lown, 1974) and if performed with direct monitoring of oxygen uptake, the test can differentiate between cardiovascular and respiratory limitations in exercise capacity (Pichurko, 2012). The results of the same test can also be used to grade the severity of the disease, predict risk of morbidity and mortality, as well as the likelihood of both future cardiac events and death within a specified timeframe (American Thoracic Society & American College of Chest Physicians, 2003; Myers et al., 1998; Stelken et al., 1996).

Laboratory tests, such as the symptom-limited exercise test, can also be used as a measure of physical fitness and functional capacity (Bruce & Hornsten, 1969), and when used in longitudinal assessment, as a measure of change in physical fitness and functional capacity over time (Bruce & Hornsten, 1969). When assessing physical fitness and functional capacity, a maximum, symptom-limited exercise test, or submaximum exercise test can be used. When the purpose of the test is to assess

physical fitness and functional capacity, the patient usually completes the test while taking their usual medications and ideally at a time that the patient usually exercises or intends to exercise (American Association of Cardiovascular and Pulmonary Rehabilitation, 2013, p. 62). The gold standard direct measure of aerobic capacity is maximum oxygen consumption (Fletcher et al., 2001; Palange et al., 2007). Direct measure of end of test or peak oxygen consumption has been shown to be reproducible in repeated testing in patients with chronic heart failure when conditions were standardised for time of day and ambient temperature (Lehmann & Kolling, 1996). Where oxygen consumption is not measured, other functional outcomes can be used to make inferences on physical fitness and functional capacity: such as the length of time and maximum stage attained during the exercise test protocol; the maximum heart rate and maximum rating of perceived exertion; and estimate of metabolic equivalent (MET) derived from peak workload.

Laboratory tests for ambulatory patients are typically conducted on a treadmill or cycle ergometer. There are many standardised protocols for treadmill and bicycle symptom-limited exercise tests, with varying stage-to-stage increments and many non-steady state test, as well as tests with variable protocols. It is recommended that the endpoint of the protocol occur within eight to 12 minutes of commencing the test (Balady et al., 2010). A protocol less than six minutes duration has been associated with a non-linear relationship between time and oxygen uptake and more than 12 minutes can introduce non-cardiopulmonary end-points such as orthopaedic reasons or peripheral muscle fatigue (Balady et al., 2010). The Bruce Protocol is one of the most commonly used standardised symptom-limited exercise test protocols, and is completed on a treadmill (American College of Sports Medicine, 2010b, p. 145;

Bruce, Kusumi, & Hosmer, 1973) (Table 1.1). One study reported a correlation of .78 between the duration of the Bruce protocol repeated within four days in 22 patients with coronary heart disease when their usual anti-anginal medication was ceased (Benhorin et al., 1993). Absolute and relative contraindications for laboratory exercise tests are shown in Table 1.2 (American College of Sports Medicine, 2010a; American Thoracic Society & American College of Chest Physicians, 2003; Fletcher G. F. et al., 1995).

Table 1.1

Bruce protocol

Stage of Bruce protocol	Speed m/s	Gradient %
I	0.76	10
II	1.12	12
III	1.52	14
IV	1.88	16
V	2.24	18
VI	2.46	20

Note. adapted from "Maximal oxygen intake and nomographic assessment of functional aerobic impairment in cardiovascular disease," by R. A. Bruce, F. Kusumi and D. Hosmer, 1973, *American Heart Journal, 85*, p. 547. Copyright 1973 by Elsevier.

Laboratory-based exercise tests such as the symptom-limited exercise test may be supervised directly by a cardiologist or the cardiologist may oversee the support staff who directly supervise the exercise test (Myers et al., 2014). Support staff include exercise physiologists, registered nurses and physiotherapists. The American Heart Association recommend that the cardiologist is responsible for the final interpretation of results (Myers et al., 2014). The time elapsed between referral for testing, testing, interpretation of results and results being made available to the staff supervising cardiac rehabilitation programs is likely to be highly variable between testing sites.

Table 1.2

Absolute and Relative Contraindications for Maximum Exercise Testing

Absolute	Relative
Acute myocardial infarction (3-5 days)	Left main coronary stenosis or its equivalent
Unstable angina	Moderate stenotic valvular heart disease
Uncontrolled arrhythmias causing symptoms or haemodynamic compromise	Severe untreated arterial hypertension at rest (> 200 mm Hg systolic, > 120 mm Hg diastolic)
Syncope	Tachyarrhythmia or bradyarrhythmia
Active endocarditis	High-degree atrioventricular block
Acute myocarditis or pericarditis	Hypertrophic cardiomyopathy
Symptomatic severe aortic stenosis	Significant pulmonary hypertension
Uncontrolled heart failure	Advanced or complicated pregnancy
Acute pulmonary embolus or pulmonary infarction	Electrolyte abnormalities
Thrombosis of lower extremities	Orthopaedic impairment that compromising exercise performance
Suspected dissecting aneurysm	
Uncontrolled asthma	
Pulmonary oedema	
Room air desaturation at rest ≤ 85% (exercise with supplemental oxygen)	
Respiratory failure	
Acute non-cardiopulmonary disorder that may affect exercise performance or be aggravated by exercise (i.e., infection, renal failure, thyrotoxicosis)	
Mental impairment leading to inability to cooperate	

American Thoracic Society and American College of Chest Physicians/2003/ATS/ACCP Statement on Cardiopulmonary exercise testing/American Journal of Respiratory and Critical Care Medicine/167/p. 227.
The *American Journal of Respiratory and Critical Care Medicine* is an official journal of the American Thoracic Society.

1.4.2 Field exercise tests.

In cardiac rehabilitation, an alternative to laboratory exercise testing is field exercise testing. The outcome of field tests can infer a level of physical fitness and functional capacity and, if repeated, monitor progress over time for both individual and group, as well as for research, such as on effects of differing exercise regimens. The result of the field exercise test may also assist the therapist in setting exercise levels for a safe and effective cardiac rehabilitation exercise program. While field exercise tests have been used to predict laboratory-based test results, including maximum oxygen consumption (Cahalin, Mathier, Semigran, Dec, & DiSalve, 1996; Cheetham, Taylor, Burke, O'Driscoll, & Green, 2005; Fowler, Singh, & Revill, 2005; Green, Watts, Rankin, Wong, & O'Driscoll, 2001; Guazzi, Dickstein, Vicenzi, & Arena, 2009; Keell, Chambers, Francis, Edwards, & Stables, 1998; Maldonado-Martín et al., 2006; Mandic, Walker, et al., 2013; Morales et al., 1999), they do not provide information on underlying pathophysiology of exercise limitation and do not have a diagnostic or prognostic role (Rejeski et al., 2002).

Both Australian and international cardiac rehabilitation guidelines have included the use of field exercise tests as an option to measure physical fitness and functional capacity (American Association of Cardiovascular and Pulmonary Rehabilitation, 2016; Goble & Worcester, 1999, p. 162; Price et al., 2016). Commonly reported field exercise tests are the 6MWT, and the 10 m ISWT: both corridor walk tests. Other field tests that have been reported in guidelines, although infrequently, are the Chester step test and field cycle ergometry. Walking on level ground is a functional and familiar task to most patients attending cardiac rehabilitation, and may better reflect functional activities than laboratory-based tests (American Thoracic Society, 2002;

Rasekaba, Lee, Naughton, Williams, & Holland, 2009). Patients attending cardiac rehabilitation also report preferring corridor walk tests to treadmill tests (Lipkin, Scriven, Crake, & Poole-Wilson, 1986).

Corridor walk tests are performed in the clinical setting without expensive equipment. Monitoring before and after the exercise test includes measurement of heart rate, blood pressure, oxygen saturation and a rating of perceived exertion of dyspnoea and fatigue. During the corridor walk test, continuous monitoring of heart rate and oxygen saturation are common, and rating of perceived exertion and blood pressure may be monitored at intervals (Holland et al., 2014). The individual cardiovascular response to exercise provides the therapist with additional information on exercise tolerance. The field exercise test may be prematurely stopped if there is evidence of profound desaturation or onset of chest pain, intolerable dyspnoea, claudication pain, excessive sweating, or unsteadiness, excessive heart rate or blood pressure elevation or decreasing heart rate and blood pressure. These stopping criteria are similar to that for laboratory-based exercise tests (Holland et al., 2014).

Field tests, including corridor walk tests, have been shown to elicit peak heart rates and oxygen consumption values similar to symptom-limited exercise tests and cardiopulmonary exercise tests in respiratory patients (Singh et al., 2014), and in some cases for patients with heart disease (Green et al., 2001). For this reason, it is recommended that the absolute and relative contraindications and precautions for maximum exercise testing be followed for field exercise tests (Holland et al., 2014).

1.4.2.1 The 6MWT.

The 6MWT is a self-paced, time-limited field exercise test. The test was originally described for people with pulmonary disease (Butland, Pang, Gross, Woodcock, & Geddes, 1982) as an alternative measure of exercise tolerance to the 12-minute walk test (McGavin, Gupta, & McHardy, 1976). It has since been used in many other clinical settings including cardiac rehabilitation (Guyatt et al., 1985; Hamilton & Haennel, 2000; Harada, Chiu, & Steward, 1999; Zugck et al., 2000).

Since its introduction, attempts have been made to standardise test procedures (American Association of Cardiovascular and Pulmonary Rehabilitation, 2013; American Thoracic Society, 2002; Guyatt et al., 1984; Guyatt et al., 1985; Holland et al., 2014; Lipkin et al., 1986; Roul, Germain, & Bareiss, 1998; Steele, 1996). The most recent standard operating procedures were based on a systematic review of the literature in adults with chronic respiratory conditions (Holland et al., 2014), but it may be possible to utilise the practical recommendations for other groups such as cardiac rehabilitation. Table 1.3 summarises the guidelines for performance in the 6MWT for the most recent standard operating procedures (Holland et al., 2014), as well as 6MWT standardised protocols that are commonly cited in the cardiac literature.

Table 1.3

6MWT Standard Operating Procedures and Commonly Reported Protocols

Operating procedure	Requirements	Instructions	Encouragement	Assessment	Outcome	Repeated testing
AACPR 2013	Environment Indoors Flat quiet corridor, 30 m, markers 3 m Patient Use of gait aid On medication	Standardised Walk as far as possible Permitted to slow down, to stop and to rest Do not run or jog	Time remaining 1 min intervals Encouragement Standardised, 1 min intervals Patient stops to rest Standardised instruction	-	Primary distance	Follow-up repeat testing same time of day
ATS 2002	Environment Indoor or outdoor Flat quiet corridor, ≥ 30 m, markers 3 m Patient Use of gait aid On medications Rest in start position 10 mins	Standardised Walk as far as possible Permitted to slow down, to stop and to rest Do not run or jog	Time remaining 2 min intervals, and 15 seconds before completion Encouragement Standardised, 2 min intervals Patient stops to rest 1 standardised instruction	Pre-assessment HR SpO_2, dyspnoea, fatigue Monitoring during Optional SpO_2, Post-assessment HR, SpO_2, dyspnoea, fatigue, end test symptoms, performance limiters	Primary distance Secondary rest breaks, reason for rest	Repeated Testing generally not needed, if performed utilise a 1 hr rest between Follow-up repeat testing same time of day

Operating procedure	Requirements	Instructions	Encouragement	Assessment	Outcome	Repeated testing
Guyatt 1984	Environment Indoor 33 m corridor	Standardised Walk as far as possible	Time remaining No Unencouraged Group No verbal contact Encouraged Group Standardised, 30 sec intervals	-	Primary distance of 3rd 6MWT	Repeated Testing 3 tests (with 1 and 2 regarded as a practice test)
Guyatt 1985	Environment Indoor 33 m corridor	Standardised Walk as far as possible	Time remaining No Unencouraged Group No verbal contact Encouraged Group Standardised, 30 sec intervals	-	Primary distance of 3rd 6MWT	Repeated Testing 3 tests (with 1 and 2 regarded as a practice test) 30 min rest between tests
Holland 2014	Environment Indoors with temperature control Flat quiet corridor, ≥ 30 m Availability of rapid emergency response	Standardised Walk as far as possible Permitted to slow down, to stop and to rest Do not run or jog	Time remaining 1 min intervals Encouragement Standardised, 1 min intervals Patient stops to rest Standardised encouragement, 30	Pre- assessment SpO_2, HR, dyspnoea and fatigue, BP Monitoring during SpO_2, HR Post assessment SpO_2, HR,	Primary distance Secondary rest breaks including total time stopped, number of stops, average speed, SpO_2 nadir, end test HR, distance as	Repeated Testing ≥1 hr rest Follow-up repeat testing same time of day

Operating procedure	Requirements	Instructions	Encouragement	Assessment	Outcome	Repeated testing
	Patient Use of gait aid On medications		second intervals when $SpO_2 \geq 85\%$	dyspnoea, fatigue, reason for test limitation	percent predicted citing reference equation	
Lipkin 1986	Environment Indoor 20 m corridor	Standardised from McGavin et al. 1976 Walk as far as possible Permitted to slow down, to stop	Time remaining 3 min and 1 min to go Encouragement: Not standardised As needed	-	Primary distance of 3rd 6MWT	Repeated Testing 3 tests (with 1 and 2 regarded as a practice test), 3-4 hours rest
Roul 1998	Environment Flat surface	Standardised Walk as far as possible Pause to rest	Time remaining No Encouraged Group Standardised, 30 sec intervals	-	Primary distance walked calculated from pedometer Secondary symptoms reported	Repeated Testing not necessary (based on 2 tests completed same day, *ICC* .82)
Steele 1996	Environment Quiet area Record ambient temperature ≥ 30 m corridor Patient On medications	Walk as far as possible Permitted to slow down, to stop and to rest	Time 2 min intervals Encouragement Standardised 30 sec intervals	Pre-assessment Dyspnoea, BP, HR, RR Monitoring during SpO_2 first test only	Primary distance of 3rd 6MWT Secondary rest breaks, reason for rest	Repeated Testing 3 tests (with 1 and 2 regarded as a practice test), ≥ 15 min rest between tests

Operating procedure	Requirements	Instructions	Encouragement	Assessment	Outcome	Repeated testing
	2 hours following a meal			Post-assessment Dyspnoea, limiting symptoms		Follow-up same time of day

Note. AACPR = American Association of Cardiovascular and Pulmonary Rehabilitation; ATS = American Thoracic Society; SpO_2 = oxygen saturation measured via pulse oximetry; HR = heart rate; BP = blood pressure.

Most 6MWT procedures require the patient to walk as far as possible in six minutes along a flat corridor with standardised encouragement provided through the test (American Thoracic Society, 2002; Holland et al., 2014). In respiratory populations, the instruction to walk as far as possible may yield a different result to the instruction to walk as fast as possible (Holland et al., 2014; Weir et al., 2013). The corridor should be temperature controlled, free from pedestrian traffic that may interfere with test performance, and be a minimum of 30 m in length (Holland et al., 2014). The track length has been shown to affect results, patients with chronic respiratory disease walked 50 m further on a 30 m track compared with a 10 m track (Beekman et al., 2013) and 13 m further on a continuous track compared with a 30 m track (Bansal et al., 2008). If within patient, between patient, or group comparisons are to be made, track length should remain constant.

Patient motivation may affect test performance. Standard phrases for encouragement may limit this effect. Encouragement has been shown to increase the distance walked in a mixed cardiac and respiratory group by 31 m (Guyatt et al., 1984). A summary of the encouragement provided is included in Table 1.3. The most recent standard operating procedures recommend standardised encouragement at one-minute intervals (Holland et al., 2014).

The main outcome of the test is the distance walked in six minutes measured to the nearest metre or foot (Holland et al., 2014), or less commonly, as a percentage of predicted from aged-matched norms (American Thoracic Society, 2002; Holland et al., 2014). It is widely accepted that in a pulmonary population, a practice test is required for reliable results ((Holland et al., 2014; Troosters, Gosselink, & Decramer, 1999), the number of tests required in a cardiac population remains unknown. When

the test is performed more than once in a single session, the best score or the final test score or the average score of the walk tests may be recorded. The preferred interpretation of the result is unclear in the guidelines. An improvement in the test score is preferably measured by an increase in the absolute distance walked but it can also be reported as a percentage improvement or percentage of predicted score improvement (American Thoracic Society, 2002).

1.4.2.2 10 m incremental shuttle walk test.

The 10 m ISWT was designed for assessment of patients with chronic obstructive pulmonary disease (Singh, Morgan, Scott, Walters, & Hardman, 1992). However, it has since been used in a variety of populations including people with cardiac disease. This test was a modification of an earlier 20 m shuttle run designed for athletes and healthy adults (Léger & Lambert, 1982). Clinical staff, such as physiotherapists, can supervise the test.

The 10 m ISWT requires limited space and equipment to implement. The test requires a flat, indoor 10 m walking track with two cones placed 0.5 m in from both ends of the 10 m circuit; this allowed for a turning circle, as well as chairs at either end of the walking track (Singh et al., 1992). Equipment the clinician requires are; a stop watch, a recording of the test and speaker to play the recording, and appropriate equipment for clinical monitoring of cardiovascular and respiratory observations. Minimum monitoring equipment includes a sphygmomanometer and pulse oximeter.

The 10 m ISWT is a standardised symptom-limited externally paced incremental test. Immediately prior to beginning the test, standardised instructions are played from a

recording. The test is externally paced, with signal beeps from a recording at regular intervals to indicate when the participant should commence the next shuttle. A triple beep sounds to signal progression to the next level and an increase in walk speed. The 10 m ISWT uses proportionally slower speeds than the original 20 m shuttle run (Singh et al. 1992). In level 1, walk speed begins at 0.5 m/s and increases each minute by 0.17 m/s for a maximum of 12 minutes (Table 1.4). In level 12, there are 14 shuttles. Participants can stop the test at any time, otherwise the test is stopped when participants are unable to maintain the required pace or cannot keep going.

Table 1.4

Summary of the 10 m ISWT

10 m ISWT level	Maximum shuttles in level	Maximum cumulative shuttles	Time to complete 1 shuttle (sec)	Maximum speed m/s	Maximum speed km/h	Maximum distance (m)
1	3	3	20.0	0.50	1.8	30
2	4	7	15.0	0.67	2.4	70
3	5	12	12.0	0.83	3.0	120
4	6	18	10.0	1.00	3.6	180
5	7	25	8.6	1.17	4.2	250
6	8	33	7.5	1.33	4.8	330
7	9	42	6.7	1.50	5.4	420
8	10	52	6.0	1.67	6.0	520
9	11	63	5.5	1.83	6.6	630
10	12	75	5.0	2.00	7.2	750
11	13	88	4.6	2.17	7.8	880
12	14	102	4.3	2.30	8.4	1020

Note. Calculations were based on the description of the test provided by Singh et al. (1992).

The main outcome measures from the 10 m ISWT are the number of shuttles (i.e., completed 10 m laps), the distance walked based on the number of completed shuttle

and the time duration. It is usually only necessary to report on one of these outcome measures, and the distance walked based on the number of completed shuttles is the most commonly used outcome. When more than one test is completed in a single session, the best test score is reported. An improvement in the test result is seen by completing a greater number of shuttles, an increase in the distance walked or the time of test.

Information obtained from the 10 m ISWT is used to make inferences on physical fitness and functional capacity, as well as the patient's cardiovascular response to exercise. The results may guide the implementation of an exercise program and monitor change. The 10 m ISWT does not provide diagnostic information or underlying pathophysiological changes.

Corridor walk tests are commonly used in cardiac rehabilitation in Australia. The selection of the most appropriate field test is usually dependent on the resources available, such as equipment, therapist knowledge and performance, and patient preference and ability. Measurement properties of the tests are not commonly considered or even well understood by many clinicians. A test that is too difficult will create floor effects and the result will underestimate the physical fitness and functional capacity, and performance will increase the risk of an adverse response. If the test is not challenging, a ceiling effect will occur, or the test result will not reflect the ability the patient. In addition, if the test is too long factors other than fitness, such as peripheral musculoskeletal fatigue or loss of motivation.

1.5 Theoretical Framework of Measurement Properties

In cardiac rehabilitation, outcomes from the field exercise test are used to describe the physical fitness and functional capacity of an individual or group at one point in time and guide individual exercise prescription. If the exercise test is repeated at the end of cardiac rehabilitation, it may detect change in an individual or group. The field test needs to be appropriate for the heterogeneous group of patients attending cardiac rehabilitation and the cardiac rehabilitation program that may vary in length from a few weeks to a few months. Determining the most appropriate field exercise test for the measurement of physical fitness and functional capacity in cardiac rehabilitation will be based on an evaluation of the measurement properties, interpretability and clinical utility of the tests. The measurement properties of any field exercise test in cardiac rehabilitation include the reliability, validity and responsiveness. The clinical utility includes the practicality and feasibility of the test as well as the patient acceptance of the field exercise test in cardiac rehabilitation. The framework proposed by the COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) group (Mokkink et al., 2010c) will be used in this thesis.

The COSMIN group developed consensus-based standards after a four-round Delphi study on the taxonomy, terminology and definition of measurement properties of patient-reported health outcomes (Mokkink et al., 2010c) and later published a checklist for evaluating methodological quality of studies reporting measurement properties of patient-reported health outcomes (Terwee et al., 2012) with a particular focus on questionnaires. The taxonomy and checklist have been important not only in standardised critical evaluation of studies on measurement properties but also in the design of studies on measurement properties, and these have extended to field

exercise tests, for example the 6MWT in paediatric populations (Bartels, De Groot, & Terwee, 2013). Since the introduction of the taxonomy and checklist, there has been an improvement in the quality of systematic reviews on health-related outcome measurement instruments (Terwee et al., 2016).

Reliability, validity and responsiveness were the three domains of measurement properties proposed by the COSMIN group. The reliability domain of measurement properties refers to the repeatability of the test score and consistency of measurement when real change has not occurred. Two aspects of this domain that are important in describing the reliability of field exercise tests are relative reliability and measurement error, also known as the absolute reliability. The validity domain of measurement properties refers to how well the field exercise test measures the construct of physical fitness and functional capacity. Construct validity and criterion validity are important aspects of the validity of a field exercise test. Responsiveness is a form of longitudinal validity and reflects the degree to which the field exercise test is able to detect change in physical fitness and functional capacity over time, such as over a cardiac rehabilitation intervention.

Interpretability was identified as an important characteristic of field exercise tests and is the degree to which qualitative meaning can be applied to the field exercise test results or the change scores (Mokkink et al., 2010c). The COSMIN group recognised interpretability as an important characteristic of a measurement tool but did not consider it a measurement property (Mokkink et al 2010a). The focus of interpretability is describing the MIC as well as understanding the presence of significant floor and ceiling effects. Participants in cardiac rehabilitation are a heterogeneous group with a broad range of physical abilities. The best test to measure

physical fitness and functional capacity will need to provide a physical challenge to participants with a wide range of fitness abilities, for example, those participants who are barely able to walk with a gait aid, to those able to run comfortably.

The clinical utility and patient acceptance of the test are other characteristics that are important in selecting a best test. Clinical utility includes the feasibility, both cost and time to complete the test and the practicality of the test including the equipment and space, such as corridor length, required to complete the test. Patient acceptance includes the patient comfort and willingness to complete the test as well as whether the patient believes his or her performance in the test is an adequate reflection of their physical fitness and functional capacity, in other words face validity. If a participant believes that a test does not provide an adequate physical challenge the test will lack face validity, similarly a test will lack face validity from a patient's perspective if a patient believes that the test was too difficult and they could not demonstrate their abilities.

1.6 Local Considerations

The research contributing to this thesis was conducted in the City of Greater Bendigo, in the Loddon Mallee Region of Victoria, Australia. The Loddon Mallee region covers 58,961 square kilometres of land in north western Victoria (Department of Health and Human Services, 2014). The City of Greater Bendigo is a large inner regional Victorian city of Australia with an approximate population of 101,000 people (Department of Health, 2013).

Residents living in the local area were at greater risk of cardiac disease than other regions across Australia. An overview of the health of the Greater City of Bendigo compared with the Victorian population using aggregated data from 2008-2012 showed a standardised morbidity ratio for an acute myocardial infarction was 1.33 (National Heart Foundation, 2012). The standardised morbidity ratio was higher for NSTEMI at 1.42, unstable angina at 1.43 and 0.99 for chronic heart failure (National Heart Foundation, 2012). In 2008, 61% of the local population did not meet the minimum guidelines for exercise (National Heart Foundation, 2012).

The studies reported in this thesis were conducted in three departments of two health and human service providers in the Greater City of Bendigo: Bendigo Health, and St John of God Health Care, Bendigo. In 2013, Bendigo Health was a 653 bed, high acuity, teaching regional hospital (Bendigo Health, 2013a) and provided a cardiac rehabilitation service to inpatients, community patients and carers of people diagnosed with heart disease (Bendigo Health, 2013b). The comprehensive cardiac rehabilitation program was conducted in the physiotherapy department of Bendigo Health. A cardiac rehabilitation nurse coordinated the program. The cardiac rehabilitation nurse and a physiotherapist completed baseline and discharge assessments. The physiotherapy assessment included a field exercise test; either an incremental bike test, 6MWT or 10 m ISWT. There were three classes scheduled in a week and patients typically attended one class each week for eight to 12 weeks, the program could be extended for an individual if clinically indicated. Each class consisted of a 60- minute exercise session and a 60- minute education session. Exercise classes offered individualised low to moderate intensity exercise programs with a focus on aerobic training. The cardiac rehabilitation nurse, a physiotherapist,

exercise physiologist and other support staff including allied health assistants and volunteers, supervised the exercise classes.

In 2013, St John of God Health Care, Bendigo Hospital was a 121 bed, private hospital (St John of God Health Care, 2010b), and provided a cardiac rehabilitation service to privately insured outpatients diagnosed with cardiac disease (St John of God Health Care, 2010a). The comprehensive cardiac rehabilitation program was conducted in the Allied Health department, and was coordinated by the cardiac rehabilitation nurse. At baseline, the coordinator completed a initial assessment. Assessment of physical fitness and functional capacity was by subjective questioning, and an exercise test was not usually completed. The program consisted of a 60-minute exercise session and a 60- minute education session. There were three exercise classes scheduled during the week and patients attended one or two classes per week. Classes offered individualised low to moderate intensity exercise programs with a focus on aerobic training. Private health insurance providers covered the cost of the eight to 10 sessions of the program. On completion of these sessions, an application could be made to the private insurer to extend the program for that patient. The cardiac rehabilitation nurse, a physiotherapist or exercise physiologist and other support staff including volunteers, supervised the exercise classes.

Central Victorian Cardiology, located in St John of God Health Care, Bendigo, provided diagnostic cardiology services (CVC Victoria, 2014a). One diagnostic test regularly performed was a stress echocardiogram (CVC Victoria, 2014b). This test requires two appointments; first, for a baseline ultrasound of the heart muscle and valve function, and second, to complete a symptom-limited exercise test immediately followed by a second cardiac ultrasound. The Bruce or modified Bruce protocol was

the preferred symptom-limited exercise test. Test selection was at the discretion of the cardiologist. An echo-cardiographer completed the ultrasound and a medical doctor supervised the symptom-limited exercise test.

Measurement of oxygen uptake, such as during the cardiopulmonary exercise test, was not available to cardiac patients within the Loddon Mallee Region. If measurement of oxygen uptake was required during exercise testing, patients were required to travel to Melbourne, the closest metropolitan city, a distance of 150 km and approximate drive time of two hours.

1.7 Thesis Aims

The primary aim of the research presented in this thesis was to identify the best field exercise test or tests for measurement of physical fitness and functional capacity in cardiac rehabilitation programs. In order to do this, studies were designed to address the ten secondary aims presented in Table 1.5.

The initial stages of the research of this thesis evaluated the measurement properties of the two most common field exercises tests. Initially, the retest reliability was assessed in the form of relative reliability and measurement error. If the results indicated sufficient levels of retest reliability, research continued to review the validity and interpretability of the tests. Validity included criterion validity, construct validity and responsiveness. Following research in the local context, a systematic review was completed to synthesise the available literature on the measurement properties of field exercise tests used in cardiac rehabilitation contexts internationally.

Table 1.5

Summary of Secondary Research Aims and Research Designs

Secondary research aims	Study	Type of research	Study design
1. To determine the relative reliability of the 6MWT in a cardiac rehabilitation population	1	Test retest	Prospective
2. To determine the measurement error of the 6MWT in a cardiac rehabilitation population	1	Test retest	Prospective
3. To determine if the time between repeated tests affects the test retest reliability of the 6MWT	1	Test retest	Prospective
4. To determine the relative reliability of the 10 m ISWT in a cardiac rehabilitation population	2, 3b	Test retest	Prospective
5. To determine the measurement error of the 10 m ISWT in a cardiac rehabilitation population	2, 4	Test retest	Prospective
6. To determine the criterion validity of the 10 m ISWT in patients with diagnosed and treated cardiac disease	3a	Comparison with a gold standard	Prospective
7. To determine the construct validity of the 10 m ISWT when used in cardiac rehabilitation programs	3b	*A-priori* hypothesis testing	Prospective
8. To determine the responsiveness of the 10 m ISWT when used in cardiac rehabilitation	3b	Pre- and post-intervention, *a-priori* hypothesis testing	Prospective
9. To determine the interpretability of the 10 m ISWT when used in cardiac rehabilitation.	3b	Pre- and post-intervention	Prospective

Secondary research aims	Study	Type of research	Study design
10. To synthesise the available evidence on measurement properties of commonly used field exercise tests used in cardiac rehabilitation	4	Systematic review	

1.8 Thesis Overview

Chapter 2 and 3 describes the test retest reliability of the 6MWT and the 10 m ISWT. Chapter 2 reports findings of the first study, a test retest study on the 6MWT in cardiac rehabilitation. This chapter also investigates the effect of the time between test scheduling. Chapter 3 reports findings of the second study and Part B of the third study, a test retest study on the 10 m ISWT in cardiac rehabilitation. In both chapters, retest reliability is expressed two ways, as relative reliability and as measurement error. The chapters follow the framework of the COSMIN group (Mokkink et al., 2010b, 2010c).

Chapter 4 and 5 describes the validity of the 10 m ISWT in cardiac rehabilitation. Chapter 4 presents the results of Part A of the third study, a comparison of the 10 m ISWT with a gold standard. Evidence for both the concurrent criterion validity is reported as well as a framework to determine the evidence for the predictive criterion validity of the 10 m ISWT in cardiac rehabilitation. Chapter 5 presents the results from Part B of the third study, on the evidence for the construct validity and responsiveness of the 10 m ISWT in a cardiac rehabilitation program using *a-priori* hypothesis testing.

Chapter 6 presents the findings of the interpretability of the 10 m ISWT in cardiac rehabilitation. Interpretability is reported as the MIC and calculated in three ways: 95% limit cut-off point, receiver-operating curve (ROC) analysis and predictive modelling.

Chapter 7 presents the synthesis evidence of measurement properties of field exercise tests used in cardiac rehabilitation. This chapter is a systematic review of the literature.

Chapter 8 discusses the findings, makes recommendations regarding the best field exercise test in cardiac rehabilitation, reviews possible reasons for this and the clinical implications of the findings of this thesis for the use and interpretation of field exercise test in cardiac rehabilitation. Strengths and limitations of the research as well as future research direction are included.

Chapter 2: Retest Reliability of the Six Minute Walk Test

2.1 Chapter Aims

This study addressed secondary research aims 1 and 2. The first aim was to determine the evidence for relative reliability of the 6MWT when up to three tests are performed in a mixed cardiac rehabilitation population. The second aim was to determine the evidence for measurement error of the 6MWT when up to three tests are performed in a mixed cardiac rehabilitation population. The results of this chapter have been published (Appendix 1) (Hanson, McBurney, & Taylor, 2012), and are presented in an expanded format in this chapter.

2.2 Introduction

The 6MWT is a field exercise test that can be used to measure physical fitness and functional capacity in patients who attend cardiac rehabilitation, and was described in Chapter 1, Section 1.2.1. If performed at the start of the program, the results can be used to describe physical fitness and functional capacity of an individual or group, and assist in prescription of an individual exercise program. When the 6MWT is repeated at the end of cardiac rehabilitation, the test can provide information on individual and group change. In cardiac rehabilitation, it remains unknown how many tests are required for a reliable test result.

Previous studies have generally provided evidence to support the relative reliability of the 6MWT in cardiac rehabilitation. The reliability coefficients between the first and

second walk have varied from .78 (Guazzi et al., 2009) to .80 (Demers, McKelvie, Negassa, & Yusuf, 2001; Ingle et al., 2005), to .88 (Nogueira, Leal, Pulz, Nogueira, & Filho, 2006), to .93 (Carvalho et al., 2011) to .96 (Cahalin et al., 1996) to .99 (Gary et al., 2004). Over three 6MWTs the reported intraclass correlation coefficient (*ICC*) was .97 (Hamilton & Haennel, 2000). The reliability coefficients used in these studies varied, with the *ICC* being the most common, but generally not well described. Carvalho et al. (2011) used a Spearman's non parametric correlation (rho) and Guazzi et al. (2009) was not clear which correlation coefficient was used. Due to the limited description of the type of *ICC* used in the analysis of the studies it is not possible to determine if results can be generalised beyond the study.

There was limited information describing the measurement error of the 6MWT in cardiac rehabilitation. Between the first and second walk the standard error of measurement (*SEM*) ranged from 15 (Opasich et al., 1998; Pinna et al., 2000) to 19 m (Olper, Cervi, Santi, Meloni, & Gatti, 2011). The limits of agreement or individual 95% confidence intervals in cardiac disease populations were reported in five studies (Ingle et al., 2005; Nogueira et al., 2006; Patrick, 2008; Pulz et al., 2008). The limits of agreement between the first and second walk were wide and varied. The minimal improvement to be 95% of real change above the measurement error, in other words an estimation of the smallest detectable change (SDC), ranged from 48 (Pulz et al., 2008) to 146 (Ingle et al., 2005) m. When an additional 6MWT was performed, the limits of agreement narrowed and the minimal improvement to be 95% confident of real change was 50 m (Nogueira et al., 2006). No studies were found that investigated the measurement error for the group, in other words to determine how much a group

of people with cardiac disease would need to change, on average, to be confident that change exceeded measurement error.

The presence of systematic change between repeated testing has been investigated. One study demonstrated that the difference between walk distance on repeated tests was not statistically significant (Kervio, Ville, Leclercq, Daubert, & Carré, 2004a). Whereas, other researchers have demonstrated a statistically significant improvement between the first and second 6MWT (Gayda, Temfemo, Choquet, & Ahmaidi, 2004; Guyatt et al., 1985; Hamilton & Haennel, 2000; Kristjánsdóttir, Ragnarsdóttir, Einarsson, & Torfason, 2004; Lipkin et al., 1986; Nogueira et al., 2006; Patrick, 2005). Two studies have reported the presence of statistically significant change between the second and third 6MWTs (Guyatt et al., 1985; Hamilton & Haennel, 2000). Guyatt et al. (1985) reported a statistically significant increase in distance walked when the first two 6MWTs were compared with four subsequent 6MWTs in patients with chronic heart failure. A plateau was observed in the final four-6MWT distances, and within-person standard deviation was less than 6% of the mean distance in the final four tests. The researchers concluded that the 6MWT was acceptable after two practice tests were performed, and recommended three 6MWTs were necessary (Guyatt et al., 1985). Hamilton and Haennel (2000) also reported a significant increase in the distance between the second and third 6MWT. In contrast, Patrick (2005), who studied a population with chronic heart failure, and Nogueira et al. (2006), who studied patients one week after a myocardial infarction, did not observe statistically significant change between the second and third test. Kristjánsdóttir et al. (2004) and Lipkin et al. (1986) also reported the final test score as the best test but did not provide between-test comparisons.

In summary, it remains unclear how many 6MWTs are required to yield the best test score that is reliable, and where the amount of systematic and random error is acceptable. In addition, it remains unclear whether the time between tests affects the test score. The preliminary aim of this research was to determine if the time between testing affects the reliability of test scores. Following this the two main aims were: to determine the relative reliability of the 6MWT in a mixed cardiac rehabilitation population when up to three tests are performed; and to estimate the measurement error of the 6MWT in a mixed cardiac rehabilitation population when up to three tests are performed.

2.3 Method

Research used a quantitative design involving repeated testing on groups of participants referred to the cardiac rehabilitation program.

2.3.1 Ethics approval and consent.

Ethics approval was granted by La Trobe University, Faculty of Health Sciences, Faculty Human Ethics Committee (reference number FHEC06/174) and Bendigo Health Human Research Ethics Committee (Reference number 24/2006) (see Appendix 2). All participants enrolled in the investigation read and signed a Participant Information and Consent Form that was approved by the relevant ethics committees (see Appendix 3).

2.3.2 Participants.

2.3.2.1 Eligibility criteria for participants.

All adults with coronary artery disease referred to cardiac rehabilitation irrespective of severity or duration of the condition were eligible to participate in this study. This included those adults with acute coronary syndrome, stable angina pectoris, following revascularisation procedures including coronary artery bypass graft surgery or percutaneous transluminal coronary angioplasty or coronary artery stent; coronary artery disease as diagnosed by angiography, heart valve surgery, chronic heart failure; heart transplantation; following cardiac resynchronisation therapy; or implantable cardioverter defibrillator.

Participants were excluded if they had any condition where exercise would be contraindicated or if they were unable to walk for any neurological or musculoskeletal reason. Participants were also excluded if they had previously completed cardiac rehabilitation or a 6MWT, or presented to cardiac rehabilitation for risk factor reduction, with congenital heart disease, or were children or pregnant. In addition, participants were excluded if limitations in English language production or comprehension skills precluded them from understanding the consent form.

2.3.2.2 Recruitment procedures.

A consecutive series of patients referred to a cardiac rehabilitation program at a single centre who met the eligibility criteria of the study were invited to participate by the cardiac rehabilitation nurse coordinator. Referrals for patients were received from

internal hospital wards, external hospitals, cardiologists, general practitioner or self-referral. The cardiac rehabilitation nurse coordinator provided the initial screen to determine eligibility. Recruitment began in February 2007 and finished in May 2007.

2.3.2.3 Allocation of participants to groups.

After providing consent, 30 participants were randomly allocated by an independent staff member to one of three groups, labelled Group A, Group B, and Group C. The groups represented three different testing protocols. The sequence of allocation to either Group A, Group B or Group C was determined according to a random numbers generator using a concealed method (Altman & Bland, 1999).

2.3.2.4 Sample size.

Sample size estimates were completed according to the method described by Walter, Eliasziw, and Donner (1998), and based on *a-priori* set levels of optimal and minimal acceptable limits of reliability for clinical measurement. For three tests, a minimum of eight people would be needed if a minimum *ICC* level of .6 (P_0) was accepted and the hypothesis that findings from this study would be consistent with the current literature at an *ICC* of .9 (P_1), at a level of significance (α) of .05 and power of .8 (β=.2). For two tests, a minimum of 12 people would be needed if a minimum *ICC* level of .6 (P_0) was accepted and the hypothesis that findings from this study would be consistent with the current literature at an *ICC* of .9 (P_1), at a level of significance (α) of .05 and power of .8 (β=.2).

2.3.3 Procedure.

Testing took place in the physiotherapy department at the participating hospital. All participants completed three 6MWTs prior to the commencement of the cardiac rehabilitation program, labelled Walk 1, Walk 2 and Walk 3 (Table 2.1). On admission, Group A participants (n = 10) completed Walk 1, Walk 2 and Walk 3 in a single session; Group B participants (n = 10) completed Walk 1 and Walk 2 in a single session, and Walk 3 within one week and prior to commencement of cardiac rehabilitation; and Group C participants (n = 10) completed Walk 1 on the first assessment, and within one week Walk 2 and Walk 3 in a single session, prior to commencement of cardiac rehabilitation. No change in either medical or pharmacological management or exercise intervention occurred between Walk 1, Walk 2 and Walk 3. All walk tests were completed under the same conditions and supervised by the same investigator.

Table 2.1

Outline of 6MWT Schedule for Groups

Group	Initial screen	Second review (within one week)
A	Walk 1, Walk 2, Walk 3	-
B	Walk 1, Walk 2	Walk 3
C	Walk 1	Walk 2, Walk 3

2.3.3.1 Baseline demographic information.

Baseline demographic information were collected during the initial assessment and included age, date of birth, gender, height and weight, cardiac diagnosis and

intervention, cardiac signs and symptoms, cardiac medications, relevant medical history and other relevant medications, mobility and use of gait aid. This information, along with the date and time of assessment were recorded on a data-recording sheet (see Appendix 4). From the data-recording sheet, body mass index (BMI) was calculated, where BMI was equal to weight in kilograms divided by the square of height in metres. Participants were grouped according to the following recommendations: underweight (BMI <18.5 kg/m^2), healthy weight (BMI 18.5 to 24.9 kg/m^2), overweight (25.0 to 29.9 kg/m^2) and obese (BMI 30 kg/m^2 or greater) (American College of Sports Medicine, 2010a, p. 63; Kuczmarski & Flegal, 2000). A BMI greater than 40 kg/m^2 was considered morbidly obese (Anderson, Anderson, & Glanze, 1997), this has also been called Class III obesity (American College of Sports Medicine, 2010a, p. 63).

2.3.3.2 Outcome measure: 6MWT.

Participants were allowed a 20-minute seated rest break prior to commencing the first 6MWT of the session. When more than one 6MWT was completed in a single session, participants were given at least a 30-minute seated rest break between tests. This was to ensure that participants had returned to pre-exercise levels for heart rate, blood pressure, oxygen saturation, respiratory rate and rate of perceived exertion (Steele, 1996).

All 6MWTs were completed on an indoor, flat, 20 m track in a corridor in the physiotherapy department at the participating hospital. Cones were placed 0.5 m in from both ends of the 20 m track to allow for a turning circle.

Instructions and encouragement for all tests were standardised according to Guyatt et al. (1984). Participants were instructed to walk from end to end, covering as much ground as they could during the six minutes. Rest breaks were permitted, but participants were encouraged to recommence walking as soon as they were able. Encouragement was provided at 30-second intervals. The researcher faced the participant and alternated two encouraging phrases: *You're doing well* or *Keep up the good work*.

At the end of the test, the researcher called out *Stop* and the distance walked in the test was recorded to the nearest metre. A chair was taken to the participant at the end of the test for a seated recovery and monitoring.

At the completion of each test the participant rated his or her Borg rate of perceived exertion score using the 6-20 scale (Borg, 1990; Borg & Noble, 1973). For safety, heart rate and oxygen saturation measured via pulse oximetry were monitored throughout each test with additional monitoring after test completion until all values had returned to a level within 20% of the initial recordings. This information was recorded on the data-recording sheet (see Appendix 4).

Participants were not able to commence the 6MWT if resting systolic blood pressure was greater than or equal to 200 mmHg or resting diastolic blood pressure was greater than or equal to 110 mmHg (American College of Sports Medicine, 2010a). Criteria for early termination of the test included participant distress, dizziness, angina or onset of severe musculoskeletal pain (Cahalin et al., 1996), failure of the heart rate to increase with exercise, fall in oxygen saturation below 90% (American College of Sports Medicine, 2010a) or attainment of 85% of maximum heart rate (Singh et al.,

1992) using the heart rate reserve method of calculation (American College of Sports Medicine, 2010a, p. 160).

2.3.4 Statistical methods.

Participant characteristics were recorded (age, gender, weight, height, cardiac intervention and other relevant history) to allow description of the study sample. Where relevant, means were expressed followed by standard deviation. Continuous data were checked for normality. In the case of missing values, data were not imputed.

Retest reliability was expressed in two ways: first as relative reliability, as a ratio in the form of the *ICC* (Shrout & Fleiss, 1979) with corresponding confidence intervals (CI), and second as absolute reliability using the standard error of measurement (*SEM*), and the 95% confidence intervals for the individual and the group.

2.3.4.1 Between-group scores and reliability.

Test scores across the three walks for Group A, Group B and Group C were assessed for statistically significant differences and also relative reliability. A one-way Analysis of Variance (ANOVA) was applied to determine if there were any statistically significant differences between the 6MWT distances in Group A, Group B and Group C. Further, a mixed-plot two-way ANOVA was applied to determine if there were any interaction effects between the groups (independent measure) and the three walks completed (repeated measure).

The use of the *ICC* with 95% confidence intervals is recommended for interpreting relative reliability for continuous data (Mokkink et al., 2010b; Shrout & Fleiss, 1979). Unlike the Pearson or Spearman correlation coefficient, the *ICC* can take into account both systematic error between repeated measures and random error or unpredictable variability. The *ICC* model 2 was described by Shrout and Fleiss and assumes that both raters and participants are randomly chosen (Shrout & Fleiss, 1979). Unlike other models such as the model 3, results can be generalised to other raters and subjects with similar characteristics (Shrout & Fleiss, 1979). This model has also been defined by McGraw and Wong (McGraw & Wong, 1996) as *ICC*(C,1) and *ICC*(A,1) when single measures are used where *C* refers to measures for consistence and *A* refers to measures for absolute agreement (McGraw & Wong, 1996). In IBM SPSS version 22 the equivalent model for single measures is the 2-way random effects model, selected for consistency or absolute agreement. The IBM SPSS 2-way random model was used to calculate the *ICC* with the 95% confidence intervals; for both consistency ($ICC_{\text{consistency}}$) and absolute agreement ($ICC_{\text{agreement}}$).

The model 2 $ICC_{\text{consistency}}$ gives a relative index of the between subjects variance to the between subjects plus error variance. The error variance is defined as the residual variance ($\sigma^2_{residual}$) calculated by the interaction between the subjects and raters. The error variance does not include the variability in repetition or practice and it is therefore unaffected by systematic change between time points in test retest reliability (de Vet, Terwee, Mokkink, & Knol, 2011, p. 104). $ICC_{\text{consistency}}$ reflects the relative ranking of participants and is based on the following formula (de Vet, Terwee, Knol, & Bouter, 2006):

$$ICC_{consistency} = \frac{\sigma^2_{subjects}}{\sigma^2_{subjects} + \sigma^2_{residual}}$$

Where $\sigma^2_{subjects}$ is the between subjects variance and is calculated by $\frac{BMS-EMS}{K}$, and the error variance is defined as the residual variance ($\sigma^2_{residual}$) and is equal to the EMS value from the ANOVA table.

The Model 2 $ICC_{\text{agreement}}$ gives a relative index of the between subjects variance to the between subjects plus error variance, where the error variance is defined as the residual variance plus test retest variance. It provides additional information of the absolute agreement of the scores of repeated tests, in other words the absolute variability in repetition, and is based on the following (de Vet et al., 2006):

$$ICC_{agreement} = \frac{\sigma^2_{subjects}}{\sigma^2_{subjects} + (\sigma^2_{observation} + \sigma^2_{residual})}$$

Where $\sigma^2_{subjects}$ is the between subjects variance, $\sigma^2_{residual}$ is the residual variance and $\sigma^2_{observation}$ is the variance due to systematic differences between time points and is calculated by $\frac{OMS-EMS}{n}$ from the ANOVA table. In absolute agreement, the variability in repetition will affect the $ICC_{\text{agreement}}$ score.

The *ICC* can vary between 0 and 1. An *ICC* of 0 indicates the error variance is large relative to the subject variance and it is assumed there is no reliability. Whereas an *ICC* of 1 indicates perfect reliability or agreement and the error variance is negligent. The *ICC* was interpreted with the following guidelines: good reliability was a score

greater than 0.75, moderate reliability between 0.50 and 0.75 and poor reliability less than 0.5 (Portney & Watkins, 2000, p. 565).

2.3.4.2 6MWT distance.

Change between scores for Walk 1, Walk 2 and Walk 3 was assessed for statistical significance with a repeated measures ANOVA. *Post hoc* comparisons (least significant difference) were performed to test for statistically significant differences between the walk tests (i.e., Walk 1 and Walk 2, Walk 1 and Walk 3, Walk 2 and Walk 3).

2.3.4.3 Relative reliability.

The relative reliability refers to the degree that subjects can be distinguished from each other in the presence of measurement error (de Vet et al., 2006; Terwee et al., 2007). The relative reliability of the sample was interpreted using both the $ICC_{consistency}$ and $ICC_{agreement}$ model with 95% confidence intervals using the method described in Section 2.3.4.1. The $ICC_{consistency}$ and $ICC_{agreement}$ of the sample were calculated across the three walks (Walk 1, Walk 2 and Walk 3), and for all possible pairs of walk tests (Walk 1 and Walk 2, Walk 2 and Walk 3, and Walk 1 and Walk 3).

2.3.4.4 Measurement error.

Measurement error reflects the degree of precision of a score or groups of scores and provides a measure of the systematic and random error of a test score not attributed to the construct measured (Mokkink et al., 2010c). For the individual scores, measurement error was derived from the *SEM*. The measurement error around change

scores were expressed in terms of distance walked using the 95% confidence intervals for the group (Taylor, Dodd, & Graham, 2004) and the 95% confidence intervals for the individual, otherwise known as the limits of agreement (de Vet et al., 2006; Terwee et al., 2007) or least significant difference (Tilloston & Burton, 1991).

3.3.4.4.1 Measurement error for an individual score.

The *SEM* estimates, in the unit of measurement, the error component of an observed score (Guilford, 1954, p. 389; Harvill, 1991; Streiner & Norman, 2008, p. 191; Weir, 2005), or more specifically, it is the standard deviation of the distribution of a test score when no change in status or learning effect has occurred (Weir, 2005; Wyrwich, 2004). It can range from 0 to the standard deviation of the observed score. A score of 0 indicates perfect consistency with the true score equal to the observed score, this would occur when the correlation coefficient is equal to 1.0. A *SEM* equal to the standard deviation of the observed score indicates no consistency and would occur when the correlation coefficient equalled 0 (Streiner & Norman, 2008, p. 191).

To estimate measurement error, the SEM was derived using the square root of the error variance (de Vet et al., 2006; de Vet et al., 2011, p. 111; Guilford, 1954, p. 389). Similar to *ICC*, researchers distinguish between the traditional *SEM*, defined as *SEM* for consistency ($SEM_{\text{consistency}}$), and the *SEM* for absolute agreement ($SEM_{\text{agreement}}$) (de Vet et al., 2006; de Vet et al., 2011, p.111). The $SEM_{\text{consistency}}$ was derived using the following equation:

$$SEM_{consistency} = \sqrt{\sigma^2_{residual}}$$

Where $\sigma^2_{residual}$ is the variance due to the interaction between patients and

observations. The $SEM_{consistency}$ as a percentage of the grand mean was also expressed. This provided information on the relative size of the $SEM_{consistency}$.

The $SEM_{agreement}$ takes into account the systematic error in test and retest scores or in other words variability in repetition, and was derived using the following equation (de Vet et al., 2006):

$$SEM_{agreement} = \sqrt{(\sigma^2_{observations} + \sigma^2_{residual})}$$

Where $\sigma^2_{observations}$ is the variance due to systematic difference between the repeated and $\sigma^2_{residual}$ is the variance due to the interaction between subjects and observations. The $SEM_{agreement}$ as a percentage of the grand mean was also expressed. This provided information on the relative size of the $SEM_{agreement}$.

3.3.4.4.2 Measurement error for change scores.

The measurement error around the change scores were calculated for the group and the individual using 95% confidence intervals (Taylor, Dodd, et al., 2004). The 95% confidence intervals estimate the amount of change that would be required to reflect true change over measurement error with 95% confidence. The 95% confidence intervals were calculated for all combinations of pairs of walk tests. Confidence intervals for the group mean scores were calculated using the following equation (Taylor, Dodd, et al., 2004):

$$95\%CI_{group} = M_{diff} \pm \frac{t_{0.975} \times SD_{diff}}{\sqrt{n}}$$

Where M_{diff} is the mean difference of retest minus test scores and takes into account the systematic error, SD_{diff} is the standard deviation of the differences between retest and test scores, n is the number of participants and $t_{0.975}$ is the critical value for t with a two-tailed test at that sample size.

To determine the degree of change required in an individual to be 95% confident of real change (Altman & Bland, 1999) the 95% confidence intervals were recalculated for $n = 1$ (Taylor, Dodd, et al., 2004):

$$95\%\ CI_{individual} = M_{diff} \pm t_{0.975} \times SD_{diff}$$

Data for absolute agreement for the individual were presented graphically by plotting individual mean test and retest scores against the corresponding individual change scores for all combinations of pairs of walk tests as well as the 95% limits of agreement (Bland & Altman, 1986). Change scores outside the 95% limits of agreement were considered real change and change scores within the limits of agreement could not be distinguished from measurement error.

2.4 Results

Thirty consecutive participants who met the eligibility criteria of the study consented to participate in the study (Figure 2.1). All ($n = 30$) participants completed the walk tests after randomisation into one of three groups prior to the commencement of cardiac rehabilitation.

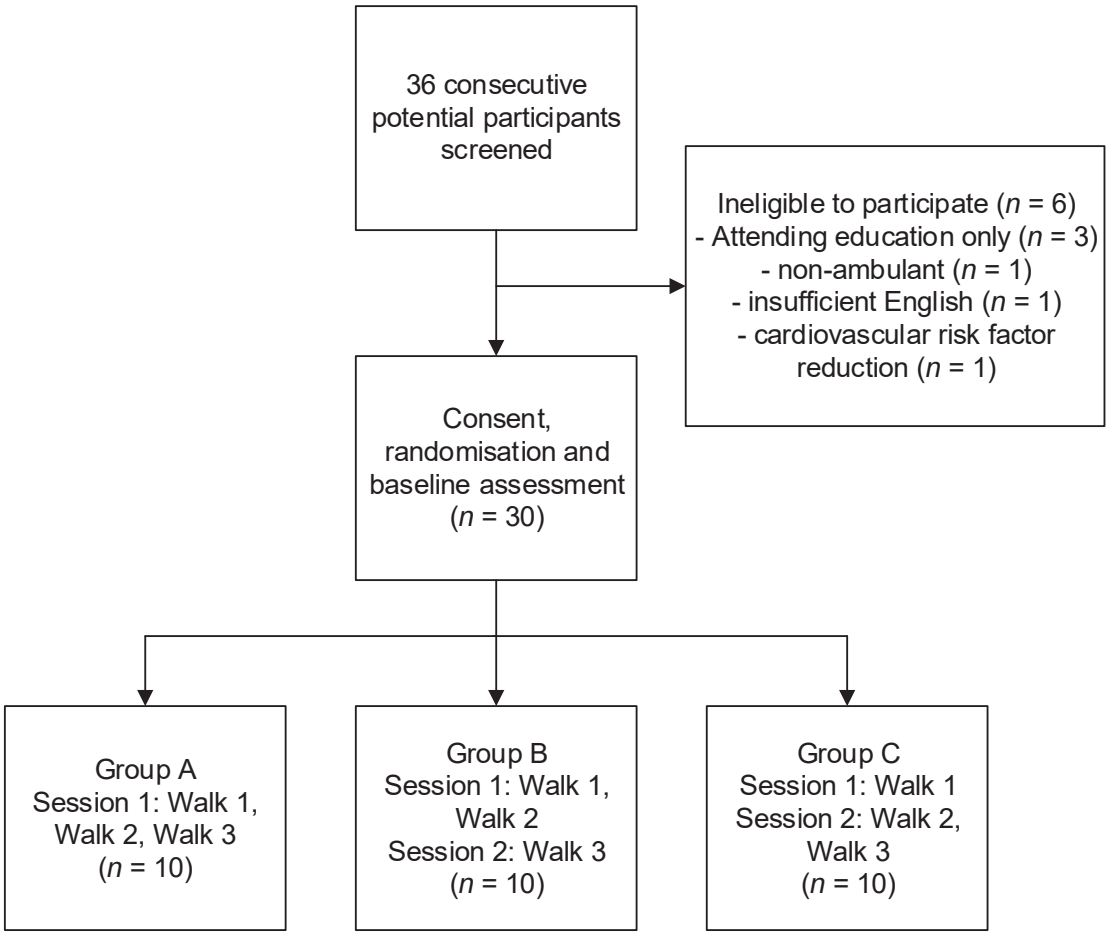


Figure 2.1. Participant recruitment and participation for test retest reliability of the 6MWT.

2.4.1 Characteristics of the participants in the sample.

2.4.1.1 Baseline demographic characteristics of participants.

Table 2.2 shows the baseline demographic characteristics for all participants (n = 30) who completed the study. The mean age of participants was 63 years (*SD* 8) with the youngest participant 49 years and the oldest 79 years. There were more men (n = 24) than women (n = 6). Twenty of the participants underwent a revascularisation procedure prior to commencing cardiac rehabilitation; this included percutaneous and open interventions.

Table 2.2

Baseline Demographic Characteristics of the Sample

Characteristics	Group A *n* = 10	Group B *n* = 10	Group C *n* = 10	Total *n* = 30
Age mean (*SD*)	61 (8)	64 (8)	65 (7)	63 (8)
Gender male:female *n*	7:3	8:2	9:1	24:6
Intervention *n* (%)				
Revascularisation Surgery	6 (60)	5 (50)	9 (90)	20 (67)
Medical management	3 (30)	3 (30)	1 (10)	7 (23)
Heart Valve Surgery	1 (10)	2 (20)	0 (0)	3 (10)
Height mean (*SD*) (cm)	171 (7)	173 (7)	179 (9)	174 (8)
Weight mean (*SD*) (kg)	84 (19)	81 (8)	86 (8)	84 (13)
BMI mean (*SD*) (kg/m^2)	29 (6)	27 (2)	27 (3)	28 (4)
BMI category *n* (%)				
Underweight	1 (10)	0 (0)	0 (0)	1 (3)
Healthy	2 (20)	1 (10)	3 (30)	6 (20)
Overweight	2 (20)	9 (90)	4 (40)	15 (50)
Obese	5 (50)	0 (0)	3 (30)	8 (27)

Note. BMI = body mass index.

Of the participants, 11 (37%) were referred to cardiac rehabilitation following a percutaneous intervention, 10 (33%) after coronary artery bypass graft surgery, six (20%) following commencement of medical management for coronary artery disease, two (7%) following mitral valve replacement and one (3%) following an aortic valve replacement. The mean time elapsed since an acute event was 27 days (*SD* 11). All participants were able to walk independently without a gait aid. No participant had a history of falls.

The mean height of participants was 174 cm (*SD* 8), and the average weight was 84 kg (*SD* 13). There were no statistically significant between-group differences in participant height ($F(2,29) = 2.629$, $p = .091$) or weight ($F(2,29) = .406$, $p = .670$).

The mean BMI for the sample was 27.7 kg/m^2 (*SD* 4.0) indicating that on average the sample was overweight. Six (20%) participants were in their healthy weight range. One participant was underweight, 15 (50%) were overweight and eight (27%) were obese. No participant was categorised as morbidly obese with a BMI greater than 40 kg/m^2. There were no statistically significant differences in BMI between the three groups ($F(2,29) = .402$, $p = .673$).

2.4.1.2 Cardiovascular observations and safety.

All participants completed the three 6MWTs safely without complications, and no test was prematurely stopped. Participants made a full recovery between walk tests, in terms of heart rate, blood pressure, respiratory rate, oxygen saturation and rate of perceived exertion. All cardiovascular observations returned to baseline scores after each walk (Table 2.3). On three occasions, participants required up to five minutes for their heart rate to return to within 20% of baseline.

Table 2.3

Recovery and Symptoms Reported at end of the 6MWT

Recovery and end test symptoms	Walk 1	Walk 2	Walk 3
2 min recovery[a] *n* (%)	28 (93)	29 (97)	30 (100)
Symptoms reported at end of test *n* (%)			
Thoracic pain (non-cardiac)	1 (3)	3 (10)	1 (3)
Fatigue	2 (7)	3 (10)	3 (10)
Lower limb pain	3 (10)	3 (10)	3 (10)
No symptoms	18 (60)	15 (50)	18 (60)
Other	1 (3)	2 (7)	1 (3)
Shortness of breath	5 (17)	6 (20)	4 (13)

Note. [a] 2 min recovery was defined as heart rate within 20% of baseline measure at two minutes.

2.4.2 Between-group differences and reliability.

Mean scores for each 6MWT across Group A, Group B and Group C are shown in Table 2.4. There was a statistically significant increase in the mean distance walked over the three walks (Walk 1, Walk 2 and Walk 3) for each group. *Post hoc* analysis showed that this difference was significant for all walk test score comparisons. However, there were no statistically significant between-group differences for Walk 1 ($F(1,9) = 0.154$, $p = .8$), Walk 2 ($F(1,9) = 0.160$, $p = .85$) or Walk 3 ($F(1,9) = 0.021$, $p = .98$). There were no significant interaction effects ($F(4,54) = 1.67$, $p = .17$).

Across the three walks, good relative retest reliability was shown for Group A, Group B and Group C. The *ICC* model 2 and 95% confidence bands suggest a wide overlap in relative reliability across the three groups (Table 2.4).

As there were no significant between-group differences or interaction effects between the three groups (Group A, Group B and Group C) and similar results for relative reliability for the three groups it was decided that only data for the combined sample ($n = 30$) will be presented.

Table 2.4

The 6MWT Distances Walked at Baseline with Associated ICC

						ICC for consistency			*ICC* for absolute agreement		
	6MWT mean distance (m) (*SD*)						95% CI			95% CI	
Group	Walk 1	Walk 2	Walk 3	*F*(2,18)	*p*	*ICC*	LL	UL	*ICC*	LL	UL
Group A	437 (81)	505 (71)	528 (63)	67.4	< .001	.94	.83	.98	.66	.06	.91
Group B	440 (69)	487 (61)	532 (59)	31.5	< .001	.83	.59	.95	.54	.04	.96
Group C	455 (80)	495 (78)	526 (93)	29.9	< .001	.94	.84	.98	.80	.22	.95

Note. *ICC* = intraclass correlation coefficient; CI = confidence interval; LL = lower limit; UL = upper limit. Group A (n = 10) completed three 6MWTs during the first session; Group B (n = 10) completed two 6MWTs during the first session; Group C (n = 10) completed one 6MWTs during the first session.

2.4.3 6MWT distance.

The mean distance walked improved over the three walks. Participants (n = 30) walked a mean distance of 444 m (*SD* 75) for Walk 1, 496 m (*SD* 68) for Walk 2 and 529 m (*SD* 71) for Walk 3. The difference between the three walks was statistically significant ($F(2,58) = 109.5$ $p < .001$). *Post hoc* analysis showed that this difference was significant for all walk test score comparisons (i.e., Walk 1 and Walk 2; Walk 2 and Walk 3 and Walk 1 and Walk 3).

Twenty nine of the 30 participants walked further in Walk 2 compared with Walk 1, and 29 of the participants walked further in Walk 3 compared with Walk 2 and Walk 3 compared with Walk 1. Fourteen participants walked 50 m or more further in Walk 2 compared with Walk 1, 26 participants walked 50 m or more further in Walk 3 compared with Walk 1, and 7 participants walked 50 m or more further in Walk 3 compared with Walk 2.

2.4.4 Relative test retest reliability.

For the combined sample (n = 30) good relative retest reliability was shown for Walk 1, Walk 2 and Walk 3 with the $ICC_{consistency}$ = .94, 95% CI [.83, .95]. However, when the absolute scores and the systematic error of increasing distance with each walk was taken into account only fair relative retest reliability was shown for Walk 1, Walk 2 and Walk 3 with $ICC_{agreement}$ = .66, 95% CI [.10, .87]. The results for relative reliability for all pairs of walk tests are presented in Table 2.5.

Table 2.5

Relative Test Retest Reliability of the 6MWT across Three Tests

	ICC for consistency			*ICC* for absolute agreement		
		95% CI			95% CI	
6MWT	*ICC*	*LL*	*UL*	*ICC*	*LL*	*UL*
Walk 1, 2 and 3	.94	.83	.95	.66	.10	.87
Walk 1 and 2	.93	.86	.97	.74	0	.93
Walk 1 and 3	.87	.74	.93	.52	0	.84
Walk 2 and 3	.91	.81	.95	.82	.18	.94

Note. ICC = intraclass correlation coefficient; CI = confidence interval; LL = lower limit; UL = upper limit.

2.4.5 Measurement error of the 6MWT in cardiac rehabilitation.

2.4.5.1 Measurement error for an individual score.

The $SEM_{consistency}$ ranged from 19 to 27 m or 4 to 5 per cent of the grand mean. However, when test retest variability or systematic error was accounted for in the measurement error the $SEM_{agreement}$ ranged from 32 to 66 m or 6 to 13% of the grand mean. The results of the *SEM* for all combination of walk tests are shown in Table 2.6.

Table 2.6

Summary of Measurement Error for the Individual Scores

	SEM for consistency		*SEM* for absolute agreement	
6MWT	$SEM_{consistency}$ (m)	% of grand mean	$SEM_{agreement}$ (m)	% of grand mean
Walk 1 and 2	19	4	41	9
Walk 1 and 3	27	5	66	13
Walk 2 and 3	21	4	32	6

Note. SEM = standard error of measurement.

2.4.5.2 Measurement error for the change score.

For the group, a change of at least 63 m would be required as an indication of real change, over and above measurement error with 95% confidence, between the first and the second walk test (Table 2.7). If the first test is regarded as a practice test, a change of at least 45 m would be required for real change, over and above measurement error, with 95% confidence, between the second and third 6MWTs completed before cardiac rehabilitation. A summary of the 95% confidence intervals for the group is presented in Table 2.7.

Table 2.7

Summary of Measurement Error for the Change Scores

		95% CI_{group} (m)		95% $CI_{individual}$ (m)	
6MWT	M_{diff} (m) (SD_{diff})	LL	UL	LL	UL
Walk 1 and 2	52 (26)	42	62	-2	106
Walk 1 and 3	85 (38)	71	99	8	162
Walk 2 and 3	33 (30)	22	44	-29	95

Note. M_{diff} = mean difference; SD_{dif} = standard deviation of the difference; CI = confidence interval; LL = lower limit; UL = upper limit.

For the individual in this population, a change of at least 107 m from the first to the second walk test would be required as an indication of real change, over and above measurement error with 95% confidence. A change of 96 m would be required as an indication of real change, over and above measurement error with 95% confidence, in an individual attending cardiac rehabilitation between the second and third 6MWTs, with the first test being considered a practice test. This is equivalent to walking an additional five laps of a 20 m corridor in six minutes, or an increased mean walking

speed by 0.27 m/sec over the six minutes. A summary of the 95% confidence intervals for the individual is presented in Table 2.7.

Bland-Altman plots for all combinations of pairs of walks are presented in Figure 2.2. Figure 2.2B shows two participants with differences in walk distances outside the upper limit of agreement. The two participants were men aged in their 60s, one in Group B, and the other in Group C. One participant had surgery for aortic regurgitation and one had a percutaneous intervention for coronary revascularisation following a NSTEMI. For Walk 1 and Walk 3 (Figure 2.2C) the difference in distance walked was outside the lower limit of agreement for one participant. This participant was a 72 year old man who had undergone medical management for angina and was allocated to Group B. This participant walked further in Walk 2 compared with Walk 1, however, the distance walked in Walk 3 was less than the distance walked in both Walk 1 and Walk 2.

2.5 Discussion

The purpose of this chapter was to assess if the time elapsed between testing is associated with differences in test scores and reliability and to describe the evidence for retest reliability of the 6MWT as a measure of physical fitness and functional capacity in a mixed cardiac rehabilitation program when up to three tests were performed.

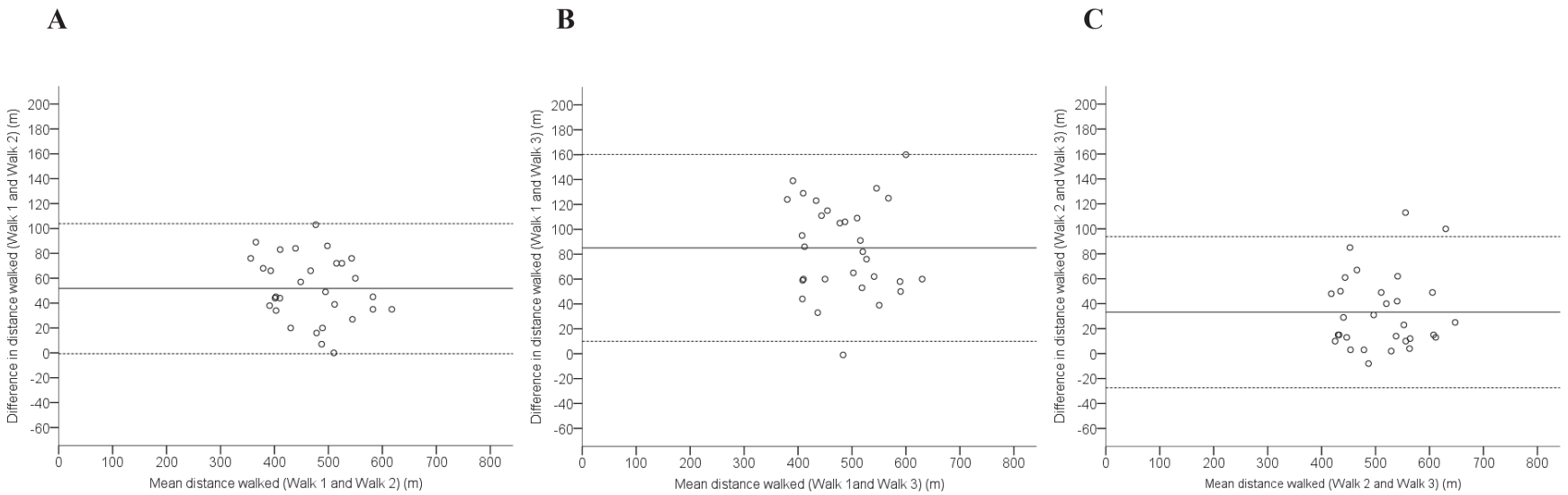


Figure 2.2. Bland Altman plots of the mean 6MWT score plotted against the difference between the retest and test score for (A) Walk 1 and Walk 2; (B) Walk 1 and Walk 2; (C) Walk 2 and Walk 3.

Key. *x*-axis is the mean score; *y*-axis the absolute difference; unbroken line = mean difference; broken line is the limits of agreement calculated by $1.96 \times SD_{diff}$.

This study provided evidence that 6MWT scores and reliability were not affected by the time elapsed between testing when the 6MWTs were completed in a one week time period and prior to the commencement of cardiac rehabilitation. No statistically significant difference in walk distance or interaction effects were found between the group that completed three tests in a single session, the group that completed two tests in the first session and a follow-up test within one week, and the group that completed one test in the first session and two follow-up tests within one week. This was the first study to investigate the effect of patterns of testing on 6MWT distance and reliability. The findings suggest that, if the 6MWT was to be used as a measure of physical fitness and functional capacity, the timing of testing prior to the commencement of cardiac rehabilitation could be scheduled to meet the needs of the patient and the clinician, without affecting test results.

Participant performance, measured by 6MWT distance, continued to increase over the three 6MWTs and these improvements were statistically significant. These results are consistent with findings by two earlier studies reporting on elderly patients with cardiac disease (Gayda et al., 2004) and patients attending cardiac rehabilitation with mild disease (Hamilton & Haennel, 2000). A potential learning effect provides one explanation for the systematic error observed over the three walks (Wu, Sanderson, & Bittner, 2003). Self-paced tests, such as the 6MWT are likely to be influenced by motivational and volitional factors (Rasekaba et al., 2009). It is possible that these factors along with anxiety or fear related to cardiac diagnosis and physical fitness, resulted in poorer initial test performances. Participants may have been gradually testing their walking abilities, and were willing to further exert themselves only after

the successful completion and recovery of a 6MWT. In addition, the 'warm up' effect of the initial tests cannot be ruled out.

The statistically significant improvements over repeated walk tests has not been consistently demonstrated across all cardiac rehabilitation groups, such as a chronic heart failure population (Patrick, 2005). It is possible that this specific population responded differently compared with a mixed cardiac rehabilitation population and, hence, may not be directly comparable. The 6MWT is a self-paced test; the participants in the current study were not limited by symptoms, whereas it is possible that patients with chronic heart failure experience performance-limiting symptoms such as breathlessness and that, the onset of these symptoms occurs at a predictable work rate.

The results of this study did not support the relative reliability of the 6MWT over three walks in this population, when the *ICC* was based on absolute agreement. However, when the *ICC* was calculated for consistency, the results indicated good relative reliability with repeated testing with an $ICC_{consistency}$ of .90 across the three walks. This finding is consistent with earlier reports in the literature (Hamilton & Haennel, 2000), suggesting that patients remain consistent in their relative ranking order. However, the use of reliability coefficients based on relative ranking or consistency alone can be misleading (Costa-Santos, Bernardes, Ayres-de-Campos, Costa, & Costa, 2011; Keating & Matyas, 1998). When absolute scores are important, such as in clinical situations when the obtained 6MWT distances may guide the implementation of an individualised exercise program, the 6MWT may not have sufficient relative reliability.

In this study, the two-way random effects $ICC_{agreement}$ scores yielded much smaller values than the two-way random effects $ICC_{consistency}$ scores. The smaller values were a reflection of the systematic change in the distances walked between Walk 1, Walk 2 and Walk 3. Previous research in the area of orthopaedics have recommended reporting of the model, type and measures used in the calculation of the *ICC* (Lee et al., 2012). The results of this research also highlighted the importance of adequately describing the model of *ICC* used.

Measurement error for the individual score and the change scores for the group and individual remained large even after repeated testing. The $SEM_{agreement}$ demonstrated considerable variability even after repeated testing. For example, the $SEM_{agreement}$ of 32 m derived from 6MWT distances in Walk 2 and Walk 3 indicated if a true walk distance score was 500 m, the observed score can be calculated with 95% confidence using the formula $500 \pm 1.96 \times SEM$ or for a true score of 500 m the observed score, with 95% confidence would fall between 437 and 563 m. This is important if using cut-off points in clinical decision making. For example, a poor performance in the 6MWT in cardiac populations has been suggested as less than 300 m and a high score greater than 450 m (Lucas et al., 1999; Zugck et al., 2000). In this study, when a practice test was not completed, the $SEM_{agreement}$ was 41 m, for a true score of 300 m the observed score, with 95% confidence would fall between 220 and 380 m. For a high score of 450 m, the observed score, with 95% confidence would fall between 370 and 530 m. An observed score of 370 or 380 m would be difficult to interpret. The large $SEM_{agreement}$ scores reflect the large degree of systematic change in repeated 6MWTs makes it difficult to interpret observed 6MWT scores.

The reported 95% confidence intervals for the individual demonstrated that if the first test was considered a practice test, an individual would need to increase his or her walk distance by more than 95 m between the second and third walk to be confident of real change with 95% confidence. This is higher than reports by Nogueira et al. (2006) who reported a distance greater than 47 m to be 95% confident of true change between the second and third walk. The wide confidence intervals may be a reflection of the heterogeneous population in cardiac rehabilitation, compared with other research investigating specific cardiac populations such as early post myocardial infarction (Nogueira et al., 2006; Roberts, Li, & Sykes, 2006) or chronic heart failure (Patrick, 2005, 2008).

This study presents new information on the measurement error of the 6MWT for patients during the pre-cardiac rehabilitation assessment in the form of the 95% confidence intervals for the group. If the first test were regarded as a practice test, a group would need to change by more than 44 m to be confident of real change with 95% confidence and this score increases to more than 62 m between the first and the second walk. Some researchers have measured 6MWT results before and after cardiac rehabilitation (Roberts et al., 2006; Tallaj et al., 2001; Wright, Khan, Gossage, & Saltissi, 2001) and reported mean change scores from 57 to 86 m in patients with left ventricular ejection fraction of equal to or greater than 40% and those with left ventricular ejection fraction less than 40% respectively (Tallaj et al., 2001). (Wright et al. (2001) and Roberts et al. (2006) reported improvements of 62 m and 67 m, respectively. Given these values, it is possible that the measurement error exceeds the increase in distance observed after cardiac rehabilitation programs when only one test is performed.

The results indicated that in this cardiac rehabilitation population, three 6MWTs repeated in relatively short time frames prior to the commencement of cardiac rehabilitation were insufficient to yield reliable scores in terms of relative reliability and absolute reliability. It is possible that adding further testing sessions, that is having a fourth or even a fifth testing session at baseline, may address the systematic error and yield reliable results. However, in a clinical setting it is unlikely that clinical staff and patients would have the ability to complete more than three tests without causing delays to care and to the start of cardiac rehabilitation. The time required to complete three walk tests with adequate rest breaks in between nears 90 minutes. Each additional test would require approximately 30 minutes additional assessment time. It is unknown what effect patient fatigue would play with these additional tests. Patrick (2005) discussed the time constraints of repeated testing in a clinical setting. The 6MWT has been shown to be well tolerated in a variety of cardiac populations during repeated same-day testing (Cahalin et al., 1996; Kervio et al., 2004a; Kristjánsdóttir et al., 2004; Nogueira et al., 2006; Patrick, 2005), with the test being safely completed three or four times in one day (Kristjánsdóttir et al., 2004; Nogueira et al., 2006; Patrick, 2005). Patrick (2005) also suggested that the time required to complete three or four tests may not be practical in a clinical setting.

2.5.1 Implications.

When used as a measure of exercise tolerance in a mixed cardiac rehabilitation population prior to commencing cardiac rehabilitation, the 6MWT was not reliable even after three tests. The relatively high $ICC_{consistency}$ demonstrated in our study is potentially misleading, as it does not take account of systematic change. The systematic measurement error was large and would likely obscure any true change in

individuals and possibly groups of patients. These results suggest that the 6MWT is not a reasonable test to use in this population because it is not practical for a clinician or patient to complete more than three 6MWTs prior to commencing a cardiac rehabilitation program (Patrick, 2005). These results suggest that, in a mixed cardiac population, if it is not feasible to complete more than three tests, the 6MWT is not a reliable test. Further research is necessary to investigate other field exercise tests that are practical, reliable and valid tests for patients receiving cardiac rehabilitation.

Chapter 3: Retest Reliability of the 10 m Incremental Shuttle Walk Test

3.1 Chapter Aims

This study addressed secondary research aims 4 and 5. The first aim was to determine the evidence for both the relative retest reliability of the 10 m ISWT when up to three tests were performed in a single session in a mixed cardiac rehabilitation population. The second aim was to determine the evidence for measurement error of the 10 m ISWT when up to three tests are performed in a mixed cardiac rehabilitation population. The results have been published (Appendix 1) (Hanson, Taylor, & McBurney, 2016) and are presented in an expanded format in this chapter.

3.2 Introduction

The 10 m ISWT is a field exercise test that can be used in cardiac rehabilitation, and supervised by clinicians such as physiotherapists. Results of the 10 m ISWT may provide clinicians with information on physical fitness and functional capacity at an individual or group level, and if repeated at the end of cardiac rehabilitation may provide a measure of change for the individual or group. A description of the 10 m ISWT was provided in Chapter 1, Section 1.4.2.2.

The 10 m ISWT may be a suitable alternative to the self-paced symptom-limited 6MWT to measure the physical fitness and functional capacity of patients attending cardiac rehabilitation. The results reported in Chapter 2 indicated that the 6MWT may not be a reliable field test in the cardiac rehabilitation population. It was demonstrated

in Chapter 2 that when the 6MWT was repeated up to three times, the systematic change remained large, and the confidence intervals for both the group and individual were wide. One possible reason proposed for these findings was an ongoing learning effect. The participants were not symptom-limited during the self-paced walk test, for example limiting shortness of breath, and after recovery, demonstrated increased walk distances on subsequent 6MWTs.

Previous research supports the relative reliability of the 10 m ISWT in the cardiac rehabilitation population, with reliability coefficients between the first and second walk ranging from .94 (Fowler et al., 2005; Jolly, Taylor, Lip, & Singh, 2008) to .98 (Green et al., 2001). Relative reliability was further improved when the 10 m ISWT was repeated, with reliability coefficients of .99 between the second and third walk (Fowler et al., 2005; Lewis, Newal, Townend, Hill, & Bonser, 2001). The method of calculation of the reliability coefficient varied between researchers. Jolly et al. (2008) reported that the *ICC* in the analysis used between and within subject means squares produced from a one-way ANOVA according to the methods described by Rankin and Stokes (1998), equivalent of an *ICC Model 1*. This model provides the most conservative estimate of all the *ICC* models (Rankin & Stokes, 1998; Shrout & Fleiss, 1979). Fowler et al. (2005) did not specify the model of *ICC* used in the research and Lewis et al. (2001) did not provide details of the type of correlation analysis used in their research. The previous research provides support for the evidence for the relative test retest reliability of the 10 m ISWT in groups of patients eligible for cardiac rehabilitation. However, as demonstrated in Chapter 2 and by previous researchers, the use of reliability coefficients when measured for consistency can be misleading in test retest reliability studies because the correlation coefficient provides a measure of

the relationship between scores and not the agreement between the scores (Bland & Altman, 1986; Costa-Santos et al., 2011; Keating & Matyas, 1998).

There is limited information about the measurement error of the 10 m ISWT in cardiac rehabilitation populations, including the number of tests required to achieve reliable scores that minimise measurement error. Three studies were located that reported measurement error of the 10 m ISWT in cardiac populations using group or individual confidence intervals (Fowler et al., 2005; Jolly et al., 2008; Pulz et al., 2008). For the cardiac rehabilitation group, after one test, an improvement in test score required to be interpreted as greater than measurement error ranged from 36 m in a mixed cardiac rehabilitation group in a single session (Jolly et al., 2008), to 56 m for a group of patients following coronary artery bypass surgery who completed the tests within one week (Fowler et al., 2005). For an individual attending cardiac rehabilitation, after one test, an improvement in test score required to be interpreted as more than measurement error ranged from 53 m in those with heart failure in a single session (Pulz et al., 2008) to 122 m (Fowler et al., 2005). Additionally, Fowler et al. (2005) reported that measurement error improved if a practice walk was included. They demonstrated if a second walk was completed within one week, an increase in walk distance of more than 5 m for groups and 21 m for individuals exceeded measurement error with 95% confidence, in patients following coronary artery bypass graft surgery (Fowler et al., 2005). The addition of a practice test may reduce the amount of change needed to exceed measurement error and improve the retest reliability of the 10 m ISWT in cardiac rehabilitation populations.

There remains uncertainty about whether there is systematic error when the 10 m ISWT is repeated in a cardiac rehabilitation population. Three studies reported

significant improvement in test scores for the 10 m ISWT between the first and second walk (Fowler et al., 2005; Jolly et al., 2008; Morales et al., 1999), whereas four studies reported no significant change in test scores (Arnott, 1997; Green et al., 2001; Payne & Skehan, 1996; Pulz et al., 2008). Two studies reported using a practice walk but did not provide any further information in their data analysis (Green et al., 2001; Lewis et al., 2001). It is possible that a practice walk may be required when using the 10 m ISWT in a mixed outpatient cardiac rehabilitation population but little information is available about the absolute reliability of the test when assessed in a single session.

The aim of this study was to determine the evidence for the relative reliability and the measurement error of the 10 m ISWT in a mixed cardiac rehabilitation population.

3.3 Method

Research used a quantitative design of repeated testing on participants referred to a cardiac rehabilitation program.

3.3.1 Ethics approval and consent.

Ethics approval was granted by La Trobe University, Faculty of Health Sciences, Faculty Human Ethics Committee (reference number FHEC07/99) and La Trobe University, University Human Ethics Committee (reference number 10-082), Bendigo Health Human Research Ethics Committee (reference number 1/2007) and St John of God Health Care Human Research Ethics Committee (reference number 438) (see Appendix 2). All participants enrolled in the investigation read and signed a

Participant Information and Consent Form that was approved by the relevant ethics committees (see Appendix 3).

3.3.2 Participants.

3.3.2.1 Eligibility criteria for participants.

The eligibility criteria were the same as those described in Chapter 3, Section 2.3.2.1.

3.3.2.2 Recruitment procedures.

A consecutive series of patients referred to cardiac rehabilitation at one of two centres who met the eligibility criteria of the study were invited to participate by the cardiac rehabilitation nurse coordinator at the site. Referrals for patients were received from internal hospital wards, external hospitals, cardiologists, general practitioner or self-referral. The cardiac rehabilitation nurse coordinator at each site provided the initial screen to determine eligibility. Recruitment for the pilot study began in October 2007 and finished in November 2007 and for the main study began in March 2011 and finished in October, 2011.

3.3.2.3 Sample size.

Sample size estimates were completed according to the method described by Walter et al. (1998) and based on *a-priori* set levels of optimal and minimal acceptable limits of reliability for clinical measurement. For two tests, a minimum of 12 people would be needed if a minimum *ICC* level of 0.6 (P_0) was accepted and the hypothesis that findings from this study would be consistent with the current literature at an *ICC* of .9

(P_1), at a level of significance (α) of .05 and power of .8 (β=.2). A minimum of 46 people would be needed for two tests if a minimum *ICC* level of 0.8 (P_0) was accepted and the hypothesis that findings from this study would be consistent with the current literature at an *ICC* of .9 (P_1).

3.3.3 Procedure.

A repeated measures design was used in the physiotherapy department at two participating hospitals. All participants completed two 10 m ISWTs in a single session before commencement of a cardiac rehabilitation program, the 10 m ISWTs were labelled ISWT 1 and ISWT 2. The first 10 participants enrolled, participated in a pilot study and completed a third 10 m ISWT in the same session as the first two tests and labelled ISWT 3. All walk tests were completed under the same conditions by the same investigator.

3.3.3.1 Baseline demographic information.

Baseline demographic information was collected during the initial assessment and included age, date of birth, gender, height and weight, cardiac diagnosis and intervention, cardiac signs and symptoms, cardiac medications, relevant medical history, and other relevant medications, mobility and use of gait aid. This information, along with the date and time of assessment were recorded on a data-recording sheet (see Appendix 4). From the data-recording sheet, BMI was calculated as per the description in Chapter 2, Section 2.3.3.1.

3.3.3.2 Outcome measure: 10 m ISWT.

Participants were allowed a 20-minute rest break prior to commencing the first 10 m ISWT and were given a 30-minute rest break between subsequent tests to ensure they had returned to pre-exercise levels for heart rate, blood pressure, oxygen saturation and rating of perceived exertion (Steele, 1996). Participants were not able to commence the 10 m ISWT if resting systolic blood pressure was greater than or equal to 200 mmHg or resting diastolic blood pressure was greater than or equal to 110 mmHg (American College of Sports Medicine, 2010a).

The 10 m ISWT protocol was administered according to the description of Singh et al. (1992) and described in Chapter 1, Section 1.4.2.2. The 10 m ISWTs were completed on an indoor, flat, 10 m track. A shuttle referred to one 10 m lap. A recording of standardised instructions for the 10 m ISWT was played immediately prior to beginning the walk test. The test commenced after a four-second count down followed by a triple beep.

No encouragement was provided during the test. To assist with familiarisation, the researcher walked alongside the participant for the first minute. After this minute the researcher stood midway and to the side of the 10 m track to observe the participant. The researcher reminded the participant to increase walking speed slightly when the triple beep sounded. If the participant reached the cone before the signal beep, the researcher reminded the participant to wait until the beep before commencing the next shuttle.

The test was stopped when the participant could no longer maintain the required pace or was more than 0.5 m from the cone after the signal beep after one opportunity to catch-up. Additional criteria for early termination of the test included patient distress, dizziness, angina, or onset of severe musculoskeletal pain, failure of the heart rate to increase with exercise, fall in oxygen saturation below 90% (American College of Sports Medicine, 2010a) or attainment of 85% of the maximum heart rate (Singh et al., 1992) using the heart rate reserve method.

The number of shuttles completed were recorded and at the completion of each test converted to the distance walked. In addition, at the completion of the test participants rated their Borg rate of perceived exertion score using the 6-20 scale (Borg, 1990; Borg & Noble, 1973). For safety, heart rate and oxygen saturation were monitored throughout each test with additional monitoring after the test until all values had returned to a level within 20% of the initial recordings. Blood pressure was recorded in a seated position before and after the test and until both diastolic and systolic values had returned to 20% of the initial recording. This information was recorded on the data-recording sheet (see Appendix 4).

3.3.4 Statistical methods.

Participant characteristics were recorded (age, gender, weight, height, cardiac intervention and other relevant history) to allow description of the study sample. Where relevant, means were expressed followed by standard deviation. Continuous data were checked for normality. In the case of missing values, data were not imputed. A one-way ANOVA was applied to determine the presence of significant

systematic change between walks. *Post hoc* comparisons (least significant difference) were performed to test for statistically significant differences between the walk tests.

Retest reliability was expressed in two ways: first as relative reliability, as a ratio in the form of the *ICC* (Shrout & Fleiss, 1979) with corresponding confidence intervals, and second as measurement error using the *SEM*, and using the 95% confidence intervals for the group and the individual. The relative reliability of the sample was interpreted using both the $ICC_{consistency}$ and $ICC_{agreement}$ two-way random effects model with 95% confidence intervals across the walks, and for all possible pairs of walk tests. The *ICC* was interpreted with the following guidelines: good reliability was a score greater than 0.75, moderate reliability between 0.50 and 0.75 and poor reliability less than 0.5 (Portney & Watkins, 2000, p. 565). A detailed description of the *ICC* and a detailed rationale for this model was provided in Chapter 2, Section 2.3.4.1.

For the measurement error of an individual score the $SEM_{consistency}$ and $SEM_{agreement}$ were derived according to the method described in Chapter 2, Section 2.3.4.4.1. The measurement error around the change scores were calculated for the group and the individual using 95% confidence intervals (Taylor, Dodd, et al., 2004) according to the method described in Chapter 2, Section 2.3.4.4.2.

3.4 Results

Sixty-two participants who met the eligibility criteria of the study consented to participate in the study (Figure 3.1). There were no dropouts. All participants who met the eligibility criteria of the study completed two 10 m ISWTs in a single session

prior to the commencement of the cardiac rehabilitation program. The first 10 participants completed a third walk test in the same session prior to the commencement of the cardiac rehabilitation program. All participants completed the required number of 10 m ISWTs to volitional exhaustion without complications or safety concerns. There were no deviations from the study.

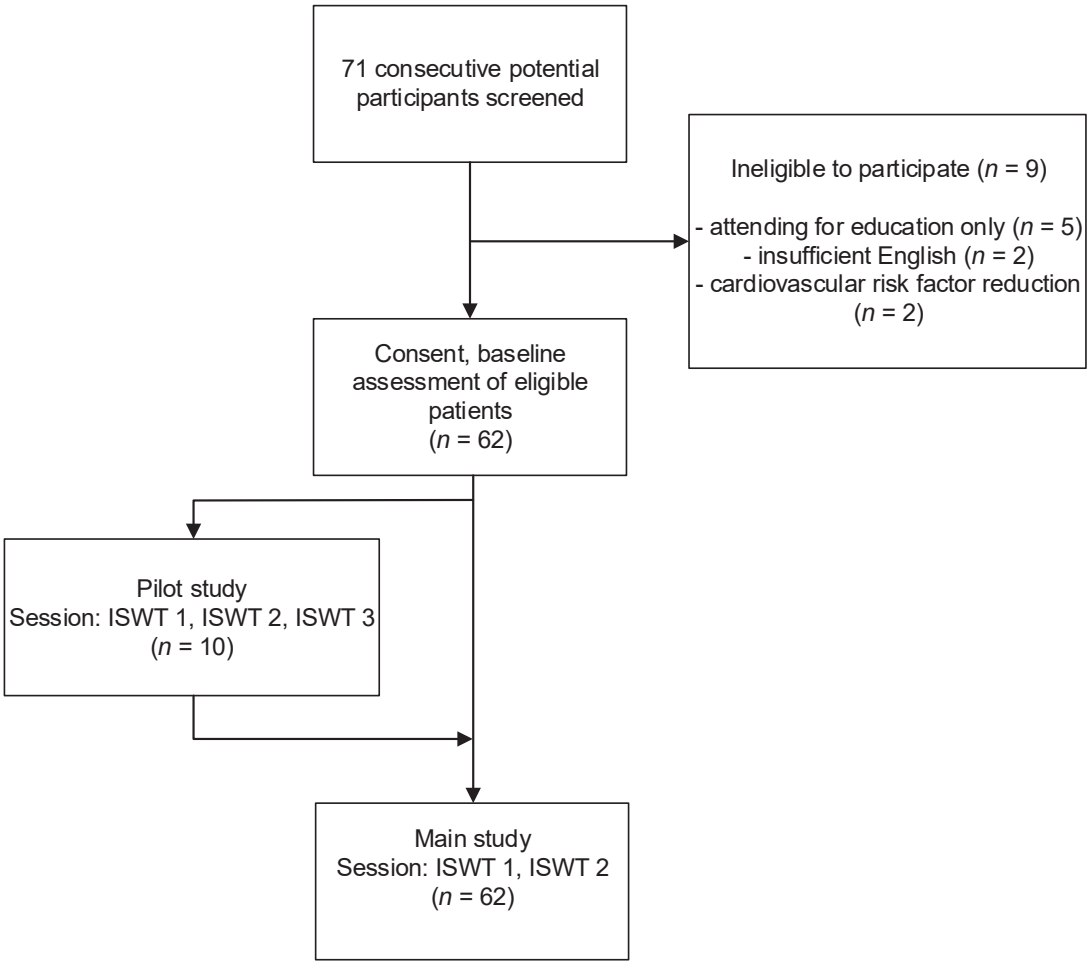


Figure 3.1. Participant recruitment and participation for the retest reliability of the 10 m ISWT.

3.4.1 Characteristics of the sample.

3.4.1.1 Baseline demographic characteristics of participants.

Table 3.1 shows the baseline demographic characteristics for all participants (n = 62), as well as the participants in the pilot study (n = 10). The mean age of participants was 68 years (*SD* 10) with the youngest participant 46 years and the oldest participant 91 years. The mean time elapsed since an acute event was 11 days (*SD* 24). All participants mobilised independently, three participants required the use of a single point stick.

Table 3.1

Characteristics of the Sample

Characteristics	Pilot study n = 10	Main study n = 62
Age mean (*SD*)	67 (10)	68 (10)
Gender male:female *n*	5:5	45:17
Intervention *n* (%)		
Revascularisation Surgery	8 (80)	39 (63)
Medical management	2 (20)	16 (26)
Other intervention	0	7 (11)
Height mean (*SD*) (cm)	161 (11)	170 (9)
Weight mean (*SD*) (kg)	83 (17)	84 (15)
BMI mean (*SD*) (kg/m^2)	32 (5)	29 (5)
BMI category *n* (%)		
Underweight	0	0
Healthy	0	10 (16)
Overweight	4 (40)	29 (47)
Obese	6 (60)	22 (35)
Morbidly obese	0	1 (2)

Note. BMI = Body mass index.

3.4.1.2 Cardiovascular observations and safety.

All participants were stable and not in breach of the safety criteria for commencing the walk tests. No participant stopped the test early due to cardiac symptoms or pain or distress. Participants made a full recovery between walk tests. All cardiovascular observations returned to baseline scores after each walk (Table 3.2).

Table 3.2

Recovery and Symptoms Reported at End of the 10 m ISWT

End of test symptoms	ISWT 1 (n = 62)	ISWT 2 (n = 62)	ISWT 3 (n = 10)
2 min Recovery *n* (%)	57 (92)	48 (77)	4 (40)
Symptoms reported at end of test *n* (%)			
No Symptoms	21 (34)	17 (27)	1 (10)
Lower limb pain	10 (16)	11 (18)	3 (3)
Shortness of breath	34 (55)	36 (58)	7 (70)
Other	1 (2)	3 (5)	0

Note. [a] 2 min recovery was defined as heart rate within 20% of baseline measure at two minutes.

3.4.2 Pilot study.

In the pilot study, the mean distance walked for ISWT 1 was 378 m (*SD* 224), for ISWT 2 was 393 m (*SD* 239) and for ISWT 3 was 398 m (*SD* 229). There was a trend for an increase in the mean distance walked from ISWT 1 to ISWT 2 to ISWT 3, but, this trend was not statistically significant ($F_{(2,18)} = 3.197$, $p = 0.065$).

For the pilot group (n = 10), good relative retest reliability was obtained over the three walks completed prior to cardiac rehabilitation, the $ICC_{consistency}$ was .994, 95% CI

[.981, .998] and $ICC_{agreement}$ was .992, 95% CI [.976, .998]. The results for relative reliability for all combinations of the walk tests for the pilot study are presented in Table 3.3.

Table 3.3

Relative Test Retest Reliability of the 10 m ISWT for the Pilot Study

	ICC for consistency			*ICC* for absolute agreement		
		95% CI			95% CI	
10 m ISWT	*ICC*	LL	UL	*ICC*	LL	UL
ISWT 1, 2 and 3	.994	.981,	.998	.992	.976,	.998
ISWT 1 and 2	.992	.967,	.998	.990	.961,	.998
ISWT 1 and 3	.992	.967,	.998	.989	.940,	.997
ISWT 2 and 3	.997	.989,	.999	.997	.990,	.999

Note. *ICC* = intraclass correlation coefficient; CI = confidence interval; LL = lower limit; UL = upper limit.

The $SEM_{consistency}$ ranged from 12 to 21 m and the $SEM_{agreement}$ ranged from 12 to 24 m. The results of the *SEM* for all combinations of walk tests are shown in Table 3.4.

Table 3.4

Summary of Measurement Error for the Individual Scores of the 10 m ISWT for the Pilot Study

	SEM for consistency		*SEM* for absolute agreement	
10 m ISWT	$SEM_{consistency}$ (m)	% of grand mean	$SEM_{agreement}$ (m)	% of grand mean
ISWT 1 and 2	21	5	23	6
ISWT 1 and 3	21	5	24	6
ISWT 2 and 3	12	3	12	3

Note. *SEM* = standard error of measurement.

For the pilot group (n = 10) a change of at least 37 m would be required as an indication of real change, over and above measurement error with 95% confidence, between the first and the second walk test. If the first 10 m ISWT is regarded as a practice test, a change of at least 18 m would be required as an indication of real change, over and above measurement error with 95% confidence, between the second and third ISWT completed before cardiac rehabilitation.

The group and individual confidence intervals for the pilot study are shown in Table 3.5. For the group, a change of at least 37 m from the first to the second walk test and 42 m between the first and the third walk test would be required as an indication of real group change, over and above measurement error with 95% confidence. A mean group change of 18 m would be required as an indication of real group change, over and above measurement error with 95% confidence, in a cardiac rehabilitation group between the second and third walk test, with the first test considered a practice test.

Table 3.5

Summary of Measurement Error for the Change Scores for the Pilot Study

		Group 95% CI (m)		Limits of Agreement (Individual 95% CI) (m)	
10 m ISWT	M_{diff} (m) (SD_{diff})	LL	UL	LL	UL
ISWT 1 and 2	15 (30)	-6	36	-53	83
ISWT 1 and 3	20 (29)	-1	41	-46	86
ISWT 2 and 3	5 (17)	-7	17	-33	43

Note. M_{diff} = mean difference; SD_{dif} = standard deviation of the difference; CI = confidence interval; LL = lower limit; UL = upper limit.

For the individual, a change of at least 84 m from the first to the second walk test and 87 m between the first and the third walk test would be required as an indication of

real change, over and above measurement error with 95% confidence. A change of 44 m would be required as an indication of real change, over and above measurement error with 95% confidence, in an individual prior to attending cardiac rehabilitation between the second and third 10 m ISWT, with the first test considered a practice test.

Bland Altman plots for all combinations of pairs of walks are presented in Figure 3.2. Figure 3.2A and 3.2B shows a participant with a difference in distance walked outside the upper limit of the limits of agreement. This participant was a 62-year-old overweight woman who was 36 days post coronary artery graft surgery. The participant completed 480 m in ISWT 1, 560 m in ISWT 2 and 570 m in ISWT 3.

The results of the pilot study demonstrated no significant difference between the three 10 m ISWTs, good relative reliability for all combinations of 10 m ISWTs, and acceptable measurement error. It was decided that the main study would investigate the test retest reliability of the 10 m ISWT when two tests are completed in a single session in a larger sample.

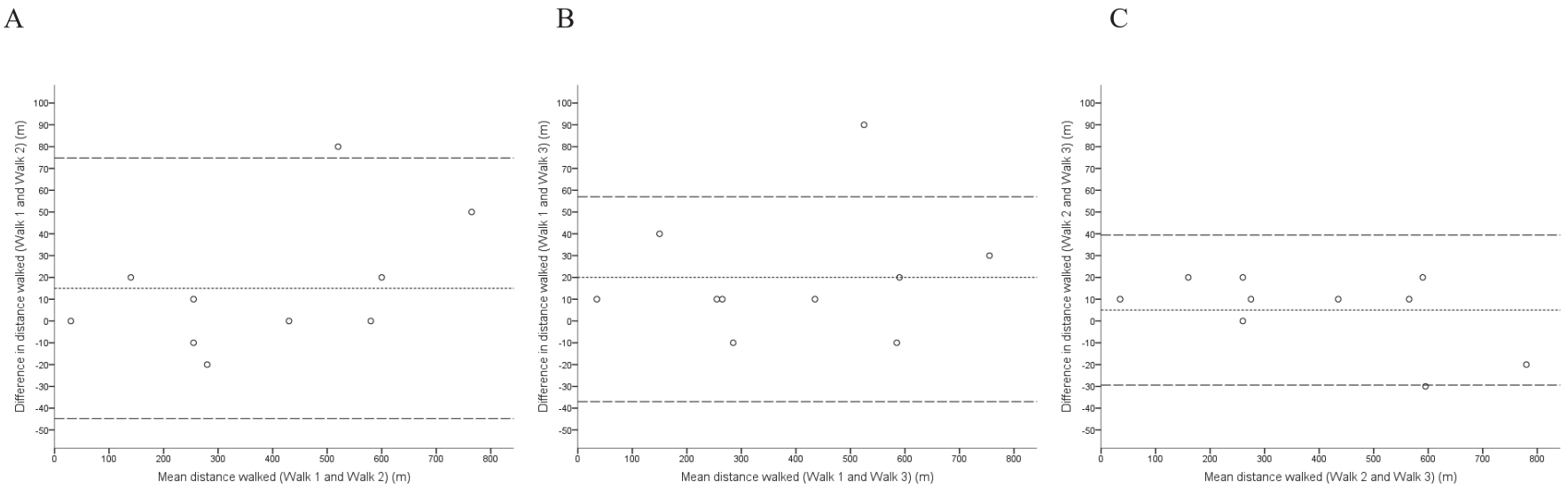


Figure 3.2. Bland Altman plots of the 10 m ISWT distance (m) for all combinations of tests during the pilot study, (A) ISWT 1 and ISWT 2, (B) ISWT 1 and ISWT 3 and (C) ISWT 2 and ISWT 3.

Key. *x*-axis is the mean score; *y*-axis the absolute difference; unbroken line = mean difference; broken line is the limits of agreement calculated by $1.96 \times SD_{diff}$.

3.4.3 10 m ISWT distance.

For the sample (n = 62), the mean distance walked improved by 17 m (*SD* 18) over the two walks. Participants (n = 62) walked a mean distance of 378 m (*SD* 173) for ISWT 1 and 395 m (*SD* 176) for ISWT 2. This difference was statistically significant ($t_{(61)}$ = 7.613 p = <.001). The distance walked in ISWT 1 ranged from 30 to 760 m and in ISWT 2 from 30 to 790 m.

Forty-seven of the 62 participants walked further in ISWT 2 compared with ISWT 1, with a mean increase of 23 m. Eleven participants walked the same distance in ISWT 1 and ISWT 2. Four participants walked a shorter distance in ISWT 2 compared with ISWT 1, with a mean decrease of 15 m.

3.4.4 Relative test retest reliability.

For the sample (n = 62) good relative retest reliability was shown for ISWT 1 and ISWT 2 for both consistency and absolute agreement. The two-way mixed effects $ICC_{consistency}$ was .995, 95% CI [.992, .997]. The two-way mixed effects $ICC_{agreement}$ was .990, 95% CI [.928, .997].

3.4.5 Measurement error.

3.4.5.1 Measurement error individual score.

The $SEM_{consistency}$ was 12 m or 3% of the grand mean. When the test retest variability was included in the measurement error the $SEM_{agreement}$ was 17 m or 4% of the grand mean.

3.4.5.2 Measurement error change scores.

For the group, a change of at least 23 m would be required as an indication of real change, over and above measurement error with 95% confidence, between the first and second walk test (Table 3.6). For the individual in this population, a change of at least 54 m from the first to the second walk test would be required as an indication of real change, over and above measurement error with 95% confidence (Table 3.6).

Table 3.6

Summary of Measurement Error for the Change Scores

		Group 95% CI (m)		Limits of Agreement (Individual 95% CI) (m)	
10 m ISWT	M_{diff} (m) (SD_{diff})	LL	UL	LL	UL
ISWT 1 and 2	17 (18)	12	22	-19	53

Note. M_{diff} = mean difference; SD_{dif} = standard deviation of the difference; CI = confidence interval; LL = lower limit; UL = upper limit.

A Bland Altman plot for the sample is presented in Figure 3.3. There was one participant with a difference in distance walked outside the upper limit of the limits of agreement. The characteristics of this participant was described in 3.4.2.

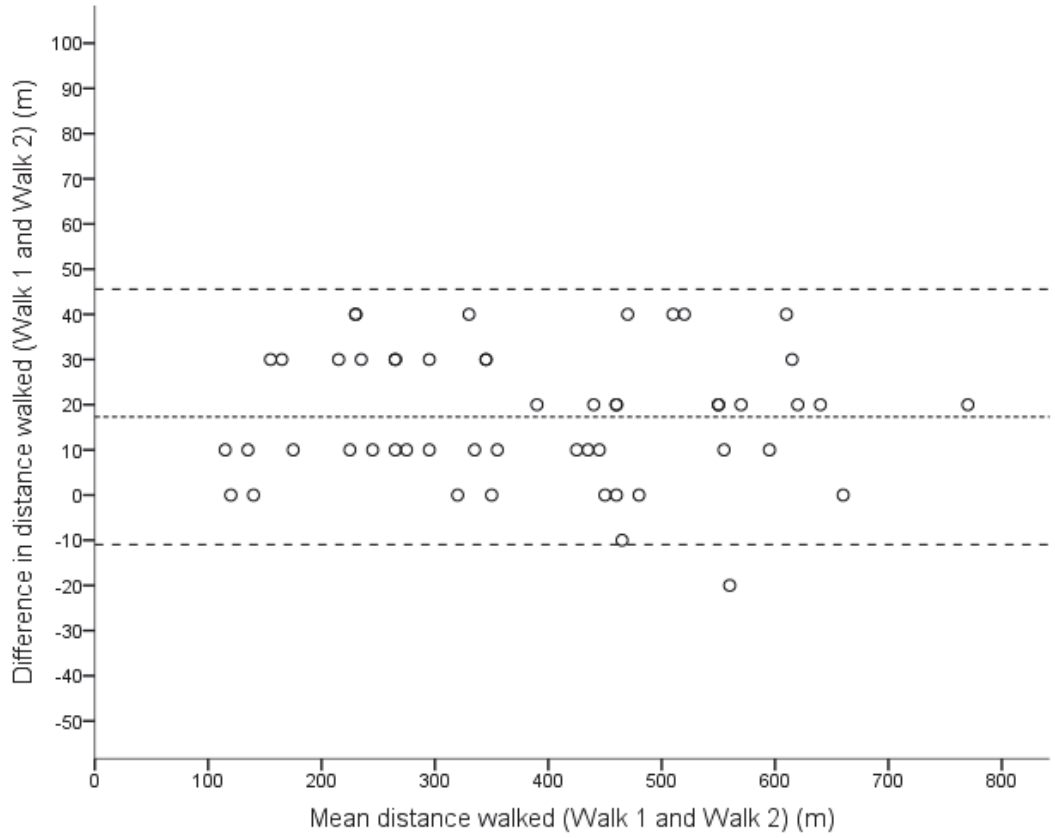


Figure 3.3. Bland Altman plot of the distance walked (m) in the two 10 m ISWTs completed in the main study.

Key. *x*-axis is the mean score; *y*-axis the absolute difference; unbroken line = mean difference; broken line is the limits of agreement calculated by $1.96 \times SD_{diff}$.

3.5 Discussion

The purpose of this study was to assess the evidence for retest reliability of the 10 m ISWT as a measure of physical fitness and functional capacity in a mixed cardiac rehabilitation population.

The pilot study results suggested there was little difference in relative reliability and measurement error between the first and second test and the first and third test. This indicated that little additional clinical information or improvements in reliability were gained by completing the third test. The $SEM_{agreement}$ for the first and second walk was

23 m and for the first and third walk was 24 m, both approximately 6% of the grand mean. The group 95% confidence intervals indicated a group would need to improve by 37 m between the first and second walk and by 42 m between the first and third walk for true change above measurement error. For the individual, an improvement of 84 m and 87 m between the first and second and first and third walks respectively to be 95% confident of true change above measurement error. This supports the notion that the third test is unnecessary.

The main study investigated the absolute and relative reliability of two 10 m ISWTs completed in a single session in a larger sample. The main study supported the reliability of the 10 m ISWT in cardiac rehabilitation population. However, further to the pilot study, the main study provided evidence that the addition of the second test may not be clinically relevant and one test may be sufficient for adequate test retest reliability in a cardiac rehabilitation population. The relative reliability expressed as the *ICC* was at least .99 between the first and second walk in the main study. The $SEM_{agreement}$ was 17 m or 4% of the grand mean and the 95% confidence intervals suggest the group would need to improve by 23 m and the individual by 54 m to be confident of true change over and above measurement error.

The results of the main study indicated there was a statistically significant difference between the distance walked between the first and the second walk. The mean difference was 17 m. This statistical significance is likely to be a result of the larger sample size, but the magnitude of the difference is unlikely to be clinically significant. There was very little difference between both the $ICC_{agreement}$ and $ICC_{consistency}$ scores and the $SEM_{agreement}$ and $SEM_{consistency}$ scores, reflecting the very small amount of systematic change seen between the first and second walk in the main study. While

there was a statistical difference in the distance walked between the first and second walk of the main study, the results are unlikely to be clinically important and do not affect the reliability of the test.

Two studies have reported group improvements of more than 58 m following a six-week intervention (Fowler et al., 2005; Robinson, Samani, & Singh, 2011). Both studies reported at least one practice 10 m ISWT in the initial baseline testing, but no practice test at follow-up, after the intervention. The results presented in this chapter suggest that a single 10 m ISWT is sufficiently reliable to detect this amount of typical change observed after a short cardiac rehabilitation intervention, and for that change to be interpreted as real change over and above measurement error both for individuals and in the evaluation of group programs.

The results suggest that the 10 m ISWT has a higher level of retest reliability compared with the 6MWT. The results of Chapter 2 showed measurement error of the 6MWT remained high even between the second and third test. For the 6MWT, when a practice walk was included an improvement of 45 m for the group and 96 m for the individual was required to overcome measurement error with 95% confidence. While in this chapter, when a practice walk was included, the distance walked in the 10 m ISWT would need to improve by 18 m for the group, and 44 m for the individual to be 95% confident of real change. If no practice walk was included, the group would need to improve by 23 m and the individual by 54 m to be 95% confident of real change. The research presented in this chapter demonstrates the support for the retest reliability of the 10 m ISWT when a single test is performed.

3.5.1 Implications.

The 10 m ISWT is safe and feasible when used as a field exercise test in a mixed outpatient cardiac rehabilitation population. The results of this chapter indicate good retest reliability of the 10 m ISWT when a single 10 m ISWT was performed and slightly better scores for retest reliability when two tests were completed with the first regarded as a practice test. It appears that the benefits to retest reliability when a second test is performed may not be clinically important. The choice of completing one or two tests may be left to the clinician or researcher to balance the retest reliability with the time taken to complete a second test. The pilot study results suggested that there was little difference in relative reliability and measurement error between the first and second test and the first and third test, indicating that no additional clinical information or improvements in reliability are gained by completing more than one practice test. The externally paced and incremental 10 m ISWT appears to have better retest reliability than the self-paced 6MWT in this population. The following empirical studies presented in this thesis will focus on the validity and interpretation of the 10 m ISWT rather than the 6MWT.

Chapter 4: Criterion Validity of the 10 m ISWT in Cardiac Rehabilitation

4.1 Chapter Aims

This study addressed secondary research aim 6 of this thesis, specifically to determine the evidence for criterion validity of the 10 m ISWT in a mixed cardiac rehabilitation population. To do this, a study of the association between the outcomes of the 10 m ISWT and a symptom-limited exercise test was completed. Participants completed a multistage treadmill exercise stress test using the Bruce Protocol in a cardiology laboratory and within a one-week period also completed two 10 m ISWTs in a single session, with the order randomised. The main outcome measure was distance walked and test time, the secondary outcome measures were maximum heart rate, and rate pressure product and oxygen saturation.

4.2 Introduction

Chapter 3 reported research that demonstrated the 10 m ISWT was highly reliable in a mixed cardiac rehabilitation population when either a single test was performed or two tests, with the first regarded as a practice test. The 10 m ISWT has been described in detail in Chapter 1, Section 1.4.2.2. The test outcomes provide objective measures that can be used to make inferences about physical fitness and functional capacity (Solway, Brooks, Lacasse, & Thomas, 2001). Unlike the symptom-limited exercise test, the 10 m ISWT is not generally used as a diagnostic tool or prognostic tool (Rasekaba et al., 2009).

The gold standard for measuring physical fitness and functional capacity of patients with cardiac disease are the symptom-limited exercise test with ECG monitoring or the cardiopulmonary exercise test (American College of Sports Medicine, 2010b, p. 111; Balady et al., 2007; Palange et al., 2007; Piepoli et al., 2010; Thompson, Arena, Riebe, & Pescatello, 2013). Despite these recommendations, it is not always possible to complete a laboratory test at entry and exit to a cardiac rehabilitation program (Simms et al., 2007). Field exercise tests, including the 6MWT and 10 m ISWT may be a suitable alternative for the entry assessment (Taylor, Bell, & Lough, 2010). These alternatives for exercise testing have been included in national guidelines for cardiac rehabilitation programs in Australia, New Zealand and throughout the UK (Goble & Worcester, 1999; New Zealand Guidelines Group & Heart Foundation, 2002; Price et al., 2016). The predictive function of the 10 m ISWT is unknown.

The investigation of criterion validity, or the degree to which outcome measures of the 10 m ISWT are associated with outcome measures of a gold standard (Mokkink et al., 2010c) in cardiac rehabilitation populations, has been reported for patients with chronic heart failure (Green et al., 2001; Lewis et al., 2001; Morales et al., 1999; Pulz et al., 2008), coronary artery bypass surgery (Fowler et al., 2005), and elderly stable coronary artery disease patients (Mandic, Walker, et al., 2013). All followed the protocol described by Singh et al. (1992) for the 10 m ISWT. All compared the 10 m ISWT distance walked, number of shuttles, or peak 10 m ISWT speed with the peak oxygen consumption during a symptom-limited exercise test; with moderate to high correlations ($.72 \leq r \leq .87$) and 52% to 76% of the variation in the criterion measure predicted from the 10 m ISWT distance or number of shuttles walked (Fowler et al.,

2005; Green et al., 2013; Lewis et al., 2001; Mandic, Walker, et al., 2013; Pulz et al., 2008).

There is limited published research comparing other functional outcome measures from a symptom-limited exercise test with the 10 m ISWT. Measuring oxygen consumption is not always possible in cardiology laboratories. The associated costs of staff, equipment outlay and equipment maintenance are often prohibitive (Chatterjee, Sengupta, Nag, Kumar, & Rudra, 2013). Functional outcomes of a stress test include duration, maximal heart rate and blood pressure, oxygen saturation, symptoms and limiting factors, and ECG changes. These functional measures are commonly reported in patient referrals to cardiac rehabilitation and where maximal exertion was attained, these functional outcome measures have been considered as useful as oxygen consumption (Bruce, Blackmon, Jones, & Strait, 1963; Bruce & Hornsten, 1969).

The association between the 10 m ISWT and functional outcome measures of symptom-limited exercise tests in patients with chronic heart failure has been reported to be moderate, with between 29% and 46% of the variation in the maximum heart rate of the symptom-limited exercise test able to be predicted from variation in peak heart rate in the 10 m ISWT (Green et al., 2001; Lewis et al., 2001; Mandic, Walker, et al., 2013; Morales et al., 1999; Pulz et al., 2008). There is no information on the association, specifically the concurrent validity or the predictive validity, of the functional outcomes of the 10 m ISWT in a mixed cardiac rehabilitation population.

The aim of this research was to determine the evidence for concurrent criterion validity of the 10 m ISWT as an objective measure of physical fitness and functional capacity in cardiac rehabilitation. This study aimed to determine first, if those who

perform well on a symptom-limited exercise test also perform well on the 10 m ISWT, and if functional variables such as peak exercise heart rate and exercise time achieved on the symptom-limited exercise test are associated with equivalent outcomes on the 10 m ISWT; and second, if the 10 m ISWT distance could predict the symptom-limited exercise test time.

4.3 Method

4.3.1 Ethics approval and consent.

Ethics approval was granted by La Trobe University, University Human Ethics Committee (reference number 10-082), and St John of God Health Care Human Research Ethics Committee (reference number 438) (see Appendix 2). All participants enrolled in the investigation read and signed a Participant Information and Consent Form that was approved by the relevant ethics committees (see Appendix 3).

4.3.2 Participants.

4.3.2.1 Eligibility criteria for participants.

All adults with stable and treated heart disease, irrespective of severity or duration of the condition, who were referred to complete a treadmill symptom-limited exercise test as part of a stress echocardiogram investigation in the cardiology department, were eligible to participate in this study. This included those adults with acute coronary syndrome, stable angina pectoris, following revascularisation procedures

including coronary artery bypass graft surgery or percutaneous transluminal coronary angioplasty or coronary artery stent; coronary artery disease as diagnosed by angiography, heart valve surgery, chronic heart failure; heart transplantation; following cardiac resynchronisation therapy; or implantable cardioverter defibrillator.

Participants were excluded if they had any condition where exercise would be contraindicated, if they were unable to walk for any neurological or musculoskeletal reason, or completed a symptom-limited exercise using a protocol other than the Bruce protocol. Participants were also excluded if they were diagnosed with cardiovascular risk factors in the absence of diagnosed cardiovascular disease, with congenital heart disease, were children or pregnant, were required to cease their usual cardiac medications for the symptom-limited exercise test. Participants who had previously completed a 10 m ISWT or had any condition where a symptom-limited exercise test would be an absolute or relative contraindication (Fletcher et al., 2001) were also excluded. In addition, participants were excluded if limitations in English language production or comprehension skills precluded them from understanding the consent form.

4.3.2.2 Recruitment procedures.

A consecutive series of patients referred to the cardiology centre of a regional hospital for a symptom-limited exercise test who met the eligibility criteria of the study were invited to participate by the manager of the centre. Referrals for patients were received from three local cardiologists. The manager of the department provided the initial screen to determine eligibility. Recruitment began in June 2014 and finished in July 2014.

4.3.2.3 Allocation of participants to groups.

After providing consent, 15 participants were randomly allocated by an independent staff member to one of two groups, labelled Group A and Group B, representing two different testing protocols. Group A completed the symptom-limited exercise test first and the two 10 m ISWTs within one week. Group B first completed the two 10 m ISWTs in a single session and then the symptom-limited exercise test within one week. The sequence of allocation to either Group A or Group B was determined according to a random numbers generator using a concealed method. The sequence was generated using the Website, randomization.com, using permuted blocks of four (Dallal, 2008).

4.3.2.4 Sample size.

Sample size estimates were completed according to the method described by Howell (Howell, 2012, p. 258). The following formula was used:

$$n = (\frac{\delta}{\rho_1})^2 + 1$$

Where *n* is the sample size; δ is a constant and for a power of .80 equals 2.8; and ρ_1, an estimate of the correlation in the population, was based on the minimum correlation of .75 which was defined as the lower limit of a strong correlation (Portney & Watkins, 2000, p. 565) and greater than a suggested minimal acceptable correlation of .70 (Terwee et al., 2007). Based on this formula, the recommended sample size estimate was 15 for concurrent criterion validity.

Recommendations for the sample size for prediction equations vary from five to 10 cases per independent variable (Norman & Streiner, 2008, p. 157) to 50 cases per independent variable (Tabachnick & Fidell, 2007, p. 123). Based on this, the sample size estimate of 15 will be used to explore the evidence for the predictive criterion validity of the 10 m ISWT in patients with cardiac disease.

4.3.3 Procedure.

4.3.3.1 Baseline demographic information.

Baseline demographic information was collected during the initial assessment and included age, date of birth, gender, height, weight, BMI, cardiac diagnosis and intervention, cardiac signs and symptoms, cardiac medications, relevant medical history and other relevant medications, usual exercise tolerance, mobility and use of gait aid. This information along with the date and time of assessment were recorded on a data-recording sheet (see Appendix 4). From the data-recording sheet, BMI was calculated as per the description in Chapter 2, Section 2.3.3.1.

4.3.3.2 Outcome measure: symptom-limited exercise test.

Participants completed their symptom-limited exercise test, as part of a stress echocardiogram, under the supervision of a medical doctor and followed the usual practice guidelines for the clinic. The symptom-limited exercise test was administered according to the Bruce protocol (Bruce et al., 1973) also described in Chapter 1, Section 1.4.1. The test was completed in a cardiology laboratory with temperature and humidity control. The motorised treadmill was capable of speeds up to 20 km/hr and

an incline greater than 22%. For safety, the treadmill had a front rail, two side rails and an emergency stop button.

Prior to commencing the test, each participant rested for a minimum of 30 minutes. A resting supine ECG was performed. Limb electrodes were then moved to the torso, and a modified limb lead supine ECG was performed and compared with the standard ECG. Blood pressure was also taken prior to commencing the exercise test. Standardised instructions were given to the patient including instruction on the safe and correct use of the treadmill. The test continued until the end, or terminated if there were any significant ECG changes, a reduction in blood pressure with increasing workload, unreasonable hypertension, onset of angina or increasing angina, patient reporting symptoms of distress or dizziness, excessive shortness of breath or claudication, or changes in general appearance.

For monitoring and safety, blood pressure was recorded once during every stage of the Bruce protocol. Continuous monitoring of anterior, lateral, and inferior myocardial zones occurred through a three lead ECG and at the end of each minute of exercise a 12 lead ECG. Participants were instructed to inform the supervising medical doctor of any symptoms that developed through the test. At completion of the test, blood pressure, a 12 lead ECG recording and peak rating of perceived exertion were recorded, followed by an echocardiogram.

The primary outcome measure used for comparison with the 10 m ISWT was the duration of the symptom-limited exercise test. Secondary outcome measures used were peak exercise heart rate, oxygen saturation measured via pulse oximetry and rate pressure product calculated by the formula:

$$Rate\ pressure\ product = peak\ heart\ rate\ \times peak\ systolic\ blood\ pressure$$

Rate pressure product provides information on the work of the myocardium (May & Nagle, 1984) and myocardial oxygen consumption (Gobel, Nordstrom, Nelson, Jorgensen, & Wang, 1978).

4.3.3.3 Outcome measure: 10 m ISWT.

The procedure for the 10 m ISWT was completed as per the description in Chapter 3, Section 3.3.3.2. The results of Chapter 3 indicated that the retest reliability of 10 m ISWT was supported with an $ICC_{consistency}$ of .99 in a cardiac rehabilitation population.

The information from the testing of the 10 m ISWT was recorded on the data-recording sheet (see Appendix 4). The primary outcome measure used was the distance walked calculated from the number of completed shuttles, and the secondary outcome measures were peak heart rate, oxygen saturation and the rate pressure product.

4.3.4 Statistical methods.

Data were analysed using the statistical package for the social sciences (Version 22, IBM SPSS Statistics for Windows, Version 22.0. Armonk, NY: IBM Corp). Where relevant, means were expressed followed by standard deviation. In the case of missing values, data were not imputed. The significance level was set at a two tailed $p < 0.05$. After checking for any series effects due to the order of testing using a mixed plot two-way ANOVA, the groups were combined for testing of criterion validity.

4.3.4.1 Concurrent criterion validity of the 10 m ISWT.

The strength of the relationship between the symptom-limited exercise test and the two trials of the 10 m ISWT for all primary and secondary outcome measures was assessed with the Pearson's product-moment correlation coefficient (*r*) with 95% confidence intervals. The 95% confidence intervals for the Pearson's product-moment correlation coefficient were calculated using the process described by Glass and Hopkins (1996, p. 357) using Fisher *r* to *Z* transformation. The correlation coefficient was interpreted with the following guidelines: strong, a correlation greater than 0.75, moderate between 0.50 and 0.75 and weak less than 0.5 (Portney & Watkins, 2000, p. 565). The significance level was set at a two tailed $p < 0.05$. Terwee et al. (2007) in their recommendation of evaluating criterion validity, interpret a reasonable correlation score as at least .70.

Bland Altman plots graphically displayed the concurrent criterion-related validity of the 10 m ISWT. The limits of agreement were defined as the mean difference ± 2 standard deviations.

The variation in the symptom-limited exercise test explained by the 10 m ISWT was summarised by the coefficient of determination (r^2) (Portney & Watkins, 2009, p. 551). The proportion of the variation in the symptom-limited exercise test not explained by the 10 m ISWT was calculated using $1 - r^2$. This was calculated for all primary and secondary outcome measures.

4.3.4.2 Predictive criterion validity of the 10 m ISWT.

Predictive criterion validity of the 10 m ISWT was explored using standard error of estimate and linear regression with 95% confidence intervals for the slope (*b*) for the 10 m ISWT distance and the symptom-limited exercise test duration. Confidence intervals were calculated around the regression equation for the group and individual. The relationship was also demonstrated visually, using scatterplots.

The standard error of estimate was calculated for all primary and secondary outcome measures (Altman & Gardner, 2000, p. 91; Portney & Watkins, 2009, p. 552). A higher standard error of estimate represents greater variability around the regression line and therefore greater error in the prediction of a regression equation. The standard error of estimate ($s_{Y \cdot X}$) was calculated using the equation:

$$s_{Y \cdot X} = S_Y \sqrt{(1 - r^2)\frac{(N-1)}{(N-2)}}$$

Where S_Y is the standard deviation of the symptom-limited exercise test score, and r^2 the coefficient of determination (Howell, 2012, p. 269).

Linear regression calculated values for the y intercept (*a*) and slope (*b*). The 95% confidence intervals for the slope were calculated according to Altman and Gardner (Altman & Gardner, 2000, p. 76). Confidence intervals for the intercept (*a*) were not calculated as previous researchers have reported the clinical interpretation to be meaningless (Altman & Gardner, 2000, p. 77).

The group and individual confidence intervals for the symptom-limited exercise test for a specific 10 m ISWT score were calculated (Altman & Gardner, 2000, p. 76; Howell, 2012, p. 275). Individual confidence limits are otherwise known as the prediction interval. The distances selected for the 10 m ISWT related to the distance walked at the completion of each level of the 10 m ISWT test. The formula and calculation are presented in Appendix 5.

4.4 Results

4.4.1 Characteristics of the sample.

Twelve men and three women participated in this study with a mean age of 65 years (*SD* 8). The demographic characteristics of the participants are summarised in Table 4.1. There were no statistically significant differences between Group A and Group B in terms of age, height, weight, BMI, and waist or hip circumference.

4.4.2 Exercise test.

All participants completed the symptom-limited exercise test and two 10 m ISWT on their usual medications within one week. The mean duration of the symptom-limited exercise test was 7.6 minutes (2.5), the duration ranged from four to 14 minutes. For the 10 m ISWTs, the distance walked in the first test was 522 m (216) and the second test 535 m (219). The distance ranged from 270 to 1020 m. The primary and secondary outcomes of the exercise tests are shown in Table 4.2. The mean difference between the first and second 10 m ISWT was 13 m (SD_{diff} 13) and this difference was statistically significant ($t(14) = 3.839$, $p = .002$).

Table 4.1

Characteristics of the Sample

Characteristics	Group A (n = 7)	Group B (n = 8)	Sample (n = 15)
Age	63 (7)	67 (8)	65 (8)
Gender *n* Male:Female	5:2	7:1	12:3
Intervention *n* (%)			
Revascularisation Procedure	6 (86)	7 (87.5)	13 (87)
Medical management	1 (14)	1 (12.5)	2 (13)
Height (cm)	174 (12)	174 (11)	174 (11)
Weight (kg)	92 (24)	89 (16)	91 (19)
BMI (kg/m^2)	30 (6)	30 (5)	30 (5)
BMI category *n* (%)			
Underweight	0	0	0
Healthy	1 (14)	2 (25)	3 (20)
Overweight	2 (29)	1 (12.5)	3 (20)
Obese	3 (43)	5 (62.5)	8 (53)
Morbidly obese	1 (14)	0	1 (7)

The one patient who completed all levels of the 10 m ISWT was a 57-year-old male, who had a percutaneous intervention. He had been attending cardiac rehabilitation for education only and within a month of his PCI, he had returned to his usual exercise activities, which included cycling more than 400 km in a week. He also achieved 14 minutes on the Bruce Protocol.

4.4.3 Between-group differences.

There was no statistically significant difference between the duration of the symptom-limited exercise test and ISWT 1 ($t(14) = .733$, $p = .476$) or the symptom-limited exercise test and ISWT 2 ($t(14) = 1.114$, $p = .284$) between groups A and B.

Table 4.2

10 m ISWT and Symptom-Limited Exercise Test Results for the Sample

Exercise test outcome	Total (n = 15)	Group A (n = 7)	Group B (n = 8)	t(13)	p
ISWT 1 distance (m)	522 (216)	566 (229)	484 (211)	.721	.48
ISWT 1 duration (min)	7.8 (2.0)	8.3 (2.0)	7.4 (2.0)	.798	.44
ISWT 1 peak HR (beats/min)	111 (11)	115 (10)	108 (12)	1.283	.22
ISWT 1 RPP (mmHg/beats/min)	16833 (2397)	16980 (2094)	16704 (2774)	.215	.83
ISWT 1 SpO_2 (%)	97 (1)	97 (1)	97 (1)	.142	.89
ISWT 2 distance (m)	535 (219)	577 (229)	499 (219)	.678	.51
ISWT 2 duration (min)	8.0 (1.9)	8.4 (2.0)	7.6 (1.9)	.718	.49
ISWT 2 peak HR (bpm)	114 (10)	117 (10)	112 (10)	.961	.35
ISWT 2 RPP (mmHg/beats/min)	17501 (2163)	17677 (1944)	17349 (2462)	.283	.78
ISWT 2 SpO_2 (%)	97 (1)	97 (1)	98 (1)	.755	.46
SLET duration (min)	7.6 (2.5)	8.4 (2.9)	6.9 (1.9)	1.23	.24
SLET peak HR (beats/min)	135 (10)	135 (10)	136 (10)	.042	.97
SLET RPP (mmHg/beats/min)	21427 (2747)	21029 (2608)	21776 (2994)	.511	.62
SLET SpO_2 (%)	95 (2)	95 (1)	95 (2)	.432	.67

Note. HR = heart rate; RPP = rate pressure product; SpO_2 = oxygen saturation measured via pulse oximetry; SLET = symptom-limited exercise test.

There were no interaction effects between the exercise test results and the order of the tests completed ($F(2,26) = 1.283$ $p = .294$). As there were no significant between-group differences or interaction effects between the groups (Group A and Group B) it was decided that only data for the combined sample ($n = 15$) will be presented.

4.4.4 Criterion validity results.

4.4.4.1 The strength of the relationship between 10 m ISWT and symptom-limited exercise test.

The association between the outcome measures for the symptom-limited exercise test and the 10 m ISWT are shown in Table 4.3. All correlations ranged from moderate to strong, with the correlation for the primary outcome measures ranging from .85 to .87 between the ISWT distance and symptom-limited exercise test duration. The correlation values were similar for the symptom-limited exercise test and first 10 m ISWT, and symptom-limited exercise test and the second 10 m ISWT.

The concurrent criterion-related validity of the 10 m ISWT against the symptom-limited exercise test is shown graphically in Figure 4.1 using Bland Altman plots. For both the first and second 10 m ISWT agreement between the symptom-limited exercise test was satisfactory, with all scores falling within the limits of agreement.

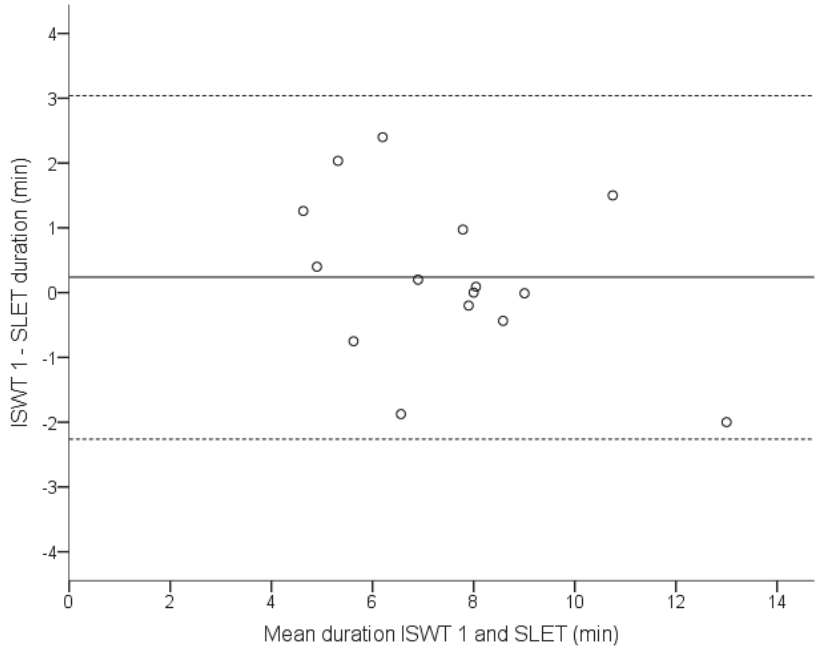
Table 4.3

Strength of the Relationship between the 10 m ISWT and the Symptom-Limited Exercise Test

Relation between outcome measures	*r*	95% CI	
		LL	UL
Primary Outcome Measures			
ISWT 1 distance and SLET duration	.87	.63	.95
ISWT 1 duration and SLET duration	.86	.62	.95
ISWT 2 distance and SLET duration	.85	.59	.95
ISWT 2 duration and SLET duration	.84	.57	.94
Secondary Outcome Measures			
ISWT 1 distance and SLET METs	.80	.49	.93
ISWT 1 duration and SLET METs	.80	.49	.93
ISWT 1 HR and SLET HR	.70	.29	.89
ISWT 1 SpO_2 and SLET SpO_2	.82	.52	.94
ISWT 1 RPP and SLET RPP	.71	.30	.89
ISWT 2 distance and SLET METs	.78	.45	.92
ISWT 2 duration and SLET METs	.77	.43	.92
ISWT 2 HR and SLET HR	.79	.47	.93
ISWT 2 SpO_2 and SLET SpO_2	.68	.26	.89
ISWT 2 RPP and SLET RPP	.71	.32	.90

Note. *r* = Pearson's product-moment correlation; CI = confidence interval; LL = lower limit; UL = upper limit; SLET = symptom-limited exercise test; METs = metabolic equivalent; HR = heart rate; SpO_2 = oxygen saturation measured by pulse oximetry; RPP = rate pressure product calculated by the formula: $peak\ heart\ rate \times peak\ systolic\ blood\ pressure$.

A



B

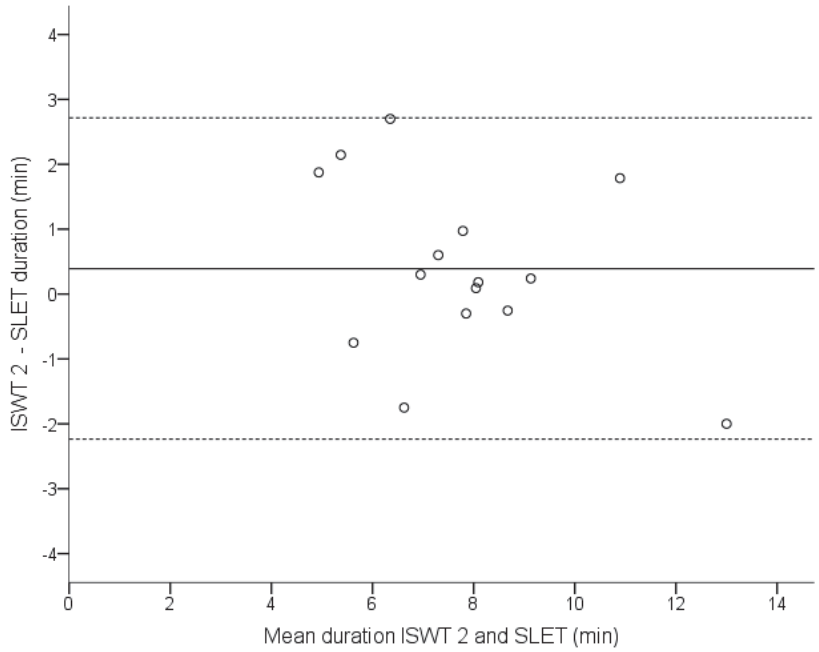


Figure 4.1. Bland Altman plots of the duration of the 10 m ISWT and the duration of the symptom-limited exercise test.

Figure 4.1A presents results for ISWT 1 and 4.1B results for ISWT 2. Key: unbroken line = mean difference; broken lines = limits of agreement calculated by $1.96 \times SD_{diff}$; SLET = symptom limited exercise test; *x*-axis = the mean score; *y*-axis = the absolute difference.

4.4.4.2 The variation of the symptom-limited exercise test score predicted by the 10 m ISWT score.

The results for coefficient of determination for all outcome measures are summarised in Table 4.4. The coefficient of determination ranged from .72 to .75 for the distance walked in the first or second 10 m ISWT and the duration of the symptom-limited exercise test. Meaning that up to 75% of the variation in the symptom-limited exercise test duration could be explained by the 10 m ISWT distance and that between 25 and 28% of the variance in the symptom-limited exercise test duration was not explained by the relationship between the 10 m ISWT distance.

Table 4.4

Coefficient of Variation for Primary and Secondary Outcome Measures for the 10 m ISWT and the Symptom-Limited Exercise Test

Relation between outcome measures	r^2	1-r^2
Primary Outcome Measures		
SLET duration and ISWT 1 distance	.75	.25
SLET duration and ISWT 1 duration	.74	.26
SLET duration and ISWT 2 distance	.72	.28
SLET duration and ISWT 2 duration	.70	.30
Secondary Outcome Measures		
SLET HR and ISWT 1 HR	.49	.51
SLET HR and ISWT 2 HR	.63	.37
SLET SpO_2 and ISWT 1 SpO_2	.67	.33
SLET SpO_2 and ISWT 2 SpO_2	.47	.53
SLET RPP and ISWT 1 RPP	.50	.50
SLET RPP and ISWT 2 RPP	.51	.49

Note. r^2 = coefficient of determination; SLET = symptom-limited exercise test; HR = heart rate; SpO_2 = oxygen saturation measured by pulse oximetry; RPP = rate pressure product calculated by the formula: $peak\ heart\ rate \times peak\ systolic\ blood\ pressure$.

4.4.4.3 Modelling the accuracy of prediction of the symptom-limited exercise test duration by the 10 m ISWT distance.

The standard error of estimate (s_{yx}) was calculated for each of the primary and secondary outcomes (Table 4.5). Also presented in Table 4.5 are the results for the linear regression equations, specifically the intercept (a) and slope (b) with 95% confidence intervals for the primary and secondary outcome measures. For the slope (b), none of the 95% confidence intervals crossed zero.

Using the distance walked in the first 10 m ISWT, the linear equation to predict the duration of the symptom-limited exercise test was:

$$SLET\ duration = 2.43 + (0.01 \times ISWT\ 1\ distance)$$

Where r = .87, 95% CI [.63, .95], standard estimate of error of 1.3.

Using the distance walked in the second 10 m ISWT, the linear equation to predict the duration of the symptom-limited exercise test was:

$$SLET\ duration = 2.48 + (0.01 \times ISWT\ 2\ distance)$$

Where r = .85, 95% CI [.59, .95], standard estimate of error of 1.4

The linear relationship between the symptom-limited exercise test duration and first and second 10 m ISWT distance are displayed using scatter plots in Figure 4.2.

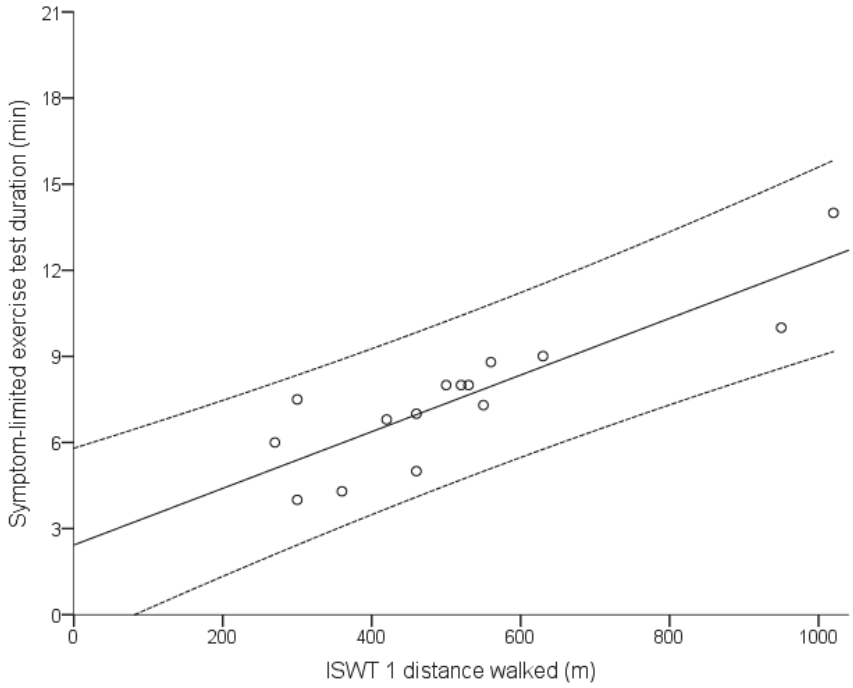
Table 4.5

Relation between Primary and Secondary Outcome Measures for the 10 m ISWT and the Symptom-Limited Exercise Test

Relation between outcome measures	M_{diff} (SD_{diff})	s_{yx}	Intercept (a)	Slope (b)	95% CI	
					LL	UL
Primary Outcome Measures						
SLET duration and ISWT 1 distance	-	1.3	2.4	0.01	.007	0.01
SLET duration and ISWT 1 duration	.24 (1.3)	1.3	-0.8	1.1	0.70	1.5
SLET duration and ISWT 2 distance	-	1.4	2.5	0.01	.006	.01
SLET duration and ISWT 2 duration	.39 (1.4)	1.5	-0.9	1.1	.6	1.5
Secondary Outcome Measures						
SLET HR and ISWT 1 HR	-24 (8)	7.1	68.0	.6	.2	1.0
SLET HR and ISWT 2 HR	-21 (6)	6.0	46.3	.8	.4	1.1
SLET SpO_2 and ISWT 1 SpO_2	2 (1)	1.1	-30.7	1.3	.7	1.8
SLET SpO_2 and ISWT 2 SpO_2	2 (1)	1.4	-42.2	1.4	.5	2.3
SLET RPP and ISWT 1 RPP	-4594 (1995)	2015.4	7783.5	.8	.3	1.3
SLET RPP and ISWT 2 RPP	-3925 (1938)	1998.6	5579.2	.9	.4	1.4

Note. M_{diff} = mean difference; SD_{dif} = standard deviation of the difference; s_{yx} = standard error of estimate; SLET = symptom-limited exercise test; HR = heart rate; SpO_2 = oxygen saturation measured by pulse oximetry; RPP = rate pressure product.

A



B

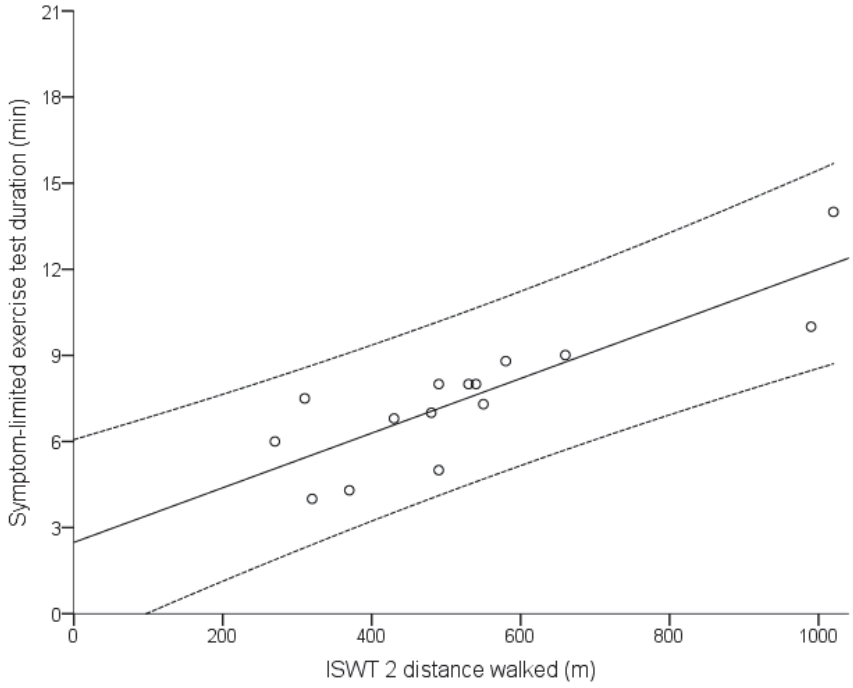


Figure 4.2. The linear relationship between the 10 m ISWT distance (m) and the symptom-limited exercise test duration is plotted with individual 95% confidence limits.

In Figure 4.2A a linear relationship of $r^2 = .75$ was observed. In Figure 4.2B a linear relationship of $r^2 = .72$ was observed.

4.4.4.3.1 Confidence limits for the symptom-limited exercise test

The predicted symptom-limited exercise test duration with group and individual confidence limits were calculated for the distance walked in the first and second 10 m ISWT are shown in Table 4.6 and 4.7, respectively. For example, for an individual who completed all 12 levels of the first 10 m ISWT (i.e., walked 1020m) could be predicted, with 95% confidence, to complete between 9.2 and 16.0 minutes of the symptom-limited exercise test when a Bruce protocol is followed. The predicted symptom-limited exercise test duration was similar for ISWT 1 and ISWT 2.

Table 4.6

Prediction of Symptom-Limited Exercise Test Duration from the First 10 m ISWT

10 m ISWT distance (m)	Estimated SLET duration (mins)	*95%* CI_{group}		*95%* $CI_{Individual}$	
		LL	UL	LL	UL
0	2.4	0.5	4.3	0	5.8
30	2.7	.9	4.5	0	6.0
70	3.1	1.4	4.8	0	6.4
120	3.6	2.0	5.2	0.4	6.8
180	4.2	2.8	5.6	1.1	7.3
250	4.9	3.7	6.1	1.9	7.9
330	5.7	4.7	6.7	2.7	8.7
420	6.6	5.8	7.4	3.7	9.5
520	7.6	6.9	8.3	4.7	10.5
630	8.7	7.9	9.5	5.8	11.6
750	9.9	8.9	11.0	6.9	12.9
880	11.2	9.8	12.6	8.1	14.3
1020	12.6	10.7	14.5	9.2	16.0

Note. SLET = symptom-limited exercise test; CI = confidence interval; LL = lower limit; UL = upper limit. If lower limit (LL) confidence interval was calculated as < 0 it was denoted as 0.

Table 4.7

Prediction of Symptom-Limited Exercise Test Duration from the Second 10 m ISWT Distance

10 m ISWT distance (m)	Estimated SLET duration (mins)	95% CI_{group}		95% $CI_{Individual}$	
		LL	UL	LL	IL
0	2.5	0.4	4.6	0	6.2
30	2.8	0.8	4.8	0	6.4
70	3.2	1.3	5.1	0	6.7
120	3.7	2.0	5.4	0.2	7.2
180	4.3	2.8	5.9	0.9	7.7
250	5.0	3.7	6.3	1.7	8.3
330	5.8	4.7	6.9	2.6	9.0
420	6.7	5.8	7.6	3.6	9.8
520	7.7	6.9	8.5	4.6	10.8
630	8.8	8.0	9.6	5.7	11.9
750	10.0	8.9	11.1	6.8	13.2
880	11.3	9.8	12.8	8.0	14.6
1020	12.7	10.8	14.6	9.1	16.3

Note. SLET = symptom-limited exercise test; CI = confidence interval; LL = lower limit; UL = upper limit. If lower limit (LL) confidence interval was calculated as < 0 it was denoted as 0.

4.5 Discussion

The purpose of this study was to determine the evidence for criterion validity of the 10 m ISWT as a measure of physical fitness and functional capacity in a mixed cardiac rehabilitation population. Methods assessed both the concurrent criterion validity and the predictive criterion validity of the 10 m ISWT to determine if the 10 m ISWT could be used as a reasonable surrogate for the symptom-limited exercise test in terms of being an objective measure of physical fitness and functional capacity The results supported the concurrent criterion validity of the 10 m ISWT as a measure

of physical fitness and functional capacity in cardiac rehabilitation when one or two tests were performed. This study provided evidence of a strong correlation between the duration of the symptom-limited exercise test and the distance walked in the 10 m ISWT (r = .85 to .87). These results were supported by the Bland Altman plots showing agreement between the duration of the symptom-limited exercise test and duration of the 10 m ISWT. The index of reliability, in other words the estimated maximum correlation was .88 (Lord & Novick, 1968, p. 72), based on the reported correlation for the symptom-limited exercise test (r = .87) and the 10 m ISWT ($ICC_{consistency}$ = .99). The results supported the concurrent criterion validity of the 10 m ISWT, and demonstrated that the two exercise tests showed consistency in ranking patients according to physical fitness and functional capacity. The results did not improve when a second 10 m ISWT was performed, which further reinforced the findings of Chapter 3, that a single walk test may be sufficient in this population.

The correlations reported in the current study were stronger than previously published correlations between symptom-limited exercise test duration and 10 m ISWT distance that ranged from .54 to .68 in patients with chronic heart failure (Green et al., 2001; Lewis et al., 2001; Pulz et al., 2008). Additionally, the current study reported correlations ranging from .69 to .79 for peak heart rate achieved in symptom-limited exercise test and 10 m ISWT, compared with earlier reports of weak to moderate associations (Green et al., 2001; Lewis et al., 2001; Morales et al., 1999). The current study used the Bruce protocol for the symptom-limited exercise test, whereas the three earlier studies used alternative protocols. Differences in the patient population may also account for differences in associations. The inclusion criteria for this study included any ambulant participant with treated and stable heart disease. In

comparison, the previous three studies used more homogenous groups of patients with heart failure. It is possible that patients with chronic heart failure respond differently to the demands of an incremental exercise test compared with the participants in this study. The current study provides evidence that the 10 m ISWT measures the same construct as the symptom-limited exercise test and may have sufficient concurrent criterion validity to provide an objective measure of physical fitness and functional capacity for patients with treated and stable heart disease attending cardiac rehabilitation.

In the current study, the 10 m ISWT mean distances walked were 522 m (*SD* 216) and 535 m (*SD* 219), greater than the previously reported studies that ranged from 401 m (*SD* 147) (Lewis et al., 2001) to 497 m (*SD* 60) (Green et al., 2001). The better performance in the 10 m ISWT in this population may be due to the stage of recovery of the participants.

The results do not support the predictive criterion validity of the 10 m ISWT in a mixed cardiac rehabilitation group when functional outcome measures such as distance walked and duration of the symptom-limited exercise test are used. This was the first study to present information on the accuracy of the predictive criterion validity of the 10 m ISWT. While a linear regression equation was presented, the 95% confidence intervals around the slope (b) were wide and made the equation difficult to interpret. In addition, the clinically relevant 95% confidence limits for individuals, or the prediction interval, remained wide. Making it difficult, in an individual, to predict with accuracy the duration of the Bruce protocol from the distance walked in the 10 m ISWT.

4.5.1 Limitations and future directions.

The relatively wide individual confidence limits may be a result of the error contribution from both exercise tests. Repeating the 10 m ISWT did not reduce confidence intervals, and the effect of repeating the symptom-limited exercise test is unknown. Another possible contributing factor for the wide confidence limits may be the small sample size. Despite justification of the sample size, it was small compared with the recommendations included in the COSMIN checklist (Terwee et al., 2012). Future research could use the model presented in this chapter with an increase in the number of participants. An increase in the number of participants should provide a more stable regression equation and may reduce the 95% confidence intervals and prediction limits (Tabachnick & Fidell, 2007). Clinically, the ability to predict the symptom-limited exercise test duration based on the 10 m ISWT distance, or in reverse, to predict the 10 m ISWT distance based on the symptom-limited exercise test time is worth exploring as it provides an important means of translating meaning between the laboratory and the clinical environment.

A larger sample size may be able to identify cut offs or thresholds in performance in the 10 m ISWT for people with cardiac disease. It would be clinically useful for a clinician to have available cut off points, for example, to identify when a patient needs a review by his or her cardiologist.

While this study used a symptom-limited exercise test protocol, the Bruce protocol, which is considered a gold standard in exercise testing (Noonan & Dean, 2000), this study relied on functional outcome measures such as distance walked and duration of test, and not measurement of maximum oxygen consumption. Cardiopulmonary

exercise testing, with specific measurement of maximum oxygen consumption and other aerobic measures is not a practical option for patients with cardiac disease who live in non-metropolitan areas of Australia. In Bendigo, and the Loddon Mallee region, this service is not available to patients with cardiac disease and patients would need to travel at least 150 km to access these medical tests. Future research could include monitoring of maximum or peak oxygen consumption.

4.5.2 Implications.

The 10 m ISWT when compared with a symptom-limited exercise test is easy to implement, requires fewer resources, it does not require expensive equipment, and when used as an estimate for physical fitness and functional capacity, does not need a medical doctor to supervise and can be conducted within the cardiac rehabilitation environment without the need for referral. Based on these findings, the 10 m ISWT measures a similar construct (physical fitness and functional capacity) to the symptom-limited exercise test, and the two tests show consistency in ranking of participants in terms of physical fitness and functional capacity. However, in absolute terms, there is emerging evidence to suggest the 10 m ISWT cannot replace the symptom-limited exercise test. Specifically, using the regression equation generated from this small sample, the distance walked in the 10 m ISWT cannot be used to accurately predict individual functional performance on the symptom-limited exercise test. While the 10 m ISWT should not be aimed at replacing the symptom-limited exercise test, it may provide a suitable cost- and time-effective clinical alternative for patients attending cardiac rehabilitation.

Chapter 5: Construct Validity and Responsiveness of the 10 m ISWT in a Cardiac Rehabilitation Program

5.1 Chapter Aims

This study addressed the secondary research aims 7 and 8 of this thesis: first ,to determine the evidence for the construct validity of the 10 m ISWT in a cardiac rehabilitation population; and second, to determine the evidence for the responsiveness of the 10 m ISWT in a cardiac rehabilitation population. The evidence for construct validity and responsiveness were assessed using *a-priori* hypothesis tests to explore the association with common assessments used in cardiac rehabilitation. The framework and definitions provided by the COSMIN taxonomy (Mokkink et al., 2010c) were followed.

5.2 Introduction

The results of Chapter 3 and 4 supported the retest reliability and the concurrent criterion validity of the 10 m ISWT as a measure of physical fitness and functional capacity in cardiac rehabilitation. Little is known about the associations between the distance walked in the 10 m ISWT and other functional outcome measures used in cardiac rehabilitation, such as alternative measures of physical fitness and functional capacity, health-related quality of life, self-efficacy, depression and measures of body composition. Furthermore, little is known about the responsiveness of the 10 m ISWT over a mixed outpatient cardiac rehabilitation program that includes a weekly-supervised exercise program.

5.2.1 Evidence for the construct validity of the 10 m ISWT in a mixed cardiac rehabilitation program.

Knowledge of the construct validity of the 10 m ISWT in cardiac rehabilitation is limited. Construct validity can be defined as the degree to which the 10 m ISWT measures the construct it is intended to measure, specifically physical fitness and functional capacity (Clark & Watson, 1995; Cronbach & Meehl, 1955; de Vet et al., 2011, p. 169; Mokkink et al., 2010b; Streiner & Norman, 2008, p. 251). The recommended method of testing construct validity is with *a-priori* hypotheses (Cronbach & Meehl, 1955) that are specific and include the expected direction and magnitude of the correlation or difference (de Vet et al., 2011, p. 169; Mokkink et al., 2010b, 2010c). The extent that predicted associations or differences between the 10 m ISWT results and other outcome measures are observed provides information on the construct validity (Mokkink et al., 2010c). It is recommended that to support the evidence for the construct validity of a test at least 75% of the hypotheses must be confirmed (Terwee et al., 2007). However, there are no standards for the number of hypotheses that need to be tested (Scholtes, Terwee, & Poolman, 2011), or the type of hypotheses such as those designed to test convergent, discriminant or known groups' validity. No study was found that investigated the construct validity of the 10 m ISWT in cardiac rehabilitation using a range of specific *a-priori* hypotheses.

Convergent construct validity of the 10 m ISWT reflects the extent to which the test correlates with other direct measures of physical fitness and functional capacity. Testing *a-priori* hypotheses relating to criterion validity can also be considered one type of convergent construct validity (Amireault & Godin, 2014). The available literature relating to the association of the 10 m ISWT with the gold standard has been

reported in Chapter 4. Chapter 4 also presented new evidence of a moderate to strong correlation ranging from .85 to .87 between the 10 m ISWT distance and the symptom-limited exercise test duration, demonstrating support for the concurrent criterion validity as well as the convergent construct validity of the 10 m ISWT.

Higher levels of physical activity including self-reported physical activity habits, may be associated with better performance on the 10 m ISWT (Mandic, Hodge, et al., 2013). A positive association between self-reported physical activity habits and the 10 m ISWT distance in patients with coronary artery disease was demonstrated during a follow-up assessment that occurred an average of 1.6 years (*SD* 0.2) after the baseline assessment ($r = .519$, $p = .002$) (Mandic, Hodge, et al., 2013). No studies were found investigating the association between physical activity and 10 m ISWT test results at commencement of cardiac rehabilitation. Further research is needed to investigate the association between physical activity and performance in the 10 m ISWT at different stages of the rehabilitation process, such as the commencement of an outpatient cardiac rehabilitation program.

Measures of physical fitness and functional capacity might be negatively associated with higher levels of adiposity or BMI. No study was found that directly measured the strength of the association between performance in the 10 m ISWT and adiposity such as BMI, waist circumference or waist-hip ratio in cardiac rehabilitation. In patients with cardiac disease, previous research has demonstrated a negative association between BMI and peak oxygen consumption on a treadmill symptom-limited exercise test (Ades et al., 2006; Horwich et al., 2001). However, in patients with chronic heart failure, no association was found between 6MWT performance and either BMI ($-.13 \leq r \leq -.11$) (Bajraktari et al., 2011; Forman et al., 2012) or waist hip ratio ($r = .11$)

(Bajraktari et al., 2011). This provides a rationale to hypothesise that there may be a negative association between the 10 m ISWT and measures of adiposity and body composition, but further investigation is required.

It is possible that that there is a positive association between improved physical fitness or functional capacity and health-related quality of life (Stewart et al., 1994), but the existing research does not extend to the 10 m ISWT in cardiac rehabilitation. The Medical Outcomes Study 36-item short form (MOS SF-36) provides estimates of eight domains of health-related quality of life, including physical function and work-related health limitations. The MOS SF-36 scale for Physical Function includes 10 items where patients rate their capacity to perform physical activities without limitations due to health. The activities range from bathing and dressing to vigorous activity (Ware & Sherbourne, 1992). The MOS SF-36 scale for Role Physical includes four questions that measure difficulty completing work or other daily activities because of physical health (Ware & Sherbourne, 1992). Higher scores in these two scales may be indirectly related to higher levels of physical fitness and functional capacity in other adult populations (Ades, Maloney, Savage, & Carhart, 1999; Anokye, Trueman, Green, Pavey, & Taylor, 2012; Bize, Johnson, & Plotnikoff, 2007; Stewart et al., 1994; Stewart, King, & Haskell, 1993). On admission to cardiac rehabilitation, previous research has demonstrated that the 6MWT distance could explain up to 39% of the variance in the MOS SF-36 scale score for Physical Functioning (Hamilton & Haennel, 2000), and between 16 and 38% of the results of a treadmill symptom-limited exercise test (Ades et al., 1999; Ades et al., 2002; Brubaker, Witta, & Angelopoulos, 2003; Jette & Downing, 1996). No association between exercise test results and the MOS SF-36 scale score for Role Physical were

found (Quittan, Sturm, Wiesinger, Pacher, & Fialka-Moser, 1999). The association between the 10 m ISWT and the MOS SF-36 scale score for either Physical Function or Role Physical in cardiac rehabilitation is unknown. Considering results from previous studies, it is possible to hypothesise that there will be a weak to moderate association between the results of the 10 m ISWT and the MOS SF-36 scale score for Physical Function and at best, a weak association between the 10 m ISWT and the MOS SF-36 scale score for Role Physical.

The association between the 10 m ISWT and the remaining seven scales scores of the MOS SF-36 is unknown. The association between exercise test results and the remaining MOS SF-36 scale scores are unlikely to have the same strength as the Physical Function scale score of the MOS SF-36 as the scales address different constructs, such as pain and social functioning. Incremental treadmill and bicycle exercise test duration were found to explain up to 15% of the variance in the MOS SF-36 scale score of General Health in patients with cardiac disease (Beniamini, Rubensein, Zaichkowsky, & Crim, 1997; Quittan et al., 1999). Somewhat unexpectedly, an incremental bicycle exercise test was found to account for 42% of the variation in the German version of the MOS SF-36 scale score for Social Competence (Quittan et al., 1999), possibly reflecting an indirect association between reduced physical fitness and functional capacity and reduced social functioning and interaction with others. In the same study, the association between the exercise test results and the remaining MOS SF-36 scale scores were small and non-significant (Quittan et al., 1999). Further research is needed to explore the strength of the association, if any, between the remaining six MOS SF-36 scales and the 10 m ISWT. Given that the remaining scales are not measures of physical fitness or functional

capacity, it could be hypothesised that there will be no significant associations between the 10 m ISWT outcomes and the remaining six MOS SF-36 scale scores.

Self-efficacy, a cognitive construct, is an important determinant of adherence to physical activity in cardiac rehabilitation (Luszczynska & Sutton, 2006; Meland, Maeland, & Laerum, 1999; Woodgate & Brawley, 2008). In cardiac rehabilitation, the associations between self-efficacy and physical fitness and functional capacity measured by a field exercise test are not well understood. No study was found investigating the association between self-efficacy and the 10 m ISWT. Everett, Salamonson, and Davidson (2009) found a significant difference in mean 6MWT distance and exercise self-efficacy scores in participants commencing an outpatient cardiac rehabilitation program, with participants who walked more than 500 m in the 6MWT scoring significantly higher on the exercise self-efficacy questionnaire than those who walked less than 400 m in the 6MWT. Another study reported that the association between the exercise tolerance self-efficacy expectation scale score and an treadmill symptom-limited exercise test outcomes at commencement and completion of cardiac rehabilitation was .39 to .66 ($p < .05$) (Cheng & Boey, 2002). It could be hypothesised that there would be a similar association between the 10 m ISWT outcome and the self-efficacy scores, but further research is needed in this area.

The impact of depression or depressive symptoms on exercise test results is unclear. No studies were found investigating the association between depression or depressive symptoms and performance in the 10 m ISWT. Reports in the literature have compared treadmill symptom-limited exercise test results in groups of patients with and without depression or depressive symptoms and results are varied. Some have reported that there is no significant difference in treadmill symptom-limited exercise

test results for patients in cardiac rehabilitation with depression or depressive-related symptoms versus those without depressive-symptoms (Hamm et al., 2004; Milani, Lavie, & Cassidy, 1996). Whereas others have reported that those with depression or depressive-symptoms have significantly lower scores on a symptom-limited exercise test in patients with chronic heart failure (Milani & Lavie, 1998) and coronary heart disease (Ruo, Rumsfeld, Pipkin, & Whooley, 2004). Ruo et al. (2004) recruited 944 participants, of whom more than half had a diagnosis of a myocardial infarction at least six months prior to commencing the study. It is possible that the presence of depression or depressive-symptoms acts as a modifier to exercise test results in the long term. However, the effect of depression or depressive-symptoms as a modifier may not be present in the pre-program cardiac rehabilitation assessment of patients following an acute event (Shen, McCreary, & Myers, 2004). It is possible that there will not be an association between the 10 m ISWT outcome and depression in pre-program assessment, and the effects of depression as a modifier on the results of the 10 m ISWT are only seen in patients who remain depressed across a cardiac rehabilitation program intervention. It is also possible that the depression scale used and the cut-off values selected to assign depressed or not depressed affect the results of the association with the 10 m ISWT.

In summary, there is limited evidence to support the construct validity of the 10 m ISWT as a measure of physical fitness and functional capacity in a mixed cardiac rehabilitation population. The current evidence shows support that the 10 m ISWT is associated with self-reported physical activity during a long-term follow-up cardiac rehabilitation assessment and supports the hypothesis that there is a positive association between physical activity and the 10 m ISWT outcomes. The association

between the 10 m ISWT and measures of adiposity, health-related quality of life, self-efficacy and depression in cardiac rehabilitation remain unknown. Previous research in related areas supports the development of hypotheses that there will be a negative association between the 10 m ISWT and measures of adiposity; a positive association between the 10 m ISWT and physical functioning areas of health-related quality of life and self-efficacy; and no significant associations between other areas of health-related quality of life measures and depression or depressive symptoms reported at the commencement of cardiac rehabilitation. Future research should be aimed at testing these hypotheses and furthering our understanding of the 10 m ISWT as a measure of the construct of physical fitness and functional capacity.

5.2.2 Responsiveness of the 10 m ISWT in an eight-week comprehensive cardiac rehabilitation program.

The 10 m ISWT can be used to evaluate changes in physical fitness and functional capacity of an individual or group who have completed a cardiac rehabilitation program. Responsiveness measures the degree to which the 10 m ISWT change score accurately reflects the change in the construct of physical fitness and functional capacity over an outpatient cardiac rehabilitation program. It can be thought of as a type of longitudinal validity (de Vet et al., 2011, p. 202; Mokkink et al., 2010c; Polit & Yang, 2016, 9.277; Streiner & Norman, 2008; Testa & Simonson, 1996). Previously, the definition of responsiveness included the detection of meaningful or important change (Guyatt, Walter, & Norman, 1987; Mokkink et al., 2010c; Terwee et al., 2007). However, the COSMIN group achieved consensus in removing reference to the meaningful nature or importance of the change score from the scope of

responsiveness testing and for it to be included in the interpretation of a score (Mokkink et al., 2010c).

Existing research investigating support for the responsiveness of the 10 m ISWT across an outpatient cardiac rehabilitation program is limited. Responsiveness can be assessed in two ways; either by a criterion-based approach, comparing the change scores with that of the change scores of a gold standard, or a construct approach whereby a set of *a-priori* hypotheses assess the validity of the change scores (de Vet et al., 2011, pp. 205-206; Mokkink et al., 2010c; Streiner & Norman, 2008, p. 267; Terwee et al., 2007). No research was found that assessed the association of the change in 10 m ISWT with the change in a symptom-limited exercise test across cardiac rehabilitation, or using a series of *a-priori* hypothesis testing to investigate the presence and degree of association of the 10 m ISWT changes scores with those of other commonly used cardiac rehabilitation outcomes.

Previous research has provided support for the internal responsiveness of the 10 m ISWT over the duration of an outpatient cardiac rehabilitation exercise intervention (Fowler et al., 2005; Frizelle et al., 2004; Tobin & Throw, 1999). Earlier research has suggested that a minimum exercise intervention of approximately eight weeks is required to observe changes in physical fitness and functional capacity (Moholdt, Madssen, Rognmo, & Aamot, 2014; Wenger & Bell, 1986). Furthermore, the largest improvements in physical fitness were reportedly observed within the first two months of exercise training (Moholdt et al., 2014). Previous research has demonstrated a moderate effect size of .55 in the change in distance walked before and after a six-week cardiac rehabilitation exercise intervention (Fowler et al., 2005), as well as significant improvements in the distance walked in the 10 m ISWT

following a cardiac rehabilitation exercise intervention ranging from 82 to 117 m (Fowler et al., 2005; Frizelle et al., 2004; Tobin & Throw, 1999). Fowler et al. (2005) reported a significant improvement of 82 m, 95% CI [53, 110] in patients who attended a six-week cardiac rehabilitation program, three 10 m ISWTs were completed during the pre-program assessment, with the mean distance ranging from 444 (*SD* 135), to 486 (*SD* 147) to 478 m (*SD* 141). Frizelle et al. (2004) demonstrated a significant change in walk distance of 109 m from 354 m (*SD* 158) during the initial assessment, to 463 m (*SD* 27.79) in the final assessment in 22 patients with an implantable cardioverter defibrillator, following a 12-week period that included six weeks of a weekly cardiac rehabilitation exercise program. Tobin and Throw (1999) reported an improvement of 117 m in patients following coronary artery graft surgery over a 12-week cardiac rehabilitation program. These results suggest that a change, that is likely greater than measurement error, occurs across cardiac rehabilitation programs in timeframes of six weeks or more. It could be hypothesised that there will be significant differences in the 10 m ISWT outcomes before and after a cardiac rehabilitation exercise program and a moderate effect size seen in programs with a minimum duration of eight weeks.

Similar to construct validity, responsiveness testing is considered robust when *a-priori* hypotheses are designed to assess the validity of the change scores (de Vet et al., 2011, pp. 205-206; Mokkink et al., 2010c; Streiner & Norman, 2008, p. 267; Terwee et al., 2007). de Vet et al. (2011, pp. 215-216) have cautioned against the use of the effect size or *p* value in isolation when investigating the responsiveness of a test. The effect size measures the magnitude rather than the validity of the change and the *p* value obtained from a paired *t*-test is dependent on the magnitude of the change

score, the standard deviation and the sample size (de Vet et al., 2011, p. 216). These tests can be used to measure the responsiveness if they are included in *a-priori* hypothesis testing (de Vet et al., 2011, p. 218).

The effect of depression or depressive symptoms on the changes in the 10 m ISWT during a cardiac rehabilitation exercise program remains unknown. Previous research has reported the association between depression and performance on incremental treadmill exercise tests with varied results. Patients with depression or depressive-mood-related-symptoms have demonstrated statistically significant improvements in exercise test performance and when compared with patients without depression were equally likely to demonstrate improvement in exercise test scores (Hamm et al., 2004; Milani et al., 1996). In both studies, there were significant improvement in the depression scores for the groups categorised at baseline as depressed (Hamm et al., 2004; Milani et al., 1996). In addition, Milani et al. (1996) reported that the group of patients identified as depressed at baseline were significantly younger than the non-depressed group. It is not known if age or the improvement in depression affected the 10 m ISWT results in this study. When age and gender were accounted for, Glazer, Emery, Frid, and Banyasz (2002) demonstrated that depression identified at baseline accounted for 9% ($p < .07$) of the variance in improvement of peak oxygen consumption, on a treadmill symptom-limited exercise test. There was no association found between treadmill time and changes in a different yet related construct, emotional state, measured by the Profile of Mood states questionnaire (Beniamini et al., 1997). Although the impact of depression on the change in 10 m ISWT over a cardiac rehabilitation program is unclear, it is possible to hypothesise that the change

in 10 m ISWT will be less, and the effect sizes smaller, in groups of patients with depression or depressive-symptoms.

Over a cardiac rehabilitation exercise program, improvements in physical fitness and functional capacity often parallel improvements in health-related quality of life (Dugmore et al., 1999; Kavanagh et al., 1996). The association between the change in 10 m ISWT outcomes with other measures of health-related quality of life or self-efficacy over a cardiac rehabilitation program remains largely unknown. Three studies have reported weak but significant positive associations between changes in disease specific health-related quality of life scales and changes in the 6MWT or treadmill cardiopulmonary exercise test in patients with chronic heart failure (Flynn, Pina, et al., 2009; Kavanagh et al., 1996) and coronary heart disease (Ades et al., 1999). Changes in the Kansas City Cardiomyopathy Questionnaire (KCCQ) summary score were positively associated with changes in exercise time on cardiopulmonary exercise test ($r = .28$, $p < .001$), peak oxygen consumption ($r = .21$, $p < .001$), and 6MWT distance ($r = .18$ $p < .001$) (Flynn, Pina, et al., 2009). Kavanagh et al. (1996) reported significant correlations between the changes in the Chronic Heart Failure Questionnaire and the changes in 6MWT and treadmill cardiopulmonary exercise test, with the correlations stronger for the 6MWT than the cardiopulmonary exercise test. Ades et al. (1999) reported a positive association between the change in MOS SF-36 scale score for Physical Function and peak oxygen uptake ($r = .22$, $p = .02$), and change in peak exercise capacity measured by METs ($r = .17$ $p = .01$). Beniamini et al. (1997) found no significant correlation between changes in treadmill time and changes in the MOS SF-36 variables. It is possible that the 6MWT being a submaximal walk test more closely relates to the subjective health-related quality of

life questionnaires than the cardiopulmonary exercise test (Beniamini et al., 1997), but it is not known if the same applies for the 10 m ISWT.

The number of scheduled cardiac rehabilitation sessions attended has been used as a crude measure of adherence (Jolly et al., 2009). It has been suggested that the self-regulatory self-efficacy skills required to attend the program are important to the success of the rehabilitation process, and higher levels of self-regulatory self-efficacy behaviours are positively associated with attendance (Woodgate, Brawley, & Weston, 2005). It is not known if there is a dose-response relationship between cardiac rehabilitation and changes in the 10 m ISWT. No studies were found that investigated the association between the number of cardiac rehabilitation sessions attended and the change in 10 m ISWT distance walked. Earlier research has reported no significant difference in the change in METs levels attained in treadmill exercise testing with adherence to cardiac rehabilitation (Fontana, Kerns, Rosenberg, Marcus, & Colonese, 1986). However, more recent studies have demonstrated a positive dose-response relationship between cardiac rehabilitation and a lower risk of death (Hammill, Curtis, Schulman, & Whellan, 2010; Suaya, Stason, Ades, Normand, & Shepard, 2009) and lower risk of future myocardial infarction (Hammill et al., 2010). The association between adherence and thorough attendance at cardiac rehabilitation and functional outcomes remains unknown. It is possible to hypothesise that those who adhere to a cardiac rehabilitation program, measured through their attendance at a program, will also demonstrate greater improvement in the 10 m ISWT scores.

It is likely that patients who rate themselves as improved after a cardiac rehabilitation program will have greater exercise test change scores than those who rate themselves as unchanged or worse (Gremeaux et al., 2011; Houchen-Wolloff, Boyce, & Singh,

2015). Houchen-Wolloff et al. (2015) reported a statistically significant difference in change scores for the distance walked in the 10 m ISWT for groups of participants who reported that they were better, slightly better and unchanged after a six-week cardiac rehabilitation program. Gremeaux et al. (2011) also reported a statistically significant difference in the change scores for the distance walked in the 6MWT for groups of participants who reported they were improved compared with those reporting no improvement or worsening of symptoms. However, they did not find a statistically significant difference in the results for the 200 m fast walk test. These results suggest that it is reasonable to hypothesise that there will be statistically significant differences in the change scores for the 10 m ISWT distance based on participants global rating of change.

There is limited evidence to support the responsiveness of the 10 m ISWT as a measure of change in physical fitness and functional capacity in a mixed cardiac rehabilitation population. Evidence supports that there is a statistically significant increase in walk distance and a moderate effect size after a cardiac rehabilitation exercise program (Fowler et al., 2005; Frizelle et al., 2004; Tobin & Throw, 1999) and that the change in the 10 m ISWT distance is positively associated with global rating of change scores after cardiac rehabilitation (Houchen-Wolloff et al., 2015). The association between the change in 10 m ISWT distance and depression scores and health-related quality of life, as well as adherence to cardiac rehabilitation, remain unknown. Previous research in related areas supports the development of hypotheses that patients without depression or depressive-symptoms will make greater improvements in the 10 m ISWT than patients with depression or depressive-symptoms and that there will be a positive association between changes in the 10 m

ISWT distance and changes in the MOS SF-36 scales scores for Physical Function and Role Physical, and adherence to cardiac rehabilitation.

The aim of this chapter was to determine the evidence for both the construct validity and the responsiveness (longitudinal validity) of the 10 m ISWT as an objective measure of physical fitness and functional capacity in a mixed cardiac rehabilitation population. This study aimed to determine: first, the evidence through *a-priori* hypothesis testing of the construct validity of the 10 m ISWT in a mixed cardiac rehabilitation population utilising the pre-program 10 m ISWT outcomes; and second, the responsiveness of the 10 m ISWT in a mixed cardiac rehabilitation population through *a-priori* hypothesis testing utilising change scores.

5.3 Method

Research used a quantitative pre-post design on participants referred to cardiac rehabilitation.

5.3.1 Ethics approval and consent.

Ethics approval was granted by La Trobe University, University Human Ethics Committee (Reference Number 10-082) and St John of God Health Care Human Research Ethics Committee (Reference Number 438) (see Appendix 2). All participants enrolled in the investigation read and signed a Participant Information and Consent Form that was approved by the relevant ethics committees (see Appendix 3).

5.3.2 Participants.

Sample size calculation was based on the recommendations that there be a minimum of 50 patients in construct validity analysis (Terwee et al., 2007).

5.3.2.2 Recruitment procedures.

The recruitment procedures followed those described in Chapter 2, Section 2.3.2.2. Recruitment began in March 2011 and finished in October 2011.

5.3.3 Procedure.

Testing took place in the physiotherapy department of the participating hospital during the pre-program cardiac rehabilitation assessment and the post-program cardiac rehabilitation assessment. All tests took place under the same conditions by the same investigator.

5.3.3.1 Pre-program assessment.

In a single session, participants completed an initial subjective assessment to record baseline demographic details, self-reported physical activity, two 10 m ISWTs (labelled ISWT 1 and ISWT 2), and completed three questionnaires: the MOS SF-36, the Brief Case Find for Depression; and an exercise self-efficacy questionnaire. Each participant completed the assessment in the order described.

Baseline demographic information collected included age, date of birth, gender, cardiac diagnosis and intervention, cardiac signs and symptoms, cardiac medications,

relevant medical history and other relevant medications, usual mobility and use of gait aid. This information along with the date and time of assessment were recorded on a data-recording sheet (see Appendix 4).

5.3.3.2 Comprehensive cardiac rehabilitation intervention.

At the commencement of the program, participants underwent an individual assessment by the cardiac rehabilitation nurse co-ordinator to gather a detailed medical history. Participants then completed an eight-week cardiac rehabilitation program that included a 60-minute exercise program and a 60-minute group education and discussion forum.

The exercise program consisted of 10 minutes warm-up activity that included balance exercises, followed by 20 minutes of individualised circuit training using mostly resistance exercises, 10 to 20 minutes of continuous aerobic activity such as walking, and a 10 minute cool-down activity. An experienced physiotherapist and cardiac rehabilitation nurse coordinator supervised the exercise class. Participants were encouraged to exercise between levels 11 and 13 on the Borg rate of perceived exertion 6-20 scale without exceeding 13, that is perceived exertion should fall between light and somewhat hard, and not exceed hard (Borg, 1990). All participants were given individualised advice on completing appropriate home exercise. The physiotherapist commenced the home exercise program at a low or moderate intensity, with the aim to progress to a total of 150 to 300 minutes of low-moderate intensity each week.

There were eight group education topics including heart disease and management, risk factor reduction, medication use, exercise and physical activity, wellness and relaxation, diet and nutrition, food label reading, and mood and emotions. Information provided during these education sessions was general in nature and other than the exercise program, no individualised or detailed intervention plans were provided. There were no specific individual interventions or actions taken to address anxiety or depressive related symptoms in participants.

5.3.2.1 Eligibility criteria for participants.

The eligibility criteria were the same as those described in Chapter 2, Section 2.3.2.1.

5.3.3.3 Post-program assessment.

In a single session, participants completed a post-program assessment, and this occurred within one week of finishing the cardiac rehabilitation program. The assessment recorded: a global rating of change score, self-reported physical activity, two 10 m ISWTs (labelled ISWT 1 and ISWT 2), the MOS SF-36 as a measure of health-related quality of life, the Brief Case Find for Depression screen for likelihood of depression and the exercise self-efficacy questionnaire, in that order.

5.3.3.4 Outcome measure: baseline demographic information.

Body composition measures during the pre-program assessment included height, weight, and BMI. Height was measured using a stadiometer and weight on digital scales that were regularly calibrated. BMI was calculated as per the description in Chapter 2, Section 2.3.3.1.

During the post-program assessment, weight was measured using the same digital scales as the pre-program assessment and BMI recalculated. Height was not reassessed during the post-program assessment.

5.3.3.5 Outcome measure: self-reported physical activity.

During the pre-program assessment, before the completion of the 10 m ISWT, participants were asked to recall their exercise activities for the preceding week. Patients did not complete a standardised questionnaire; this information was based on a one-on-one interview by the investigator with the participant. Recall questions were used as a guide to elicit information on the exercise activities that were completed. The length of time, how often and the intensity of any exercise or physical activity recalled by the participant was clarified. Questions specific to walking were asked, such as the location of walking track, time, speed and perceived intensity.

Allocation to one of three groups was based on recommendations defining sedentary, light, and moderate to vigorous intensity exercise, see Table 5.1 (Norton, Norton, & Sadgrove, 2010; Pate et al., 1995). Exercise was defined as a planned activity designed to improve or maintain physical fitness and is considered a subset of physical activity (Caspersen, Powell, & Christenson, 1985). The amount of weekly exercise reported by the participant was judged according to published guidelines.

Table 5.1

Framework for Classification of Sedentary, Light, Moderate and Vigorous Intensity Exercise

Intensity	Measures	Description:	Examples
Sedentary	METs < 1.6 RPE < 8	Activities involve little additional movement and a low energy requirement	Sitting, watching TV, reading,
Light	METs 1.6-3 RPE 8-10	Activity does not cause a noticeable change in breathing rate, low and sustainable intensity.	Slow walk 1-3 km/hr Indoor household walking Golf with buggy
Moderate	METs 3-6 RPE 11-13	Able to maintain a conversation during the activity	Brisk walk 3-6 km/hr Cycling < 15 km/hr Swimming Golf (without buggy) Social tennis
Vigorous	METs 6-9 RPE 14-16	Conversation cannot be maintained, breathing hard or panting	Jogging Brisk walk with hills or additional load Cycling > 15 km/hr or hill training Fast swimming

Note. METs = metabolic equivalent; RPE = rate of perceived exertion. The classification of sedentary, light, moderate and vigorous intensity was based on the information provided by (Norton et al., 2010, pp. 497-500; Pate et al., 1995, p. 404).

The National Heart Foundation of Australia recommended that adults with heart disease participate in a minimum of 30 minutes of physical activity at a moderate intensity on most if not all days of the week (Briffa et al., 2006). General global guidelines recommend 150 to 300 minutes of activity at a moderate intensity, or 75 to 150 minutes of activity at a vigorous intensity, or an equivalent combination for adults (Brown, Bauman, Bull, & Burton, 2012; Tremblay et al., 2011; World Health Organization, 2010) and older adults (World Health Organization, 2010). When a

decision on the level of intensity was not obvious, MET tables were used to help describe the intensity of exercise consistently (Ainsworth et al., 2000). The sedentary group had no planned regular activities designed to improve or maintain physical fitness. The low intensity group participated in planned regular low-level exercise but of insufficient frequency or duration and intensity to meet the exercise guidelines. The moderate-vigorous intensity group participated in regular exercise that was of sufficient frequency or duration and intensity to meet the exercise guidelines.

5.3.3.6 Outcome measure: 10 m ISWT.

Participants completed two 10 m ISWTs in the pre-program assessment and two 10 m ISWTs in the post-program assessment. The procedure for completing the 10 m ISWT was the same as that described in Chapter 3, Section 3.3.3.2.

5.3.3.7 Outcome measure: Medical Outcomes Study 36-item short form.

The MOS SF-36 was described and developed by Ware and Sherbourne (1992). The 36-item questionnaire includes standardised responses for participants to describe aspects of their health-related quality of life, with scoring providing information on eight scales of health status: Physical Functioning, Role Physical, Bodily Pain, General Health, Vitality, Social Functioning, Role Emotional and Mental Health (Table 5.2) (Jette & Downing, 1994; Ware & Sherbourne, 1992). The MOS SF-36 scale for Physical Functioning comprises 10 questions that relate to physical activities, five of these activities are related to moderate or vigorous intensity exercise or walking progressive distances from 100 metres to more than a kilometre. The MOS

SF-36 is a well-known and widely used generic health-related quality of life measure applicable to the cardiac rehabilitation population (Brown, 2003; Dempster & Donnelly, 2000; McKee, 2009) and recommended by the Australian Cardiac Rehabilitation Association (Goble & Worcester, 1999, p. 164).

In general, patients enrolled in an outpatient cardiac rehabilitation program present at the pre-program assessment with impaired health-related quality of life scores (Cohen et al., 1999; Jette & Downing, 1994; Morrin, Black, & Reid, 2000; Quittan et al., 1999). The Australian Bureau of Statistics have published normative data for the Australian population with and without heart disease, see Table 5.2 (Australian Bureau of Statistics, 1995). Good levels of retest reliability of the MOS SF-36 when completed within two weeks of the commencement of cardiac rehabilitation has been demonstrated, with *ICC*(2,1) for all but one of the MOS SF-36 scale scores of health status ranged from .70 to .84, and for the scale Role Emotional an *ICC* of .37 (Jette & Downing, 1994).

Table 5.2

Australian Population Norms for the MOS SF-36 for Australian Residents With and Without Heart Disease

MOS SF-36 scales of health status	No heart disease	Heart disease	Meaning of scores (Australian Bureau of Statistics, 1995; Ware & Sherbourne, 1992)
Physical Functioning	79.1	65.8	0: limited a lot in performance of all physical activities including personal activities of daily living
			100: performs all physical activities without limitation including vigorous tasks
Role Physical	77.1	51.2	0: problems with work or other daily activities due to physical health
			100: no problems with work or other daily activities as a result of physical health
Bodily Pain	74.8	60.0	0: severe and limiting pain
			100: no pain or limitation from pain
General Health	70.4	52.7	0: personal health is perceived as poor and expected to get worse
			100:personal health perceived as excellent
Vitality	64.3	54.7	0: tired and worn out all of the time
			100: feels full of life and energy all of the time
Social Functioning	84.9	73.6	0: Problems with normal social activities are extreme and frequent
			100: Undertakes normal social activities without interference
Role Emotional	82.1	70.9	0: problems with work or other daily activities due to emotional health
			100: no problems with work of other daily activities as a result of emotional health
Mental Health	76.3	71.8	0: feelings of depression and nervousness all of the time
			100: peaceful, happy and calm all of the time

Note. MOS SF-36 = Medical Outcomes Study 36-item short form. Information adapted from the Australian Bureau of Statistics (1995, p. 21) and Ware and Sherbourne (1992, p. 475)

5.3.3.8 Outcome measure: Brief Case Find for Depression.

The Brief Case Find for Depression is a simple, four-item case-finding instrument for probable depression that can be completed in a very short amount of time (Clark, McKenzie, Marshall, & Smith, 1994; Jefford et al., 2004). The test has good sensitivity and negative predictive power, but poor specificity (Jefford et al., 2004). The researcher reads the four questions to the participant: *Over the past couple of weeks, have you (A) been having restless or disturbed nights?; (B) been feeling unhappy or depressed?; (C) felt unable to overcome your difficulties?; (D) been dissatisfied with the way you've been doing things?* (Jefford et al., 2004, p. 901). Scoring for probable depression was based on the algorithm, answering yes to [A or B] AND [C or D] (Jefford et al., 2004).

5.3.3.9 Outcome measure: exercise self-efficacy scale.

The Bandura exercise self-efficacy scale is an 18-item questionnaire measuring perceived confidence in a participant's ability to exercise regularly (Bandura, 2006). The scale was modified to suit an Australian population and then the psychometric properties of the modified scale were tested in an Australian cardiac rehabilitation context (Everett et al., 2009). Good internal consistency for the 18 items with the item total scale correlations ranging from .59 to .84 was demonstrated (Everett et al., 2009). The participants had a mean score of 103 (*SD* 35) on entry to cardiac rehabilitation and there were no floor or ceiling effects (Everett et al., 2009).

After the questionnaire was completed, it was checked for errors and completeness, for example, in the case of two responses circled for one item or missing responses

the participant was asked to clarify if it was intentional or not, and if relevant given the opportunity to amend. The questionnaire was scored out of 180 points, in the case of intentionally missing responses the questionnaire was scored and then converted to a score out of 180 points.

5.3.3.10 Outcome measure: Global Rating of Change.

Participants were asked to score their Global Rating of Change in the post-program assessment before completing any of the assessment tests. Participants were asked: *Do you believe your physical fitness has improved significantly, improved a little, is about the same, is slightly worse or significantly worse?*. Results were recorded using the 5-point scale and then recoded into a 3-point scale, improved, same or worse.

5.3.4 Statistical Analysis.

Participant characteristics were recorded to allow description of the study sample. Where relevant, means were expressed followed by standard deviation. Data were not imputed in the case of missing values. The 10 m ISWT results were checked for floor and ceiling effects. Terwee et al. (2007) recommended an overall positive rating for the absence of floor and ceiling effects if there were no floor or ceiling effects present in a sample size of at least 50 patients (Terwee et al., 2007). Small floor or ceiling effects have been described as occurring between 1 to 15% of the sample and moderate floor or ceiling effects in greater than 15% of the sample (McHorney & Tarlov, 1995). The relative reliability and measurement error of the 10 m ISWT was calculated using the methods described in Chapter 2, Sections 2.3.4.3 and 2.3.4.4 for the walk tests completed in the pre-program and the post-program assessment.

The longitudinal reliability of the 10 m ISWT was measured in two ways. First, by calculating the ICC for the sample and the subsample of participants who reported themselves as unchanged on the global rating of change. The method of calculating the ICC was described in Chapter 2, Section 2.3.4.3. Second, as the reliability of the change scores, using the formula as follows:

$$Reliability\ of\ change\ score = \frac{R_{XX} + R_{YY} - 2r_{XY}}{2.0 - 2r_{XY}}$$

Where R_{XX} is the correlation of pre-program 10 m ISWT (i.e., ISWT 1 and ISWT 2) and R_{YY} is the correlation of the post-program 10 m ISWT (i.e., ISWT 3 and ISWT 4), r_{XY} is the correlation of the pre-program and post-program scores (i.e., ISWT 2 and ISWT 4).

Eleven hypotheses were determined, *a-priori*, to assess the evidence for construct validity and responsiveness. Each hypothesis followed the recommendation of the COSMIN group and included information on the expected size and direction of the association (Mokkink et al., 2010b). Each hypothesis was tested using the results of both the first and the second 10 m ISWT for the session. For a correlation using continuous data, a Pearson's product-moment correlation coefficient with 95% confidence intervals was calculated and interpreted with the following guidelines: strong, a correlation greater than 0.75, moderate between 0.50 and 0.75 and weak less than 0.5 (Portney & Watkins, 2000, p. 565). For assessing for statistical differences between groups, a significance level of $p < .05$ was set.

5.3.4.1 Evidence for construct validity of the 10 m ISWT using a-priori hypothesis testing.

Hypothesis 5.1: Participants who report regular physical exercise at a moderate-vigorous intensity will walk significantly further in the 10 m ISWT compared with those who report exercise at a low intensity who in turn will walk significantly further than those who report to be sedentary. The primary hypothesis will be tested with a one-way ANOVA and post-hoc comparisons.

Hypothesis 5.2: There will be a significant weak negative correlation ($r < |-.5|$) between measures of adiposity and the distance walked in the 10 m ISWT. Specifically, Hypothesis 5.2.1: There will be a weak negative correlation ($r < |-.5|$) of the waist-hip ratio and the distance walked in the 10 m ISWT. Hypothesis 5.2.2: There will be a weak negative ($r < |-.5|$) correlation between weight and 10 m ISWT distance; and hypothesis 5.2.3 is that there will be a weak negative correlation between BMI and 10 m ISWT distance. A Pearson's product-moment correlation with 95% confidence intervals will be used to calculate the association.

Hypothesis 5.3: There will be a significant weak-moderate positive correlation ($r < .75$) between the MOS SF-36 scale score of Physical Functioning and the distance walked in the 10 m ISWT. A Pearson's product-moment correlation with 95% confidence intervals will be used to calculate the association.

Hypothesis 5.4: There will be a weak positive correlation ($r < .50$) between the MOS SF-36 scale score for Role Physical and the distance walked in the 10 m ISWT and this correlation will be less than that demonstrated in hypothesis 5.3. This correlation

will be calculated using a Pearson's product-moment correlation coefficient with 95% confidence intervals.

Hypothesis 5.5: There will be a non-significant correlation between the scores of the remaining six MOS SF-36 scales and the 10 m ISWT, with the MOS SF-36 scale score of Role Emotional the least related to the distance walked in the 10 m ISWT. Hypothesis 5.5.1: There will be a non-significant correlation between the MOS SF-36 scale score of Bodily Pain and the 10 m ISWT distance walked at the pre-program assessment. Hypothesis 5.5.2: There will be a non-significant correlation between the MOS SF-36 scale score of General Health and the 10 m ISWT distance walked at the pre-program assessment. Hypothesis 5.5.3: There will be a non-significant correlation between the MOS SF-36 scale score of Vitality and the 10 m ISWT distance walked at the pre-program assessment. Hypothesis 5.5.4: There will be a non-significant correlation between the MOS SF-36 scale score of Social Functioning and the 10 m ISWT distance walked at the pre-program assessment. Hypothesis 5.5.5: There will be a non-significant correlation between the MOS SF-36 scale score of Role Emotional and the 10 m ISWT distance walked at the pre-program assessment. Hypothesis 5.5.6: There will be a non-significant correlation between the MOS SF-36 scale score of Mental Health and the 10 m ISWT distance walked at the pre-program assessment. Hypothesis 5.5.7: The MOS SF-36 Role Emotional scale score will be the MOS SF-36 scale least related to the distance walked in 10 m ISWT.

Hypothesis 5.6: There will be a weak correlation ($r < .50$) between the exercise self-efficacy score and the distance walked in the 10 m ISWT. This correlation will be calculated using a Pearson's product-moment correlation coefficient with 95% confidence intervals.

Hypothesis 5.7: Hypothesis 5.7.1: That the difference in the distance walked in the 10 m ISWT between participants with likely depression compared with those with unlikely depression will not be statistically significant ($p > .05$) at the baseline assessment. Hypothesis 5.7.2: That the difference in the distance walked in the 10 m ISWT between participants with likely depression compared with those with unlikely depression will be statistically significant ($p > .05$) at the post-program assessment. An independent samples *t*-test will be used to test for statistically significant differences.

5.3.4.2 Evidence for responsiveness of the 10 m ISWT over an eight-week cardiac rehabilitation program using a-priori hypothesis testing.

Hypothesis 5.8: Participants of an eight-week outpatient cardiac rehabilitation exercise program will demonstrate an improvement in the 10 m ISWT distance between their pre-program and post-program assessment, there will be a moderate effect size, and the differences and effect size will be larger in patients without likely depression than those with likely depression. The internal responsiveness of the 10 m ISWT to measure changes in physical fitness over the duration of cardiac rehabilitation will be described (Husted, Cook, Farewell, & Gladman, 2000) to measure the magnitude of change. First, a paired samples *t*-test will be used, and will focus on the statistical significance of any observed change between pre-program and post-program scores. Second, internal responsiveness will be measured through standardised effect size, calculated using the following formula:

$$ES = \frac{X_{post-program} - X_{pre-program}}{SD_{pre-program}}$$

Where $X_{pre-program}$ and $X_{post-program}$ are the mean distances walked in the 10 m ISWT during the pre-program and the post-program assessment and $SD_{pre-program}$ is the variability or standard deviation of the pre-program scores (Husted et al., 2000). Results were interpreted based on recommendations by Cohen (1977, p. 24) where less than 0.2 was a null effect size, .2 to .49 a small effect size, .50 to .79 a moderate effect size, and scores of .8 or higher a large effect size.

Hypothesis 5.9: Those unlikely to have depression will make greater improvements, across a cardiac rehabilitation program, in the 10 m ISWT distance than those with likely depression, and this difference will be statistically significant and the effect size will be larger in patients without likely depression than those with likely depression. This hypothesis will be tested using an independent samples *t*-test.

Hypothesis 5.10: There will be significant weak correlations ($r < .50$) between changes in the MOS SF-36 scale scores that relate to physical health and changes in the 10 m ISWT distance, and non-significant correlations ($p > .05$) between changes in the remaining MOS SF-36 scale scores and changes in the 10 m ISWT. Specifically, that there will be a positive weak correlation between changes in the MOS SF-36 scale score of Physical Functioning and the change in distance walked in the 10 m ISWT and the changes in the MOS SF-36 scale score of Role Physical and changes in the 10 m ISWT. The correlations between the changes scores of the remaining six MOS SF-36 scales and the change in 10 m ISWT distance will be non-

significant. These correlations will be calculated using a Pearson's product-moment correlation coefficient with 95% confidence intervals.

Hypothesis 5.11: There will be a significant weak positive association ($r < .50$) between attendance at a cardiac rehabilitation program and improvement in 10 m ISWT distance. This correlation will be calculated using a Pearson's product-moment correlation coefficient with 95% confidence intervals.

Hypothesis 5.12: Participants who report improvement in Global Rating of Change will demonstrate a greater improvement in the 10 m ISWT distance over an eight-week outpatient cardiac rehabilitation program than those who report no change or deterioration in Global Rating of Change. This difference will be statistically significant. Those who report no change or a deterioration in a Global Rating of Change will not show a statistically significant difference in the change in 10 m ISWT walk distance. External responsiveness reflects the extent that the changes in 10 m ISWT scores reflect changes in other health measurement scores. Statistical methods to measure the external responsiveness for the 10 m ISWT in cardiac rehabilitation were one-way ANOVA with three groups with post-hoc analysis.

5.4 Results

5.4.1 Characteristics of the sample.

Fifty-two patients were recruited on admission to a cardiac rehabilitation program. Tables 5.3 and 5.4 show the characteristics of the sample at baseline. Table 5.4 also shows the characteristics of the sample at the post-program assessment. The mean age

of participants was 68 years (*SD* 10) with the youngest participant 46 years and the oldest participant 91 years. All participants were able to mobilise independently, one participant required the use of a single point stick to complete the 10 m ISWTs. The pre-program assessment was completed over a mean of 29 days (*SD* 19) following the most recent cardiac intervention and this time ranged from 7 to 97 days.

Table 5.3

Characteristics of the Sample

Characteristics	Total (*n* = 52)
Age	68 (10)
Gender *n* Male:Female	40:12
Intervention *n* (%)	
Revascularisation Procedure	31 (63)
Medical management	14 (27)
Valve surgery	3 (5)
PPM and or ICD	4 (8)
BMI category *n* (%)	
Underweight	0
Healthy	10 (19)
Overweight	25 (48)
Obese	16 (31)
Morbidly obese	1 (2)

Note. PPM = permanent pacemaker; ICD = implantable cardioverter defibrillator.

The mean time that elapsed between the pre-program assessment and the post-program assessment was nine weeks, and ranged from seven to 11 weeks. One post-program assessment was completed before 8 weeks as the participant was finishing the program early for reasons unrelated to the study. Participants attended an average of 7 (*SD* 1) cardiac rehabilitation classes and this ranged from five to eight classes.

Table 5.4

Basic Pre-Program and Post-Program Objective Assessment Measures

Characteristics	Pre-program assessment	Post-program assessment
Height (cm)	172 (7)	Not assessed
Weight (kg)	84 (15)	84 (14)
BMI (kg/m^2)	29 (5)	28 (5)
Waist (cm)	99 (12)	99 (12)
Waist:Hip ratio	0.97 (0.08)	0.97 (0.07)
Males	1.00 (0.05)	0.99 (0.05)
Females	0.87 (0.07)	0.89 (0.07)
Resting HR (beats/min)	71 (11)	68 (10)
Resting BP (mmHg)	129 (14) / 75 (7)	125 (14) / 71 (8)
Resting SpO2 (%)	98 (1)	98 (1)
Resting RR (breaths/min)	17 (1)	17 (1)

Note. BMI = body mass index; HR = heart rate; BP = blood pressure; SpO_2 = oxygen saturation measured via pulse oximetry; RR = respiratory rate.

Table 5.4 shows the pre-program and post-program general assessment measures. The mean weight of participants was 84 kg (*SD* 15) at the pre-program assessment and 84 kg (*SD* 14) at the post-program assessment and ranged from 60 kg to 145 kg for both the pre-program assessment and the post-program assessment. The mean BMI for the sample was 29 kg/m^2 (*SD* 5) indicating that on average the sample was overweight. The pre-program BMI ranged from 20 to 47 kg/m^2. Nine (17%) participants were in the BMI healthy range, 26 (50%) were overweight, 16 (31%) were obese and one (2%) participant was morbidly obese. The mean BMI for the sample at the post-program assessment was 28 kg/m^2 (*SD* 5), indicating that on average the sample was overweight. The post-program BMI ranged from 20 to 47 kg/m^2. At the post-program assessment, 11 (21%) participants had a BMI in the healthy range, 25 (48%) were overweight, 15 (29%) were obese and one (2%) was morbidly obese.

5.4.1.1 10 m ISWT.

All participants completed the 10 m ISWT until volitional exhaustion and no test was stopped prematurely by the researcher. There was a positive rating for floor and ceiling effects in this sample. All participants were able to complete at least one shuttle of the 10 m ISWT. During the pre-program assessment, no participant completed all levels of the test. During the post-program assessment, one participant (2%) completed all levels of the test in both ISWT 3 and ISWT 4.

For the pre-program assessment, the mean distance walked in ISWT 1 was 378 m (*SD* 164), with a range from 110 to 760 m and for ISWT 2 the mean distance walked was 396 m (*SD* 164), with a range from 120 to 780 m (Table 5.5). At the pre-program assessment the mean difference between the distance walked in the two 10 m ISWT was 17 m (SD_{diff} 14), and this difference was significant ($t(51) = 8.648$, $p < .001$). Men were more likely to walk further than women in both ISWT 1 ($t(50) = 2.484$, $p = .016$) and ISWT 2 ($t(50) = 2.605$, $p < .012$). There were no violations to normality, skewness or kurtosis for the distance walked in ISWT 1 or ISWT 2.

Table 5.5

Results of the 10 m ISWT

	Pre-program		Post-program	
10 m ISWT outcomes	ISWT 1	ISWT 2	ISWT 3	ISWT 4
10 m ISWT distance (m)	378 (164)	396 (164)	481 (206)	497 (211)
Men 10 m ISWT distance	408 (159)	427 (156)	516 (206)	531 (210)
Women 10 m ISWT distance	280 (149)	293 (151)	363 (167)	382 (178)
End of test findings				
Peak HR (beats/min)	100 (18)	105 (18)	103 (20)	105 (25)
SBP (mmHg)	155 (20)	161 (23)	155 (29)	159 (25)
DBP (mmHg)	83 (8)	84 (9)	81 (9)	83 (9)
RR (breaths/min)	29 (3)	31 (3)	30 (3)	31 (4)
SpO_2 (%)	98 (1)	98 (1)	98 (1)	98 (1)
RPP (mmHg/beats/min) median (mode)	12 (12)	13 (12)	12 (12)	13 (12)
2-minute HR recovery [(a)] *n* (%)	46 (88%)	41 (79%)	38 (73%)	37 (71%)
End of test symptoms *n* (%)				
No symptoms	15 (29%)	13 (25%)	9 (17%)	10 (19%)
Fatigue	1 (2%)	1 (2%)	1 (2%)	1 (2%)
Lower limb pain	8 (15%)	10 (19%)	10 (19%)	12 (23%)
Shortness of breath	30 (58%)	31 (60%)	36 (69%)	33 (63%)
Other	1 (2%)	2 (4%)	0	0

Note. ISWT = 10 m incremental shuttle walk test; HR = heart rate; bpm = beats per minute; SBP = systolic blood pressure; mmHg = millimetres mercury; DBP = diastolic blood pressure; RR = respiratory rate; SpO_2 = oxygen saturation measured with pulse oximetry; RPP = rate pressure product. [(a)] 2 minute HR recovery was defined as HR within 20% of baseline measure at two minutes.

For the post-program assessment, the mean distance walked in ISWT 3 was 481 m (*SD* 206), with a range from 150 to 1020 m, and for ISWT 4 the mean distance walked was 497 m (*SD* 211), with a range from 150 to 1020 m. At the post-program assessment the mean difference between the distance walked in the two 10 m ISWTs was 16 m (SD_{diff} 21), and this difference was significant ($t(51) = 5.448$, $p < .001$). Men were more likely to walk further than women in both ISWT 3 ($t(50) = 2.349$, $p < .023$) and ISWT 4 ($t(50) = 2.248$, $p < .029$). There were no violations to

normality, skewness or kurtosis for the distance walked in ISWT 3 or ISWT 4. Results of the walk test are shown in Table 5.5.

The relative reliability expressed by the $ICC_{agreement}$ model with 95% confidence intervals of ISWT 1 and ISWT 2 was .991, 95% CI [.866, .997] and for ISWT 3 and ISWT 4 was .992, 95% CI [.968, .997]. For men, the $ICC_{agreement}$ model with 95% confidence intervals of ISWT 1 and ISWT 2 was .989, 95% CI [.823, .997] and for ISWT 3 and ISWT 4 was .993, 95% CI [.970, .997]. For women, the $ICC_{agreement}$ model with 95% confidence intervals of ISWT 1 and ISWT 2 was .993, 95% CI [.907, .998] and for ISWT 3 and ISWT 4 as .989, 95% CI [.926, .996]. The mean difference between ISWT 1 and ISWT 3 was 103 m (*SD* 72) and between ISWT 2 and ISWT 4 was 101 m (*SD* 76).

The responsiveness (longitudinal reliability) of the distance walked in the 10 m ISWT for those patients who reported an unchanged global rating of change (n = 10) was calculated both using the first test and the second test (Table 5.6). The *ICC* measured for consistency remained high, suggesting that participant ranking did not alter during cardiac rehabilitation. The lower *ICC* values when calculated for absolute values reflect the systematic change in the absolute scores across a cardiac rehabilitation intervention. The reliability of the change score was .96.

Table 5.6

Longitudinal Relative Reliability of the 10 m ISWT

	ICC for consistency			*ICC* for absolute agreement		
		95% CI			95% CI	
10 m ISWT	*ICC*	*LL*	*UL*	*ICC*	*LL*	*UL*
Global Rating of Change: Unchanged (*n* = 10)						
First test (ISWT 1 and 3)	.99	.96	.997	.92	-.01	..99
Second test (ISWT 2 and 4)	.99	.96	.998	.94	.01	.99
Whole Sample (*n* = 52)						
First test (ISWT 1 and 3)	.93	.87	.96	.81	.01	.94
Second test (ISWT 2 and 4)	.92	.86	.95	.81	.04	.94

Note. ICC = intraclass correlation coefficient; CI = confidence interval; LL = lower limit; UL = upper limit.

5.4.1.2 Self-reported physical activity.

In the pre-program assessment, 18 participants (35%) reported no regular planned physical exercise and were classified as sedentary, 25 (48%) reported regular low intensity exercise and 9 (17%) reported regular moderate-vigorous intensity exercise that met the exercise guidelines described in Section 5.3.3.5. Walking was the most common exercise with 30 participants (58%) reporting this activity in their regular exercise program. The results of the demographic characteristics according to physical activity are shown in Table 5.7.

Table 5.7

Characteristics of the Sample According to Pre-Program Reported Levels of Physical Activity

Characteristic	Sedentary	Low Intensity	Moderate-Vigorous Intensity
Age	71 (10)	69 (10)	61 (11)
Gender *n* Male:Female	11:7	21:4	8:1
Intervention *n* (%)			
Revascularisation Procedure	11 (33)	14 (42)	8 (24)
Medical management	4 (26)	9 (64)	1 (7)
Valve surgery	1 (100)	0	0
PPM and or ICD	2 (50)	2 (50)	0
Days since cardiac event	25 (12)	30 (19)	44 (34)
Weight (kg)	86 (11)	82 (17)	88 (14)
BMI ($kg.m^{-2}$)	30 (4)	27 (5)	29 (4)
Waist:Hip ratio	0.96 (0.08)	0.97 (0.08)	0.98 (0.07)
ISWT 1 distance (m)	277 (145)	376 (123)	589 (93)
ISWT 2 distance (m)	292 (144)	394 (120)	609 (92)
ISWT 3 distance (m)	366 (167)	464 (158)	759 (141)
ISWT 4 distance (m)	377 (169)	476 (155)	792 (137)

Note. PPM = permanent pacemaker; ICD = implantable cardioverter defibrillator; BMI = body mass index.

5.4.1.3 Health-related quality of life MOS SF-36.

The results of the MOS SF-36 at the pre-program assessment are shown in Table 5.8. The scores for all scale scores indicate a lower health-related quality of life compared with the Australian norm for people without heart disease, presented in Table 5.2 (Section 5.3.3.7). Floor and ceiling effects were present and the number and percentage for the pre-program and post-program assessment are shown in Table 5.8.

Table 5.8

MOS SF-36 Results at the Pre-Program Assessment and Post-Program Assessment

MOS SF-36 Scale	Pre-program Assessment				Post-program Assessment					95% CI for the mean difference	
	Mean score (*SD*)	Minimum to Maximum score	Floor *n* (%)	Ceiling *n* (%)	Mean score (*SD*)	Minimum – Maximum score	Floor *n* (%)	Ceiling *n* (%)	Mean difference	LL	UL
Physical Functioning	68 (22)	20 to 95	0 (0)	0 (0)	80 (18)	30 to 100	0 (0)	4 (8)	12	7	17
Role Physical	41 (29)	0 to 100	6 (12)	3 (6)	70 (24)	13 to 100	0 (0)	8 (15)	29	21	37
Bodily Pain	66 (25)	0 to 100	1 (2)	11 (21)	78 (20)	41 to 100	0 (0)	17 (27)	11	5	17
General Health	65 (20)	15 to 100	0 (0)	2 (4)	68 (15)	30 to 100	0 (0)	1 (2)	2	-1	6
Vitality	50 (18)	10 to 85	0 (0)	0 (0)	59 (13)	25 to 85	0 (0)	0 (0)	9	6	12
Social Functioning	63 (27)	0 to 100	1 (2)	10 (19)	82 (20)	38 to 100	0 (0)	20 (38)	19	12	26
Role Emotional	75 (30)	0 to 100	1 (2)	23 (44)	81 (20)	17 to 100	0 (0)	35 (67)	14	7	21
Mental Health	66 (14)	24 to 88	0 (0)	0 (0)	73 (11)	48 to 88	0 (0)	0 (0)	7	4	10

Note. MOS SF-36 = Medical Outcomes Study 36-item short form; *SD* = standard deviation; CI = confidence interval; LL = Lower limit; UL = Upper limit.

In the pre-program assessment, floor effects were observed in low numbers in the following MOS SF-36 scales: Role Physical (12%), Bodily Pain (2%), Social Functioning (2%) and Role Emotional (2%). Ceiling effects were observed in the pre-program MOS SF-36 scales: Bodily Pain (21%), Social Functioning (20%) Role Emotional (44%), Role Physical (6%) and General Health (4%). There were no floor effects in the post-program assessment. A number of participants experienced a ceiling effect in the post-program assessment of MOS SF-36 scales: Physical Functioning (8%), Role Physical (15%), Bodily Pain (27%), General Health (2%), Social Functioning (38%), and Role Emotional (67%) (Table 5.8).

5.4.1.4 Brief Case Find for Depression.

The Brief Case Find for Depression identified 10 participants (19%) with probable or likely depression symptoms at the pre-program assessment and eight participants (15%) with probable or likely depression symptoms at the post-program assessment. There were five men (13%) and five women (42%) in the pre-program assessment identified with probable depression and in the post-program assessment three men (8%) and five (42%) women.

5.4.1.5 Results for exercise self-efficacy.

All participants completed Bandura's Exercise self-efficacy questionnaire in the pre-program and the post-program assessment. There were no floor or ceiling effects. The mean score for the pre-program assessment was 101 (*SD* 34), with scores ranging from 21 to 163. The mean score for the post-program assessment was 116 (*SD* 33),

with scores ranging from 45 to 174. The mean difference was 15 (SD_{diff} 21) 95% CI [9, 21].

5.4.2 Evidence for construct validity using *a-priori* hypothesis testing.

5.4.2.1 Hypothesis 5.1: participants who report regular physical exercise at a moderate-vigorous intensity will walk significantly further in the 10 m ISWT compared with those who report exercise at a low intensity who in turn will walk significantly further than those who report to be sedentary.

There was a significant difference in the distance walked in the 10 m ISWT across the three groups (sedentary, low intensity and moderate-vigorous intensity) (Table 5.9). *Post hoc* comparisons showed that this was significant with all combinations, the distance walked was less in those who were sedentary than those who either participated in low intensity exercise (ISWT 1: $p = .015$; ISWT 2: $p = .012$), or moderate-vigorous exercise (ISWT 1: $p < .001$; ISWT 2: $p < .001$). In addition, those who participated in low intensity exercise walked less than those who participated in regular moderate-vigorous intensity exercise (ISWT 1: $p < .001$; ISWT 2: $p < .001$). There were no interaction effects ($F(2,49) = .286$, $p = .752$). The results supported hypothesis 5.1.

Table 5.9

Distance Walked in the 10 m ISWT According to Self-Reported Physical Exercise

	10 m ISWT distance (m) *M* (*SD*)				
Test number	Sedentary (*n* = 18)	Low Intensity (*n* = 25)	Moderate-vigorous Intensity (*n* = 9)	*F*(2, 49)	*p*
ISWT 1	277 (145)	376 (123)	588 (93)	18.1	< .001
ISWT 2	292 (120)	394 (120)	609 (92)	19.2	< .001

5.4.2.2 Hypothesis 5.2: there will be a significant weak negative correlation (r < |-.5|) between measures of adiposity and the distance walked in the 10 m ISWT.

Table 5.10 shows the results of the correlation between the measures of adiposity and the distance walked in the 10 m ISWT. There was a significant weak correlation between the waist-hip ratio and the distance walked in the 10 m ISWT, and non-significant correlation for both weight and BMI, and the distance walked in the 10 m ISWT in the pre-program assessment. *Post hoc* analysis of gender showed a poorer association between the waist:hip ratio and 10 m ISWT distance (Table 5.11). The spread of scores is shown graphically in Figure 5.1. In the first 10 m ISWT the waist:hip ratio accounted for 2% of the variation in men and 2% of the variation in women, whereas for the group as a whole the waist:hip ratio accounted for 10%. For the second walk, the waist:hip ratio accounted for 1% of the variation in men and 2% for women, whereas for the group as a whole, the waist:hip ratio accounted for 10%.

This significant weak correlation was not seen in the post-program assessment (Table 5.12). The results did not support hypothesis 5.2.

Table 5.10

Associations between the 10 m ISWT and Measures of Adiposity in the Pre-Program Assessment

	Waist-hip ratio			Weight			BMI		
		95% CI			95% CI			95% CI	
10 m ISWT	*r*	LL	UL	*r*	LL	UL	*r*	LL	UL
ISWT 1	.317	.048	.543	.157	-.121	.412	-.096	-.359	.182
ISWT 2	.321	.053	.546	.154	-.124	.410	-.092	-.356	.186

Note. *r* = Pearson's product-moment correlation coefficient; CI = confidence interval; LL = lower limit; UL = upper limit.

Table 5.11

Associations between the 10 m ISWT and Waist:Hip ratio in the Pre-Program assessment for Men and Women

	Men (*n* = 40)			Women (*n* = 12)		
		95% CI			95% CI	
10 m ISWT	*r*	LL	UL	*r*	LL	UL
ISWT 1	.125	-.194	420	.140	-.401	.609
ISWT 2	.120	-.199	416	.123	-.415	.598

Note. *r* = Pearson's product-moment correlation coefficient; CI = confidence interval; LL = lower limit; UL = upper limit.

Table 5.12

Association between the 10 m ISWT and Measures of Adiposity in the Post-Program Assessment

	Waist-hip ratio			Weight			BMI		
		95% CI			95% CI			95% CI	
10 m ISWT	*r*	LL	UL	*r*	LL	UL	*r*	LL	UL
ISWT 3	.200	-.077	.448	.164	-.114	.418	-.076	-.342	.201
ISWT 4	.172	-.106	.425	.170	-.108	.423	-.072	-.338	.205

Note. *r* = Pearson's product-moment correlation coefficient; CI = confidence interval; LL = lower limit; UL = upper limit.

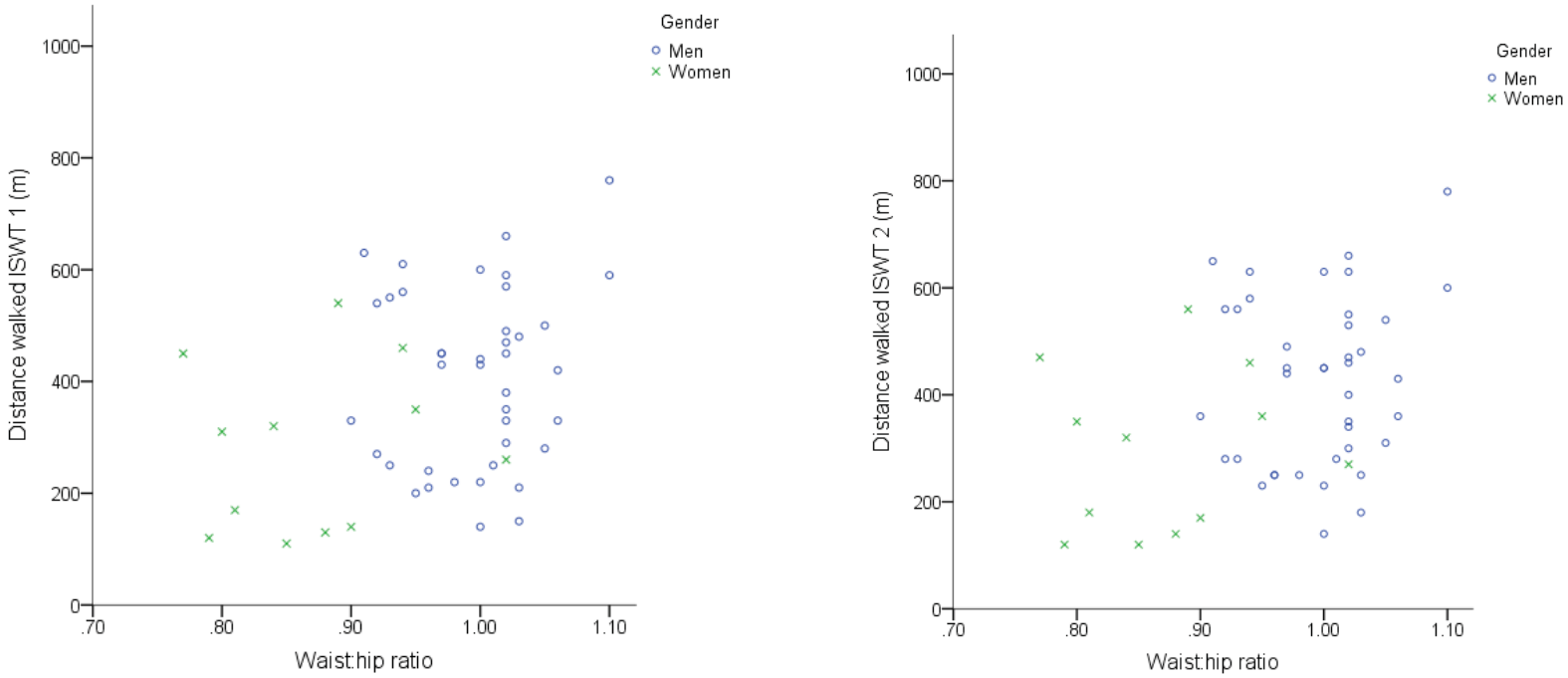


Figure 5.1. Scatter plot showing the spread of scores for the waist:hip ratio and the distance walked in the 10 m ISWT at the pre-program assessment.

5.4.2.3 Hypothesis 5.3: there will be a significant weak-moderate positive correlation (r < .75) between the MOS SF-36 scale score for Physical Functioning and the distance walked in the 10 m ISWT.

There was a moderate correlation between the MOS SF-36 scale score for Physical Functioning and the distance walked in the 10 m ISWT for the pre-program assessment (Table 5.13) and the post-program assessment (Table 5.14). The distance walked in the 10 m ISWT was able to explain up to 27% of the variance of the MOS SF-36 scale score for Physical Functioning in the pre-program assessment and up to 34% in the post-program assessment.

Gender differences were observed in the pre-program assessment but not the post-program assessment. For the men in the sample, the correlation was significant between the pre-program MOS SF-36 scale score for Physical Functioning and ISWT 1 distance was .451, 95% CI [.162, .669] and ISWT 2 was .444, 95% CI [.154, .664]. For the women in the sample the correlation was non-significant between the pre-program MOS SF-36 scale score for Physical Functioning and ISWT 1 and ISWT 2 of .393, 95% CI [-.234, .789] and .416, 95% CI [-.207, .799], respectively. In the post-program assessment, this correlation was moderate and significant for both men and women. For men, the correlation between the post-program MOS SF-36 Physical Functioning scale score and ISWT 3 and ISWT 4 was .526, 95% CI [.257, .720] and .534, 95% CI [.267, .725], respectively. For women, the correlation between the post-program MOS SF- Physical Functioning scale score, and ISWT 3 and ISWT 4 was 674, 95% CI [.163, .900] and .634, 95% CI [.094, .886], respectively.

This hypothesis was supported.

Table 5.13

Association between the 10 m ISWT and MOS SF-36 Scale Score for Physical Functioning in the Pre-Program Assessment

		95% CI		
10 m ISWT	*r*	*LL*	*UL*	r^2
ISWT 1	.515	.282	.691	.265
ISWT 2	.514	.280	.690	.264

Note. *r* = Pearson's product-moment correlation coefficient; CI = confidence interval; LL = lower limit; UL = upper limit; r^2 = coefficient of determination.

Table 5.14

Association between the 10 m ISWT and MOS SF-36 Scale Score for Physical Functioning in the Post-Program Assessment

		95% CI		
10 m ISWT	*r*	*LL*	*UL*	r^2
ISWT 3	.580	. 365	.736	.336
ISWT 4	. 576	. 360	.734	. 332

Note. *r* = Pearson's product-moment correlation coefficient; CI = confidence interval; LL = lower limit; UL = upper limit; r^2 = coefficient of determination.

5.4.2.4 Hypothesis 5.4: There will be a weak positive correlation (r < .50) between the MOS SF-36 scale score for Role Physical and the distance walked in the 10 m ISWT and this correlation will be less than that demonstrated in hypothesis 5.3.

For the pre-program assessment, there was a non-significant correlation between the MOS SF-36 scale score for Role Physical and the 10 m ISWT distance (Table 5.15). The results for the pre-program did not change when accounting for gender.

Table 5.15

Association between the 10 m ISWT and MOS SF-36 Scale Score for Role Physical in the Pre-Program Assessment

		95% CI	
10 m ISWT	*r*	LL	UL
ISWT 1	.236	-.039	.478
ISWT 2	.245	-.030	.485

Note. *r* = Pearson's product-moment correlation coefficient; CI = confidence interval; LL = lower limit; UL = upper limit.

For the post-program assessment, there was a significant weak correlation between the MOS SF-36 Role Physical scale score and the 10 m ISWT distance (Table 5.16). These results varied for gender, with a significant correlation for men (ISWT 3 r = .404, 95% CI [.106, .636] and ISWT 4 r = .414, 95% CI [.118, .643]), but not for women (ISWT 3 r = .253, 95% CI [-.375, .722], and ISWT 4 r = .260 95% CI [-.369, .726]).

Table 5.16

Association between the 10 m ISWT and MOS SF-36 Scale Score for Role Physical in the Post-Program Assessment

		95% CI		
10 m ISWT	*r*	*LL*	*UL*	r^2
ISWT 3	.366	.103	.581	.134
ISWT 4	.374	.113	.587)	.140

Note. *r* = Pearson's product-moment correlation coefficient; CI = confidence interval; LL = lower limit; UL = upper limit; r^2 = coefficient of determination.

The results partially supported hypothesis 5.4, while the correlation was weak and less than that demonstrated in hypothesis 5.3, it was also non-significant for the pre-program assessment outcomes. However, for the post-program assessment outcomes, the correlation was significant, weak, and less than that demonstrated in hypothesis 5.3.

5.4.2.5 Hypothesis 5.5: there will be a non-significant correlation between the scores of the remaining six MOS SF-36 scales and the 10 m ISWT, with the MOS SF-36 scale score for Role Emotional the least related to the distance walked in the 10 m ISWT.

There were weak and non-significant correlations between the remaining six scales of the MOS SF-36 and the 10 m ISWT distance at the pre-program assessment (Table 5.17). The correlation between the MOS SF-36 scale score for Role Emotional and the distance walked in the 10 m ISWT result was close to zero for both the first and second walk at the pre-program assessment, and the least related of the MOS SF-36 scales. The results were not affected by gender.

For the post-program assessment, results were non-significant for both ISWT 3 and ISWT 4 for five of the MOS SF-36 scales. There was a significant weak correlation between the MOS SF-36 scale score for Social Functioning and ISWT 4 (Table 5.17). This hypothesis was partially supported by the results.

Table 5.17

Associations between the 10 m ISWT and MOS SF-36 Scale Scores

	Pre-program outcomes						Post-program outcomes					
	ISWT 1			ISWT 2			ISWT 3			ISWT 4		
		95% CI			95% CI			95% CI			95% CI	
MOS SF-36 scale	*r*	LL	UL	*r*	LL	UL	*r*	LL	UL	*r*	LL	UL
Bodily Pain	.113	-.165	.374	.122	-.156	.382	.061	-0.215	.328	.081	-.196	.346
General Health	.224	-.052	.468	.249	-.026	.489	.230	-.046	.473	.253	-.021	.492
Vitality	.155	-.123	.411	.175	-.103	.427	.213	-.064	.459	.235	-.041	.477
Social Functioning	.060	-.216	.328	.078	-.199	.344	.254	-.020	.493	.278	.006	.512
Role emotional	-.048	-.317	.228	-.033	-.303	.242	.103	-.175	.366	.112	-.166	.373
Mental Health	.088	-.189	.352	.110	-.252	.445	.117	-.161	.378	.139	-.139	.397

Note. MOS SF-36 = Medical Outcomes Study 36-item short form; *r* = Pearson's product-moment correlation coefficient; CI = confidence interval; LL = lower limit; UL = upper limit.

5.4.2.6 Hypothesis 5.6: there will be a weak correlation (r < .50) between the exercise self-efficacy score and the distance walked in the 10 m ISWT.

There was a non-significant weak positive correlation between the exercise self-efficacy score and the distance walked in the 10 m ISWT in the pre-program assessment (Table 5.18). However, there was a significant weak correlation between the exercise self-efficacy score and the distance walked in the 10 m ISWT in the post-program assessment (Table 5.19).

Table 5.18

Associations between the 10 m ISWT and Exercise Self-Efficacy in the Pre-Program Assessment

		95% CI	
10 m ISWT	*r*	LL	UL
ISWT 1	.200	-.077	.448
ISWT 2	.239	-.036	.481

Note. *r* = Pearson's product-moment correlation coefficient; CI = confidence interval; LL = lower limit; UL = upper limit.

Table 5.19

Associations between the 10 m ISWT and Exercise Self-Efficacy in the Post-Program Assessment

		95% CI		
10 m ISWT	*r*	*LL*	*UL*	r^2
ISWT 3	.281	.009	.514	.079
ISWT 4	.297	.026	.527	.088

Note. *r* = Pearson's product-moment correlation coefficient; CI = confidence interval; LL = lower limit; UL = upper limit; r^2 = coefficient of determination.

There was a difference in the correlation for gender at both the pre-program and the post-program assessment. At the pre-program assessment, the mean score for men was 104 (*SD* 34) and for women 91 (*SD* 36), and at the post-program assessment, the mean score for men was 118 (*SD* 34) and for women 108 (*SD* 29). The differences in self-efficacy scores between men and women was not statistically significant for both the pre-program assessment ($t(50) = 1.090$, $p = .218$) or the post-program assessment ($t(50) = .965$, $p = .339$). For both the pre-program and the post-program assessment, the correlation was not significant for men. For women, the correlation was significant and strong for all cases in the pre-program assessment (exercise self-efficacy and ISWT 1 $r = .876$, 95% CI [.608, .965], and ISWT 2, $r = .875$, 95% CI [.605, .965]) and the post-program assessment (exercise self-efficacy and ISWT 3 $r = .764$, 95% CI [.339, .930] and exercise self-efficacy and ISWT 4, $r = .719$, 95% CI [.247, .915]) (Figure 5.2).

This hypothesis was partially supported. There was support for the post-program assessment but not for the pre-program assessment, and for the women in the sample, the association was stronger than hypothesised.

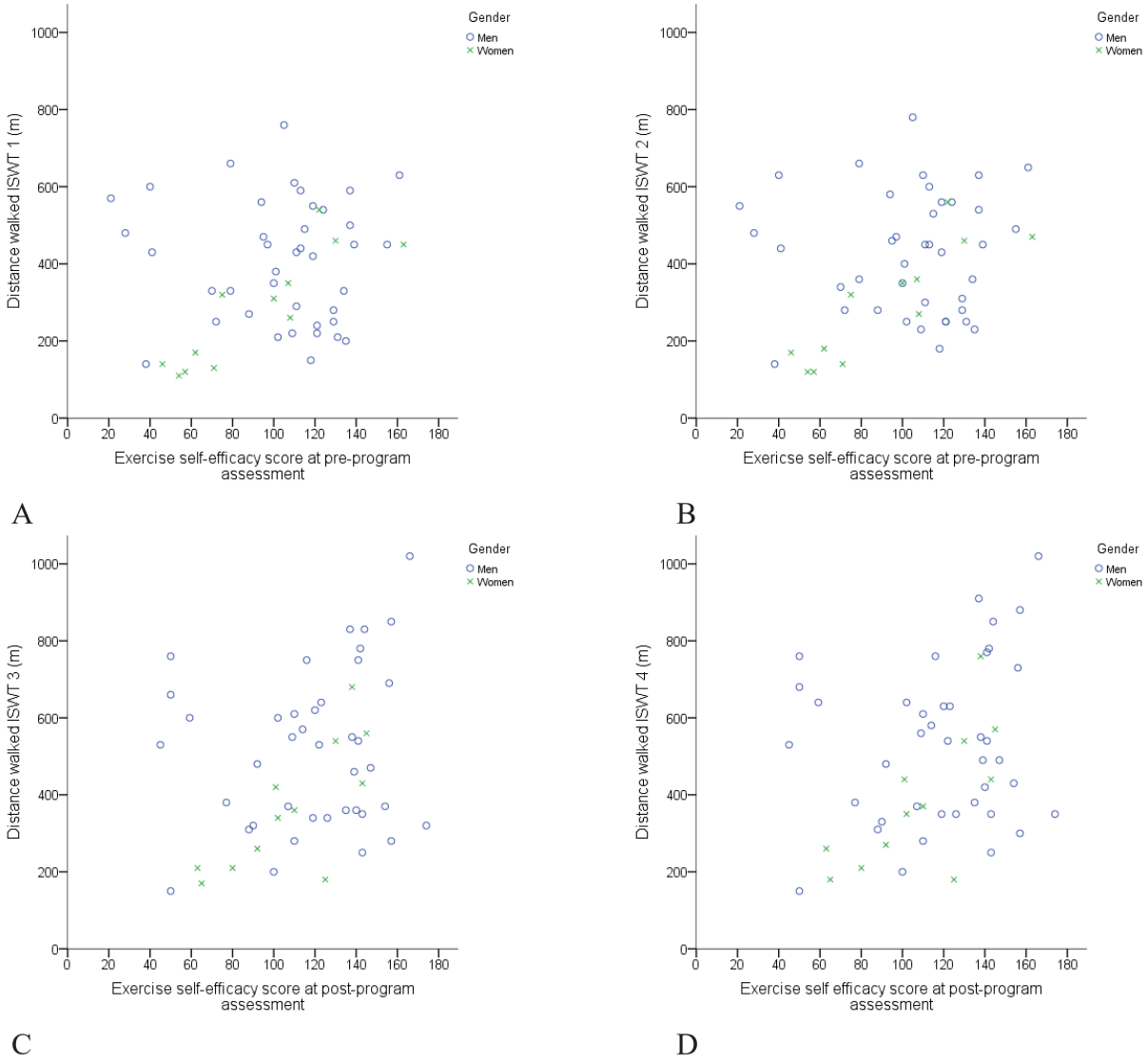


Figure 5.2. Scatter plot showing the spread of scores for the exercise self-efficacy score and the distance walked in the 10 m ISWT at the pre-program assessment and post-program assessment.

5.4.2.7 Hypothesis 5.7: the difference in the distance walked in the 10 m ISWT between participants with likely depression compared with those with unlikely depression at baseline will not be statistically significant (p > .05), however, at follow-up the difference will be statistically significant (p < .05).

At the pre-program assessment, 10 (19%) participants were likely depressed according to the Brief Case Find for Depression. The mean difference between those with likely depression and those without in the first 10 m ISWT was 86 m, 95% CI [-29, 201] and for the second 10 m ISWT was 96 m, 95% CI [-17, 210]. The distance walked in the 10 m ISWT was less in participants with probable depression than those without probable depression, although the differences were not statistically significant ($p \geq .05$) (Table 5.20).

At the post-program assessment, eight (15%) participants were likely depressed. These eight participants also scored likely depression at the pre-program assessment. The mean difference between those with likely depression and those without was 149 m, 95% *CI* [-6, 304] and 158 m, 95% *CI* [0, 317) for ISWT 3 and ISWT 4, respectively. The distance walked in the 10 m ISWT was less in participants with probable depression than those without probable depression, although the differences were not statistically significant ($p \geq .05$) (Table 5.21).

This hypothesis was partially supported.

Table 5.20

The 10 m ISWT distance in the Pre-Program Assessment based on Brief Case Find for Depression Result

	10 m ISWT distance (m)						
	Likely depression	Unlikely depression		95% CI			
Test	m (*SD*)	m (*SD*)	M_{diff}	LL	UL	*t*(50)	*p*
ISWT 1	309 (153)	395 (164)	86	-29	201	1.506	.138
ISWT 2	318 (146)	414 (164)	96	-17	210	1.701	.095

Note. CI = confidence interval; LL = lower limit; UL = upper limit.

Table 5.21

The 10 m ISWT distance in the Post-Program Assessment based on Brief Case Find for Depression Result

	10 m ISWT distance (m)						
	Likely depression	Unlikely depression		95% CI			
Test	M (*SD*)	M (*SD*)	M_{diff}	LL	UL	*t*(50)	*p*
ISWT 3	355 (168)	504 (206)	149	-6	304	1.926	.060
ISWT 4	363 (171)	521 (210)	158	0	317	2.009	.050

Note. CI = confidence interval; LL = lower limit; UL = upper limit.

5.4.3 The longitudinal construct validity of the 10 m ISWT in a mixed cardiac rehabilitation program.

5.4.3.1 Hypothesis 5.8: those who participated in an eight-week outpatient cardiac rehabilitation exercise program will demonstrate an improvement in the 10 m ISWT between their pre-program and post-program assessment, and that there will be a moderate effect size, and the differences and effect size will be larger in patients with unlikely depression compared with those with likely depression.

The results for the pre-program and post-program walk test differences and the effect sizes are shown in Table 5.22. There was a significant difference in the pre-program and post-program distances walked in the 10 m ISWT for all combinations of pre-program and post-program walk tests. A moderate effect size was noted in the pre-program and post-program comparisons. The *p*-values were smaller and the effect sizes were larger for patients without likely depression than those with likely depression for all combinations of pre-program and post-program walk tests (Table 5.23). This hypothesis was supported.

Table 5.22

Mean Difference and Effect Size for Change in the 10 m ISWT Distance over an Eight-Week Comprehensive Cardiac Rehabilitation Program

Tests	M_{diff} (SD_{diff}) (m)	t(50)	p	Effect size
ISWT 1 and 3	103 (72)	10.224	< .001	0.62
ISWT 1 and 4	118 (79)	10.831	< .001	0.72
ISWT 2 and 3	85 (70)	8.731	< .001	0.52
ISWT 2 and 4	101 (76)	9.597	< .001	0.61

Table 5.23

Mean Difference and Effect Size for Change in the 10 m ISWT Distance over an Eight-Week Comprehensive Cardiac Rehabilitation Program based on Brief Case Find for Depression Results

	Likely depression				Unlikely depression			
Tests	M_{diff} (SD_{diff}) (m)	t(7)	p	Effect size	M_{diff} (SD_{diff}) (m)	t(43)	p	Effect size
ISWT 1 and 3	63 (41)	4.394	.003	.38	110 (75)	9.714	.001	.68
ISWT 1 and 4	71 (43)	4.734	.002	.43	127 (81)	10.373	.001	.79
ISWT 2 and 3	56 (49)	3.230	.014	.36	90 (73)	8.247	.001	.57
ISWT 2 and 4	64 (52)	3.444	.011	.41	108 (78)	9.166	.001	.67

5.4.3.2 Hypothesis 5.9: those unlikely to have depression will make greater improvements, across a cardiac rehabilitation program, in the 10 m ISWT distance than those with likely depression and this difference will be statistically significant.

The mean improvement in distance walked on the 10 m ISWT between the pre-program and the post-program assessment was greater for people with unlikely depression compared with likely depression. These results did not change when the first or second test of the session was used. The differences in improvement scores were not statistically significant in three of four walk combinations and only statistically significant for the differences in improvement between ISWT 1 and ISWT 4, in other words, when one test is completed in the pre-program assessment and the results of the second test are used in the post-program assessment (Table 5.24). This hypothesis was only partially supported.

Table 5.24

Mean Differences in Change in 10 m ISWT Distance based on Brief Case Find for Depression Result in the Pre-Program Assessment

	Change in 10 m ISWT distance (m)						
	Unlikely depression	Likely depression	M_{diff}	95% CI			
Tests	M (SD_{diff})	M (SD_{diff})	(m)	LL	UL	t(50)	p
Walk 1-3	111 (76)	66 (38)	45	-5	95	1.816	.075
Walk 1-4	129 (82)	73 (40)	56	2	110	2.082	.042
Walk 2-3	92 (74)	57 (44)	35	-14	84	1.424	.161
Walk 2-4	110 (79)	64 (48)	46	-7	98	1.743	.088

Note. CI = confidence intervals; LL = lower limit; UL = upper limit.

5.4.3.3 Hypothesis 5.10: there will be significant weak correlations (r < .50) between changes in the MOS SF-36 scale scores that relate to physical health and changes in the distance walked in the 10 m ISWT, and non-significant correlations (p > .05) between changes in the remaining MOS SF-36 scale scores and changes in the 10 m ISWT.

For the sample, there was a non-significant correlation between change in the MOS SF-36 scale score Physical Functioning and the change in 10 m ISWT across a cardiac rehabilitation program (Table 5.25). There was a weak correlation between the change in MOS SF-36 scale score Role Physical and change in 10 m ISWT distance, which was significant for three of the four walk test combinations (Table 5.25). The associations between the change scores for all other MOS SF-36 scales and the change in 10 m ISWT distance were weak and non-significant (Table 5.25). There were no differences in results based on gender or likelihood of depression. The results of this hypothesis were partially supported. The results did not support the hypothesis that there would be a significant weak correlation between changes in the MOS SF-36 scale score for Physical Functioning and the change in the 10 m ISWT distance. However, the results showed partial support for a significant weak correlation between changes in the MOS SF-36 scale score for Role Physical and changes in the 10 m ISWT distance and support for the non-significant correlation between the changes in the remaining MOS SF-36 scales scores and changes in the 10 m ISWT distance.

Table 5.25

Associations between the Change in 10 m ISWT Distance Walked and the Change in MOS SF-36 Scale Scores

	ISWT 1 and 3			ISWT 1 and 4			ISWT 2 and 3			ISWT 2 and 4		
		95% CI			95% CI			95% CI			95% CI	
MOS SF-36 Scale	*r*	LL	UL	*r*	LL	UL	*r*	LL	UL	*r*	LL	UL
Physical functioning	.078	-.199	.344	.075	-.202	.341	.021	-.253	.292	.023	-.251	.294
Role Physical	.272	.001	.507	.290	.019	.522	.250	-.025	.490	.274	.001	.509
Bodily Pain	.076	-.201	.342	.100	-.178	.363	.050	-.226	.319	.079	-.198	.344
General Health	.060	-.216	.328	.107	-.171	.369	.095	-.183	.359	.141	-.137	.399
Vitality	.216	-.060	.462	.214	-.063	.460	.225	-.051	.469	.225	-.051	.469
Social Functioning	.242	-.033	.483	.218	-.058	.463	.234	-.042	.476	.212	-.065	.458
Role emotional	-.022	-.293	.252	-.039	-.309	.236	-.036	-.306	.239	-.052	-.320	.224
Mental Health	.007	-.266	.279	-.025	-.296	.250	.001	-.272	.274	-.032	-.302	.243

Note. MOS SF-36 = Medical Outcomes Study 36-item short form; *r* = Pearson's product-moment correlation coefficient; CI = confidence interval; LL = lower limit; UL = upper limit.

5.4.3.4 Hypothesis 5.11: there will be a significant weak positive association (r < .50) between attendance at a cardiac rehabilitation program and change in 10 m ISWT distance.

There was a significant weak positive correlation between the number of exercise classes attended over an eight-week comprehensive cardiac rehabilitation program and the change in the distance walked in the 10 m ISWT (Table 5.26). There were no differences in attendance rates based on gender or likelihood of depression. This hypothesis was supported.

Table 5.26

Association between the Change in 10 m ISWT Distance Walked and the Number of Cardiac Rehabilitation Sessions Attended

		95% CI		
Tests	*r*	LL	UL	r^2
ISWT 1 and 3	.472	.229	.660	.223
ISWT 1 and 4	.475	.232	.662	.226
ISWT 2 and 3	.469	.225	.658	.220
ISWT 2 and 4	.478	.236	.664	.228

Note. *r* = Pearson's product-moment correlation coefficient; CI = confidence intervals; LL = lower limit; UL = upper limit; r^2 = coefficient of determination.

5.4.3.5 Hypothesis 5.12: participants who report improvement in Global Rating of Change will demonstrate a greater improvement in the 10 m ISWT distance over an eight-week outpatient cardiac rehabilitation program than those who report no change or deterioration in Global Rating of Change. This difference will be statistically significant. Those who report no change or a deterioration in Global Rating of Change will not show a statistically significant difference in the change in 10 m ISWT distance.

There was a significant difference between the 10 m ISWT change scores and global rating of change categories for all walk test combinations (Table 5.27). *Post-hoc* comparisons showed that there was a statistically significant difference in the 10 m ISWT change score in those who reported an improvement in global rating of change compared with those who reported to be either unchanged ($p < .001$) or deteriorated ($p < .001$) for all walk combinations. *Post-hoc* comparisons showed no significant difference in the distance walked between those participants reporting no change and those who reported deterioration in their global rating of change. This hypothesis was supported.

Table 5.27

Change in 10 m ISWT Distance based on Global Rating of Change

	Change in 10 m ISWT (m) *M* (SD_{diff})				
Tests	Improvement	No Change	Deterioration	*F*(2,49)	*p*
Walk 1 and 3	143 (57)	57 (21)	17 (40)	30.318	.001
Walk 1 and 4	165 (69)	66 (22)	21 (34)	37.615	.001
Walk 2 and 3	125 (56)	40 (19)	4 (40)	28.721	.001
Walk 2 and 4	146 (57)	49 (20)	8 (35)	38.212	.001

Table 5.28 provides a summary of results of the hypothesis testing for the construct validity of the baseline scores and the longitudinal validity of the change scores.

5.5 Discussion

This study evaluated the evidence for the construct validity and responsiveness of the 10 m ISWT as a measure of physical fitness and functional capacity in a mixed outpatient cardiac rehabilitation population. The COSMIN framework was used. Overall, 92% of the *a-priori* hypotheses were at least partially supported. One (8%) hypothesis was rejected. The results provide evidence of construct validity and responsiveness to support the use of the 10 m ISWT as part of a pre-program assessment for a comprehensive cardiac rehabilitation program and as a test responsive to changes in physical fitness and functional capacity over a mixed outpatient cardiac rehabilitation program.

Table 5.28

Summary of the Results of Hypothesis Testing

	Hypothesis	Accepted
5.1	Participants who report regular physical exercise at a moderate-vigorous intensity will walk significantly further in the 10 m ISWT compared with those who report exercise at a low intensity who in turn will walk significantly further than those who report to be sedentary.	Yes
5.2	There will be a significant weak negative correlation between measures of adiposity and the 10 m ISWT distance walked.	No
5.3	There will be a significant weak-moderate positive correlation between the MOS SF-36 scale score for Physical Functioning and the distance walked in the 10 m ISWT.	Partial
5.4	There will be a weak positive correlation between the MOS SF-36 scale score for Role Physical and the distance walked in the 10 m ISWT and this correlation will be less than that demonstrated in hypothesis 5.3.	Partial
5.5	There will be a non-significant correlation between the scores of the remaining six MOS SF-36 scale scores and the 10 m ISWT, with the SF-36 scale score for Role Emotional the least related to the distance walked in the 10 m ISWT.	Yes
5.6	There will be a weak correlation between the exercise self-efficacy score and the distance walked in the 10 m ISWT.	Partial
5.7	The distance walked in the 10 m ISWT will be less in participants with likely depression compared with those with unlikely depression and that this difference in distance will not be significant at baseline, however, will be significant at follow-up after cardiac rehabilitation.	Partial
5.8	Those who participate in an eight-week outpatient cardiac rehabilitation exercise program will demonstrate an improvement in the 10 m ISWT between their pre-program and post-program assessment, and that there will be a moderate effect size, the differences and the effect size will be larger in patients without likely depression than those with likely depression.	Yes

	Hypothesis	Accepted
5.9	Those with unlikely depression will make greater improvements across a cardiac rehabilitation program in the 10 m ISWT distance than those with likely depression, this difference will be statistically significant, and the effect size will be larger in patients without likely depression than those with likely depression.	Partial
5.10	There will be significant weak correlations between changes in the MOS SF-36 scale scores that relate to physical health and changes in the 10 m ISWT distance, and non-significant correlations between changes in the remaining MOS SF-36 scale scores and changes in the 10 m ISWT.	Partial
5.11	There will be a significant weak positive correlation between attendance at a cardiac rehabilitation program and improvement in 10 m ISWT results.	Yes
5.12	Participants who report improvement in Global Rating of Change will demonstrate a greater improvement in the 10 m ISWT distance over an eight-week outpatient cardiac rehabilitation program than those who report no change or deterioration in a Global Rating of Change. This difference will be statistically significant. Those who report no change or a deterioration in Global Rating of Change will not show a statistically significant difference in the change in 10 m ISWT walk distance.	Yes

5.5.1 Evidence for construct validity of the 10 m ISWT.

This chapter provides further support that the distance walked in the 10 m ISWT provides information on physical fitness and functional capacity of patients attending a mixed outpatient cardiac rehabilitation program. In this current study, seven hypotheses relating to construct validity were tested and the results demonstrated convergent themes between the 10 m ISWT and commonly used outcomes relating to physical fitness or functional capacity in a mixed cardiac rehabilitation group, as well as divergent themes between the 10 m ISWT and commonly used outcome measures that do not specifically relate to physical fitness or functional capacity. There were two significant positive associations between the distance walked in the 10 m ISWT and commonly used outcome measures, specifically the self-reporting of the amount of weekly exercise and physical functioning and the reporting of functional activity through the MOS SF-36 Physical Function scale score. Those patients who reportedly met national exercise guidelines, and exercised regularly at a moderate intensity, walked further in the 10 m ISWT than those patients who reported exercise at an insufficient intensity, frequency or duration to meet the national exercise guidelines, or those who were sedentary. The hypothesis that those who participate in some form of exercise despite being less than the national guidelines for intensity, frequency or duration will perform better in the 10 m ISWT than those who were sedentary was also supported. The accuracy of participant reporting is unknown, and it is possible that participants in this study over-reported their participation in physical activity (Beyler, Nusser, Fuller, & Gregory, 2008). Nonetheless, these findings support previous reports in the literature that even low-intensity exercise can be beneficial in improving physical fitness and functional capacity in patients with cardiac disease

(Wright et al., 2001), and presents new evidence that the 10 m ISWT is able to differentiate between these groups of people.

Functional activity was measured by the MOS SF-36 scale Physical Functioning and it was found to correlate moderately with the 10 m ISWT. Five of the 10 questions in MOS SF-36 scale Physical Functioning specifically relate to how a participant's health limits him or her in vigorous activities such as running, lifting heavy objects, participating in strenuous sport; and moderate activities, such as moving a table, pushing a vacuum cleaner, bowling or playing golf, lifting or carrying groceries, climbing several flights of stairs and walking more than a kilometre. These results are consistent with earlier research presenting a significant moderate association between the distance walked in the 6MWT and the MOS SF-36 Physical Functioning scale in cardiac rehabilitation (Hamilton & Haennel, 2000). There appears to be a stronger association when using clinically-available outcome measures such as distance walked in functional walk tests, such as the 6MWT or the 10 m ISWT, compared with less functional testing such as incremental treadmill tests that include both incline and speed increases (Brubaker et al., 2003; Jette & Downing, 1996) or other physiological outcome measures, such as peak oxygen uptake (Jette & Downing, 1996).

The hypothesis that there would be no significant difference in the distance walked in the 10 m ISWT for those who presented to a mixed outpatient cardiac rehabilitation program with likely depression compared with those with unlikely depression was observed. However, the hypothesis that there would be a significant difference in the distance walked between the two groups at completion of cardiac rehabilitation was not observed. The likelihood of depression or depressive symptoms did not result in a statistically significant change in the 10 m ISWT during the pre-program assessment

and the post-program assessment, although there appeared to be a non-significant trend that those with likely depression walked less on the 10 m ISWT compared with those with unlikely depression. It is possible that the non-significant findings are a reflection that the length of time the participants experienced depressive symptoms was insufficient to affect physical fitness and functional activity, or it is possible that the findings were a result of type II errors and a significant difference may have been detected in a larger sample size. There was a greater difference in scores in the post-program assessment and this difference approached significance for both ISWT 3 and ISWT 4 and may reflect the additional time that the participants had experienced depressive related symptoms. In addition, the BCD has been reported to have poor specificity (Jefford et al., 2004); meaning that patients may have been incorrectly identified as having likely depression and, it is not known how many of the 10 patients were incorrectly identified. Additionally, it is not known how many of these patients were experiencing major depressive symptoms compared with minor depressive symptoms, and the impact of major depression on the walk distance score compared with minor depressive-related symptoms. It is not known whether using a different tool to detect depression in patients attending cardiac rehabilitation would have resulted in different findings. The uneven group sizes may have resulted in an underestimation of the true differences. These results suggest that at baseline there were no statistically significant differences in walk distance for those with likely depression, compared with those without likely depression identified using the BCD and further investigation is needed.

There were three unexpected findings in the study. First, that the results did not support an inverse association between the 10 m ISWT distance walked and measures

of adiposity; second, that there was a non-significant association between the pre-program MOS SF-36 Role Physical scale score and the distance walked in the 10 m ISWT, and third, the gender differences from the association between the exercise self-efficacy score and the distance walked in the 10 m ISWT. For measures of adiposity, the results demonstrated a weak positive association between the waist-hip ratio and the distance walked in the 10 m ISWT. Further analysis showed this result was potentially misleading, and that when men and women were analysed independently the results showed a lack of association. The positive association seen in the sample is likely to be a result of the confounding effect of gender groups: the men in this sample were more likely to walk significantly further in the 10 m ISWT than women; and men are more likely to have upper-body obesity and thus higher waist:hip ratios than women who typically have lower-body obesity (Vague, 1956). The absence of the association between the 10 m ISWT and BMI when taking account of gender may be a result of the homogeneous group, with 81% of participants being overweight or obese.

The association between the pre-program MOS SF-36 scale score for Role Physical and the distance walked in the 10 m ISWT was not supported. It is possible that the MOS SF-36 scale score for Role Physical was more strongly affected by other factors such as individual reservations about returning to usual activities, professional advice given on movement restrictions, or return to activities. These findings support the findings of earlier research that found limitations in the completion of heavy housework was due to reservations around the diagnosis of heart disease rather than exercise capacity (Neill et al., 1985). In comparison, there was a positive weak association between MOS SF-36 scale score for Role Physical and the 10 m ISWT

distance at the post-program assessment, suggesting that people who experience fewer limitations with their work or other daily activities due to their physical health also perform better on the 10 m ISWT. It is possible that over the duration of cardiac rehabilitation the movement and other restrictions, as well some of the fear and avoidance behaviours, were lessened in the higher functioning group, and that this was reflected in the weak correlation between the two outcomes.

The study reported a weak association between the exercise self-efficacy scores and the distance walked in the 10 m ISWT at the post-program assessment, but not with the pre-program scores. Unexpectedly, strong correlations were observed during both the pre-program and the post-program assessment for the women in the study. It is possible that the women in this study were more accurately able to determine how different factors affected their ability to exercise on most days of the week. This finding remains unexplained and requires further investigation.

The evidence for the construct validity was supported through *a-priori* hypothesis testing. This chapter also demonstrated that there was support for the construct validity when one 10 m ISWT was performed and provided further evidence that one 10 m ISWT is sufficient in this population.

5.5.2 Evidence for responsiveness of the 10 m ISWT.

The second aim of this chapter was to assess the responsiveness or longitudinal construct validity of the 10 m ISWT as a measure of physical fitness and functional capacity in a mixed cardiac rehabilitation group. The results support the usefulness of the 10 m ISWT as a measure of change in physical fitness and functional capacity

over an eight-week mixed outpatient cardiac rehabilitation program. The *a-priori* hypothesis testing yielded support for both the internal and the external responsiveness of the 10 m ISWT in this population and context.

The internal responsiveness was supported by the statistically significant change in walk test scores as well as the moderate effect size across the outpatient program. In this study, average walk distance improved by just over 100 m or 10 shuttles. These findings are consistent with the range of distances reported in earlier research. The improvements seen in this study were greater than the 82 m and moderate effect size reported by Fowler et al. (Fowler et al., 2005) after a six-week program, and less than the improvements reported following 12-week programs (Frizelle et al., 2004; Tobin & Throw, 1999). Additional support for the 10 m ISWT as a responsive tool in this context was the positive rating for floor and ceiling effects. All participants completed at least one shuttle and during the pre-program assessment, no participant completed all levels of the test. The results support the 10 m ISWT as a suitable tool to detect change when it does occur in patients with cardiac disease who attend an eight-week program.

The hypothesis that the presence of depression or depressive symptoms would affect the change in 10 m ISWT results was mixed, but in general showed support that those with unlikely depression performed better in a cardiac rehabilitation program than those with likely depression. As expected the *p*-values were smaller and the effect sizes greater for the group of patients with unlikely depression compared with the group of patients with likely depression. The change in the distance walked between the first pre-program and the first post-program 10 m ISWT, as well as the second pre-program and the second post-program 10 m ISWT, were not statistically

significant but the *p*-values both approached .05. These results may have been affected by the small sample size of the likely depressed group. These results support previous reports in the literature that patients with depression who attend cardiac rehabilitation and do not receive specific intervention for their depression do not achieve the same outcomes as those patients without depression (Glazer et al., 2002).

The results of this study support the external responsiveness of the 10 m ISWT in a mixed cardiac rehabilitation group. The hypothesis, that there would be a statistically significant difference in the change in 10 m ISWT distance for the group who rated themselves as improved compared with the groups who rated themselves as unchanged or deteriorated, was supported. This is the second study that has presented information on the external responsiveness of the 10 m ISWT in a mixed cardiac rehabilitation program and the results of this study support the results of the earlier study (Houchen-Wolloff et al., 2015).

In addition, this is the first study to report the associations between attendance at cardiac rehabilitation and change in 10 m ISWT distance. While attendance is a crude measure of adherence, it can be considered a reflection of self-regulatory self-efficacy (Woodgate et al., 2005). In this study, there was a weak but significant association between attendance and change in 10 m ISWT distance walked. Further research is required to understand this association.

One unexpected result was finding no association between the change in MOS SF-36 Physical Function scale score and the change in 10 m ISWT. No explanation for this finding can be offered, and further research is needed to understand the relationship

between changes in health-related quality of life, in particular self-reporting of physical function, and the change in the 10 m ISWT.

The responsiveness or longitudinal construct validity was supported by the results of various other *a-priori* hypothesis testing comparing the change scores of the distance walked in the 10 m ISWT with scores from other common assessment scores. Furthermore, the results of this chapter provide support that the level of systematic error in the 10 m ISWT does not affect the validity when a single test is performed. In general, there were no differences in the results of the responsiveness when one test or a second test was performed during an assessment. There appeared to be variation in results if a different number of tests were performed, such as one test during the pre-assessment and two tests during the post-assessment. The results reinforce the findings of Chapter 4 that a single 10 m ISWT is sufficient in this population.

5.5.3 Limitations and future directions.

This study was limited to measuring responsiveness through a longitudinal construct validity design. An alternative research design would be to measure responsiveness with a criterion approach (de Vet et al., 2011, p. 206). However, the use of laboratory exercise testing in pre-program and post-program cardiac rehabilitation in this region is extremely rare and recruitment would not have been practical. Additionally, it would be useful to test further a number of the hypotheses relating to gender, depression and exercise self-efficacy with a larger sample size: Terwee et al. (2007) recommends 50 participants in each group for analysis. While the self-reported measures of home exercise provided a useful analysis in this study, it would be useful to compare the results of the 10 m ISWT with results of more formal measures of

physical activity, including exercise, such as with the use of accelerometers or standardised questionnaires.

5.5.4 Implications.

This study adds to the current knowledge of the 10 m ISWT in a mixed cardiac rehabilitation population. It followed the recommendations of the COSMIN group, and through hypothesis testing demonstrated support that the 10 m ISWT measures what it purports to measure and is responsive in a mixed outpatient cardiac rehabilitation program. It provides new knowledge of construct validity of the 10 m ISWT, such as the relationship between functional outcomes of the 10 m ISWT with reported home exercise activity, attendance at cardiac rehabilitation and other cardiac rehabilitation assessment tools. The results suggest that the 10 m ISWT is an appropriate tool to measure responsiveness in a program and can be used to measure program effectiveness, and when functional outcome measures are used, the response can be seen after an eight-week program.

Chapter 6: The Interpretation of Change in the 10 m ISWT Distance in Cardiac Rehabilitation

6.1 Chapter Aims

This study addressed secondary research aim 9, specifically to assess the interpretability of the 10 m ISWT when used in cardiac rehabilitation. The framework and definitions provided by the COSMIN group (Mokkink et al., 2010c) were followed to estimate the SDC, and the MIC in the 10 m ISWT over an eight-week comprehensive cardiac rehabilitation program. In addition, a description of the presence of any significant floor and ceiling effects.

6.2 Introduction

Meaningful change has emerged as a key concept in interpretability of measurement, and has been defined by the COSMIN group as the degree to which qualitative meaning can be applied to test scores (Mokkink et al., 2010a, 2010c). Meaningful change provides information beyond the statistical significance of the change score. Two key concepts are the SDC, which is the smallest amount of change that can be detected over and above measurement error (de Vet et al., 2011, p. 258; Terwee et al., 2007), and the MIC. The MIC is the smallest amount of change that is meaningful or perceived as important and is typically from the perspective of the patient, but it may be from the perspective of the clinician or an institution.

The SDC for the 10 m ISWT distance walked has been reported in Chapter 3 and in previous literature as the 95% confidence intervals for the individual or the group. In Chapter 3 the SDC for the group (n = 62) was 23 m and the individual was 54 m after a single walk. Previous research in different cardiac rehabilitation populations have reported the SDC for the group to range from 26 to 56 m (Fowler et al., 2005; Jolly et al., 2008; Pepera, McAllister, & Sandercock, 2010) and for the individual to range from 53 to 203 m (Fowler et al., 2005; Pepera et al., 2010; Pulz et al., 2008).

One study has reported the MIC in distance walked in the 10 m ISWT following a six-week outpatient cardiac rehabilitation exercise intervention (Houchen-Wolloff et al., 2015). The MIC was calculated two ways and the results varied. The MIC from the anchor-based method that used a global rating of change score as the anchor was 70 m. In comparison, the MIC from the distribution-based method, that calculated the standard deviation of the change score and effect size, was 37 m with a small effect size. The authors concluded that the larger MIC, calculated using an anchor-based approach, was a result of patient expectations of improvement in exercise ability over a cardiac rehabilitation course (Houchen-Wolloff et al., 2015). These results suggest that patients who attend cardiac rehabilitation need to improve by more than three shuttles and possibly up to seven on the 10 m ISWT to rate themselves as improved. Further research is needed to investigate the MIC in this population.

Previous research has reported the statistical significance of the change in the 10 m ISWT distance after a cardiac rehabilitation intervention (Arnold, Sewell, & Singh, 2007; Fowler et al., 2005; Robinson et al., 2011). In Chapter 5, a mean difference of 103 m, 95% CI [82, 123] was reported after an eight-week comprehensive cardiac rehabilitation program. Arnold et al. (2007) found improvements of 101 m, 95% CI

[81, 121] and 88 m, 95% CI [20, 107] in patients attending once-weekly and twice-weekly six-week cardiac rehabilitation programs. Robinson et al. (2011) reported an improvement of approximately 60 m following either a traditional or a fast-tracked six-week cardiac rehabilitation program and this improvement was statistically significant. The researchers also reported that those participants who went on to complete a maintenance program and were still exercising at six months, demonstrated a statistically significant mean improvement of 40 m, (p = .03) 95% CI [4, 77]. Fowler et al. (2005) reported an improvement of 82 m, 95% CI [53, 110] following a six-week cardiac rehabilitation exercise intervention. Another study demonstrated a non-significant mean improvement of 7 m, 95% CI [-29, 44] over an eight-week period where participants were involved in an ongoing program designed to maintain functional capacity (Pepera et al., 2010). While the changes in 10 m ISWT walk distance after a cardiac rehabilitation program designed to improve physical fitness and functional capacity are often statistically significant, the meaningfulness or interpretability of the change scores, that is whether the changes exceed the SDC or MIC, remains unknown.

The purpose of this study was to contribute to the limited research on the interpretability of change scores for the 10 m ISWT in a cardiac rehabilitation population. Specifically, this research aimed to determine the SDC, as well as the MIC, using a variety of methods in a prospective cohort of patients referred to a mixed outpatient cardiac rehabilitation program.

6.3 Method

6.3.1 Ethics approval and consent.

Ethics approval was granted by La Trobe University, University Human Ethics Committee (reference number 10-082), and St John of God Health Care Human Research Ethics Committee (reference number 438) (see Appendix 2). All participants enrolled in the investigation read and signed a Participant Information and Consent Form that was approved by the relevant ethics committees (see Appendix 3).

6.3.2 Participants.

Participants were recruited on admission to a cardiac rehabilitation program. Eligibility criteria for participants have been described in Chapter 3, Section 3.3.2. The recruitment of participants for this study was described in Chapter 5, Section 5.3.2.2.

6.3.3 Procedure.

The procedure for this study was described in detail in Chapter 5, Section 5.3.3.

6.3.3.1 Pre-program assessment.

The pre-program assessment was described in Chapter 5, Section 5.3.3.1. Participants completed an initial subject and physical assessment to record baseline details and then two 10 m ISWTs in a single session (ISWT 1 and ISWT 2).

6.3.3.2 Intervention.

Participants completed an eight-week outpatient comprehensive cardiac rehabilitation program as described in Chapter 5, Section 5.3.3.2.

6.3.3.3 Post-program assessment.

In a single session, participants completed a subjective and physical assessment to record general information, a specific assessment of a global rating of change score, and two 10 m ISWTs in a single session (ISWT 3 and ISWT 4). The post-program assessment was described in detail in Chapter 5, Section 5.3.3.3.

6.3.3.4 Outcome measures.

6.3.3.4.1 Baseline demographic information.

Baseline demographic information occurred during the initial assessment of patients as described in Chapter 5, Section 5.3.3.4.

6.3.3.4.2 10 m ISWT.

The two 10 m ISWTs were completed according to the description in Chapter 3, Section 3.3.3.2 and Chapter 5, Section 5.3.3.6.

6.3.3.4.3 Global rating of change scale.

During the post-program assessment, before the completion of the 10 m ISWTs, participants were asked to score their global rating of change on a 5-point Likert scale (de Vet et al., 2011, pp. 214-215). The response options were: improved significantly, improved a little, about the same, slightly worse, or significantly worse. The results of the global rating of change were then recoded into three categories: improved, unchanged and deteriorated. Improved significantly and improved a little were recoded to improved, and slightly worse or significantly worse were recoded to deteriorated.

6.3.4 Statistical analysis.

Participant characteristics were recorded to allow description of the study sample according to the global rating of change, including the percentage of respondents with the lowest and highest possible scores (de Vet et al., 2011, p. 228). The scores and change scores for the three groups of participants were described and checked for any statistically significant difference in the distance walked between the three groups.

6.3.4.1 Smallest detectable change.

The SDC was calculated using the 95% confidence intervals or limits of agreement for both the group, as well as the individuals (de Vet et al., 2011, pp. 242-243). Confidence intervals for the group mean scores were calculated using the following equation (Taylor, Dodd, et al., 2004):

$$95\%\ CI_{group} = M_{diff} \pm \frac{t_{0.975} \times SD_{diff}}{\sqrt{n}}$$

Where M_{diff} is the mean difference of retest minus test scores and takes into account the systematic error, SD_{diff} is the standard deviation of the differences between retest and test scores, n is the number of participants and $t_{0.975}$ is the critical value for t with a two-tailed test at that sample size.

The 95% confidence intervals or limits of agreement for the individual change score were calculated for $n = 1$ (Taylor, Dodd, et al., 2004):

$$95\%\ CI_{individual} = M_{diff} \pm t_{0.975} \times SD_{diff}$$

The SDC was interpreted as the change that falls outside of these limits of agreement.

6.3.4.3 Reliable change index.

The reliable change index (RCI) is another method for observing if the change that has occurred in an individual is beyond measurement error or the SDC, in other words true change. The RCI for an individual is calculated using the formula: $(X_{Post} - X_{Pre})/\sqrt{2 \times SEM^2}$ (Jacobson & Truax, 1991). An RCI of 1.96 or greater reflect real change with 95% confidence. Using the RCI, the change score that must be exceeded to be interpreted as true change with 95% confidence was calculated by:

$$95\%\ Cutoff = \pm 1.96 \times \sqrt{2 \times SEM_{Absolute}{}^2}$$

The number of participants who exceeded the 95% limit cut-off points were reported for the change scores of the first pre- and post- walks and the second pre- and post-walks.

6.3.4.4 Minimal important change.

Both the MIC reflecting participants' perceived improvement in physical fitness and functional capacity over a cardiac rehabilitation program, and the MIC reflecting participants' perceived deterioration in physical fitness and functional capacity over a cardiac rehabilitation program, were estimated. The MIC for improvement and deterioration perceived by the participant was determined three ways: first, by using 95% limit cut-off points ($MIC_{95\%Cut\ off}$); second by using receiver operating curve (ROC) analysis (MIC_{ROC}) (de Vet et al., 2007; de Vet et al., 2011, p. 246); and third, using predictive modelling (MIC_{pred}) (Terluin, Eekhout, Terwee, & De Vet, 2015).

The MIC values were calculated for the change scores between the first walks completed before and after cardiac rehabilitation (i.e., ISWT 1 and ISWT 3) and the second walks completed before and after cardiac rehabilitation (i.e., ISWT 2 and ISWT 4).

6.3.4.4.1 Minimum important change using 95% limit cut-off points.

The MIC values were first estimated using the 95% limit cut-off points ($MIC_{95\%cutoff}$). These scores reflect the variability of the sample who rated themselves as unchanged according to the global rating of change anchor (de Vet et al., 2007), and were

unrelated to the distribution of the sample who rated themselves as improved or deteriorated. They were calculated using the following formula:

$$95\%\ limit\ cutoff\ point = Mean\ change \pm 1.645 \times SD_{change}$$

Where 1.645 was a constant corresponding to the 5% upper and lower limit of a one-tailed test (de Vet et al., 2007).

6.3.4.4.2 Minimal important change using ROC analysis.

The MIC values for improvement and deterioration were also estimated using ROC curve analysis (MIC_{ROC}). This method used the global rating of change scale as an external criterion as an anchor to determine the optimal ROC curve cut-off points for discriminating between patients who had improved or deteriorated, and those patients who were unchanged (de Vet et al., 2007). Equal weighting was placed on minimising the false positives and false negatives, so the ROC curve cut-off was the value that corresponded with the least misclassification, determined by the sum of percentages of false positive and false negative classification. In other words, the optimal cut-off point maximised the score for $[1 - sensitivity] + [1 - specificity]$ (de Vet et al., 2007).

6.3.4.4.3 Minimal important change using predictive modelling.

The third method to predict the MIC was using predictive modelling (MIC_{pred}). Here, the MIC_{pred} with 95% confidence intervals for improvement and deterioration were calculated according to the method described by Terluin et al. (2015). Logistic

regression analysis was performed for the improved and unchanged groups and the deteriorated and unchanged groups for the equation:

$$\ln(odds_{post}) = C + B_X \times X$$

Where $odds_{post}$ was the post-test odds of being improved (or deteriorated) based on an initial pre-cardiac rehabilitation 10 m ISWT distance, C represented the intercept, B_X the regression coefficient of the change score and X the change score. Using this method, Terluin et al. (2015) defined the MIC_{pred} as the score that corresponded to a likelihood ratio (LR) of one, meaning that the probability of belonging to the improved (or deteriorated) group was equal to the probability of being improved (or deteriorated):

$$X = \frac{(\ln(odds_{pre}) - C)}{B_X}$$

Where, $odds_{pre}$ equals 1 for a LR of 1, given the formula: $LR = odds_{post}/odds_{pre}$.

The MIC_{pred} 95% confidence intervals were calculated according to Terluin et al. (2015). In addition, the likelihood ratio and post-test probability with 95% confidence intervals for improvement and deterioration were calculated to develop a predictive framework for interpreting the change scores in a cardiac rehabilitation population.

6.4 Results

6.4.1 Characteristics of the sample.

Fifty-two patients completed two 10 m ISWTs prior to, and two 10 m ISWTs after an outpatient cardiac rehabilitation program. The characteristics of the sample were reported in Chapter 5, Section 5.4.1.

6.4.1.1 Global rating of change.

The perceived global rating of change was assessed at the post-program assessment. Of the 52 participants, 32 reported an improvement in their global rating of change, 10 reported their global rating of change was unchanged and 10 reported deterioration in their global rating of change (Table 6.1).

Table 6.1

Characteristics of the Sample According to Global Rating of Change

	Global Rating of Change score				
	Deteriorated	Unchanged	Improved	*F*(2,49)	*p*
Gender (M:F) (*n*)	8:2	7:3	25:7	-	-
Age years (*SD*)	72 (13)	73 (10)	65 (9)	3.201	.049
Height cm (*SD*)	173 (7)	166 (9)	173 (6)	4.051	.024
Weight kg (*SD*)	84 (11)	76 (11)	87 (11)	2.456	.096
BMI kg.m-2 (*SD*)	28 (3)	28 (5)	29 (5)	.367	.694

There was a significant difference between the improved and unchanged group for age ($p = .035$), but not for other comparisons. For height, there was a significant

difference between both the deteriorated and unchanged group (p = .033), and the improved and unchanged groups (p = .008). There was no difference in global rating of change according to gender χ^2 (2, n = 52) = .349, p = .840.

6.4.1.2 10 m ISWT results according to the global rating of change.

The distance walked in the 10 m ISWTs according to global rating of change are summarised in Table 6.2. There were no statistically significant differences in walk distances at pre-program assessment (ISWT 1 and ISWT 2) according to the global rating of change: deteriorated, unchanged, and improved. There were statistically significant differences in the post-program 10 m ISWT distances (ISWT 3 and ISWT 4) according to the global rating of change: deteriorated, unchanged, and improved (Table 6.2). For ISWT 3, *post hoc* analysis showed a statistically significant difference between the improved group and the unchanged group (M_{diff} 172 m p = .012, 95% CI [39, 305]), and the improved group and deteriorated group (M_{diff} 238 m p = .001, 95% CI [105, 371]). For ISWT 4, *post hoc* analysis showed a statistically significant difference between the improved group and the unchanged group (M_{diff} = 184 m, p = .008, 95% CI [50, 318]), and the improved group and deteriorated group (M_{diff} = 255 m, p < .001, 95% CI [121, 389]). For both ISWT 3 and ISWT 4, the difference between the distance walked for the deteriorated and unchanged groups were not statistically significant.

Table 6.2

10 m ISWT Distance based on Global Rating of Change

	10 m ISWT (m) *M* (*SD*)				
Test	Deteriorated	Unchanged	Improved	F(2,49)	*p*
ISWT 1	305 (131)	331 (142)	416 (172)	2.387	.103
ISWT 2	318 (130)	348 (139)	435 (172)	2.625	.083
ISWT 3	322 (130)	388 (145)	560 (205)	8.056	.001
ISWT 4	326 (128)	397 (151)	581 (205)	9.138	< .001

Table 6.3 summarises the change scores across the cardiac rehabilitation program according to the global rating of change. There was a statistically significant difference for improvement in walk distance between the three groups over an eight-week cardiac rehabilitation program. When the first tests were used (i.e., ISWT 1 and ISWT 3), *post hoc* analysis showed a statistically significant difference in the change scores of the improved group, compared with both the unchanged group (M_{diff} = 86 m, p < .001, 95% CI [90, 162]), and the deteriorated group (M_{diff} = 126 m, p < .001 95% CI [91, 162]). There was no statistically significant change in the deteriorated and unchanged group (M_{diff} = 40 m, p = .076, 95% CI [-4, 84]). When the second tests were used (i.e., ISWT 2 and ISWT 4), *post hoc* analysis showed a statistically significant difference in the change scores of the improved group, compared with both the unchanged group (M_{diff} = 97 m, p < .001, 95% CI [62, 132]), and the deteriorated group (M_{diff} = 138 m, p < .001, 95% CI [103, 173]). There was no statistically significant change in the deteriorated and unchanged group (M_{diff} = 41 m, p = .063, 95% CI [-2, 84]).

Table 6.3

Change in 10 m ISWT Distance Walked based on the Global Rating of Change

	10 m ISWT change scores m (SD_{diff})				
Tests	Deteriorated	Unchanged	Improved	$F(2,49)$	p
ΔISWT 1 to 3	17 (40)	57 (21)	143 (52)	30.318,	< .001
ΔISWT 2 to 4	8 (35)	49 (20)	146 (57)	38.212	< .001

Note. A positive score indicates an increase in distance walked post-program

For all groups, the mean change in 10 m ISWT distance increased over an eight-week cardiac rehabilitation program (Table 6.3). There was a significant increase between the pre-program and post-program change in 10 m ISWT distances for the improved group (Walk 1 and 3: $t(31) = 14.207$, $p < .001$, 95% CI [123, 164], and Walk 2 and 4: $t(31) = 14.537$, $p < .001$, 95% CI [125, 166]), and the unchanged group (Walk 1 and 3: $t(9) = 8.761$, $p < .001$, 95% CI [42, 72] and $t(9) = 7.869$, $p < .001$, 95% CI [35, 63]). For the deteriorated group, there were no significant changes between pre-program and post-program distances (Walk 1 and 3: $t(9) = 1.353$, p < .209, 95% CI [-11, 45], and Walk 2 and 4: $t(9) = .732$, $p = .483$, 95% CI [-17, 33]).

6.4.2 Smallest detectable change.

The SDC scores were derived from the 95% confidence intervals of the 10 m ISWT distance walked in the pre-program assessment (Table 6.4). Using the results from the two 10 m ISWTs that occurred at the pre-program assessment (ISWT 1 and ISWT 2), the SDC for the group was 22 m and for the individual was 47 m. The results were similar using the results from the two 10 m ISWTs that occurred at the post-program assessment (ISWT 3 and ISWT 4), the SDC for the group was 22 m and for the individual was 58 m.

Table 6.4

The 95% Confidence Intervals for the Group and Individual (Limits of Agreement) for the Change in 10 m ISWT Distance when Repeated in a Single Session

Tests	M_{diff} (m) (SD_{change})	Group 95% CI		Individual 95% CI	
		LL	UL	LL	UL
ΔISWT 1 and 2	17 (14)	13	21	-12	46
ΔISWT 3 and 4	16 (21)	10	21	-26	57

6.4.3 Reliable change index.

The cut-off value for the reliable change score with 95% confidence was 44 m, and was based on the $SEM_{absolute}$ of 16 m for the sample. The percentage of participants whose change scores exceeded this cut off when one test was performed (i.e., ISWT 1 and ISWT 3) was 83% (n = 43) and when the results of the second test were used (i.e., ISWT 2 and ISWT 4) was 77% (n = 40).

6.4.4 Minimal important change.

The MIC was calculated for improvement and deterioration over an eight-week comprehensive cardiac rehabilitation program. The predicted MIC for participants to report an improvement ranged from 70 to 92 m. The predicted MIC for patients to report unchanged, and not a deterioration, ranged from 16 to 42 m (Table 6.5).

Table 6.5

The Minimal Important Change in 10 m ISWT Walk over an Eight-Week Comprehensive cardiac rehabilitation program

	Improvement					Deterioration				
	$MIC_{95\%cutoff}$	MIC_{ROC}	MIC_{Pred}	95% CI		$MIC_{95\%cutoff}$	MIC_{ROC}	MIC_{Pred}	95% CI	
Tests	(m)	(m)	(m)	LL	UL	(m)	(m)	(m)	LL	UL
ΔISWT 1 to 3	92	85	71	45	89	22	Unable to determine	42	-793	90
ΔISWT 2 to 4	82	85	70	43 to	86	16	25	32	-19	71

Note. $MIC_{95\%cutoff}$ = minimum important change calculated using 95% limit cut-off points; MIC_{ROC} = minimum important change calculated using ROC analysis; MIC_{Pred} = minimum important change calculated using predictive modelling; CI = confidence intervals; LL = lower limit; UL = upper limit.

The sensitivity and specificity at the MIC_{ROC} for improvement was good. Sensitivity was 91% and specificity was 100%. In other words, assuming a MIC_{ROC} of 85 m represents meaningful improvement: 9% of the patients who rated themselves as improved were misclassified as unchanged. All patients who rated themselves as unchanged were correctly classified.

The results of the logistic regression are presented in Tables 6.6 and 6.7. The predictive framework showing meaning to the change scores is shown in Tables 6.8 and 6.9. For example, a patient with a change score of 100 m, based on the results of a single test in each session (i.e., ISWT 1 and ISWT 3), has a likelihood ratio of 8, 95% CI [2, 40] and a probability of being improved of 89%, 95% CI [62, 98].

Table 6.6

Logistic Regression results for the Groups that Rated Global Rating of Change as Improved and Unchanged

Test	Intercept (C)	se (C)	Regression coefficient (*B*)	se (*B*)	*r* (C-B)
ΔISWT 1 to 3	-5.119	1.985	.072	.025	-.959
ΔISWT 2 to 4	-8.509	3.816	.122	.052	-.983

Note. Intercept (C) = constant of the regression equation; se (C) = standard error of the intercept (C); se (B) = standard error of the regression coefficient; *r*(C-B) = correlation coefficient between the intercept (C) and the regression coefficient (B).

Table 6.7

Logistic Regression Results for the Groups that Rated Global Rating of Change was Deteriorated and Unchanged

	Intercept (C)	Se (C)	Regression coefficient *(B)*	SE *(B)*	*r* (*C*-*B*)
ΔISWT 1 to 3	2.553	1.506	-.061	.031	-.929
ΔISWT 2 to 4	2.542	1.386	-.079	.038	-.903

Note. Intercept (C) = constant of the regression equation; se (C) = standard error of the intercept (C); se (B) = standard error of the regression coefficient; *r*(C-B) = correlation coefficient between the intercept (C) and the regression coefficient (B).

6.5 Discussion

This study presented thresholds for the MIC of the 10 m ISWT distance perceived by a patient over a mixed outpatient cardiac rehabilitation program. The MIC for patients to perceive any improvement in physical fitness and functional capacity ranged from 70 to 92 m. The MIC for patients who reported any deterioration in physical fitness and functional capacity ranged from 16 to 42 m. The MIC thresholds variation were due to the statistical method used and whether one or two tests were performed in each session.

Table 6.8

Likelihood Ratios of Patients Improved and Unchanged Related to Pre-Program to Post-Program Change in 10 m ISWT Distance

	ISWT 1 and ISWT 3						ISWT 2 and ISWT 4					
Change Score	Likelihood Ratio	95% CI		Post-test Probability	95% CI		Likelihood Ratio	95% CI		Post-test Probability	95% CI	
		LL	UL		LL	UL		LL	UL		LL	UL
0	.006	0	0.29	.006	0	.23	0	0	0.36	0	0	0.26
20	.025	.001	.49	.025	.001	.33	.002	0	.56	.002	0	.36
40	.11	.013	.86	.10	.013	.46	.027	.001	.93	.026	.001	.48
60	.45	.12	1.75	.31	.10	.64	.31	.048	1.93	.23	.046	.66
80	1.90	.62	5.81	.66	.38	.85	3.49	.712	17.13	.78	.42	.95
100	8	1.61	39.96	.89	.62	.98	40.1	1.7	935.5	.98	.63	.999
120	33.82	3.02	378.53	.97	.75	1	459.9	2.9	73042.4	.998	.74	1
140	142.74	5.17	3938.32	099	.84	1	5276	4.6	6092672	1	.82	1
160	602.45	8.56	42416.2	0.998	.90	1	60536	7	5206932	1	.88	1
			2						73			
180	2542.7		464165.	1	.93	1	694537	11	4499547	1	.92	1
			94						6743)			
200	10732.16	22.5	5123073	1	.96	1	7968446	16	3911507	1	.94	1
			.5						709395			

Note. CI = confidence interval; LL = lower limit; UL = upper limit.

Table 6.9

Likelihood Ratios of Patients Deteriorated and Unchanged, Related to Pre-Program to Post-Program Change in 10 m ISWT Distance

	ISWT 1 and ISWT 3						ISWT 2 and ISWT 4					
Change Score	Likelihood Ratio	95% CI		Post-test Probability	95% CI		Likelihood Ratio	95% CI		Post-test Probability	95% CI	
		LL	UL		LL	UL		LL	UL		LL	UL
-60	499	.77	325630	1	.434	1	1454	1.30	1625888	1	.57	1
-40	147	.75	29134	.99	.43	1	299	1.16	77515	1	.54	1
-20	43.5	.72	2639	.98	.42	1	61.7	1.01	3766	.98	.50	1
0	12.8	.67	246	.93	.40	1	12.7	.84	192	.93	.46	1
20	3.73	.58	24.8	.79	.37	.96	2.62	.58	11.9	.72	.37	.92
40	1.12	.36	3.49	.53	.26	.78	.54	.15	1.94	.35	1.30	.66
60	.33	.08	1.36	.25	.074	.58	.11	.011	1.14	.10	.011	.533
80	.098	.009	1.06	.089	.009	.51	.023	.001	.92	.022	.001	.48
100	.029	.001	.962	.028	.001	.49	.005	.00003	.78	.005	.00003	.44

Note. CI = confidence interval; LL = lower limit; UL = upper limit.

This was the first study to use predictive modelling to report the MIC with 95% confidence intervals for the 10 m ISWT distance over the duration of an eight-week comprehensive cardiac rehabilitation program. When binary logistic regression was used, the MIC for improvement in physical fitness and functional capacity was 71 m, 95% CI [45, 89] when one test was performed, and 70 m, 95% CI [43, 86] when a second test was performed in a single session. In other words, when one test is performed in each session, a change score of 71 m means the likelihood of being improved exceeds the average probability of being unchanged (Terluin et al., 2015). These findings are consistent with the previously published MIC of 70 m for patients to rate their performance as slightly better, and 85 m for patients to rate their performance as better over a six-week cardiac rehabilitation program (Houchen-Wolloff et al., 2015).

The MIC obtained from binary logistic regression was similar when one test or two tests were performed, thereby providing further support that a practice walk test is not required. However, the $MIC_{95\%cutoff}$ ranged from 92 m when one walk test was performed and 85 m when a second walk test was performed. The $MIC_{95\%cutoff}$ is a distribution-based method for calculation of the MIC, and is based on the variability of the sample who rated their physical fitness and functional capacity as unchanged. The variation may have been a result of the small sample size for the unchanged group (n = 10) and would need further investigation.

The threshold for the MIC for improvement in physical fitness and functional capacity over an outpatient cardiac rehabilitation was higher than the SDC. For the individual, the limits of agreement calculated between the first and second walk, completed before cardiac rehabilitation, were -12 to 46 m. A change score within this range on

the 10 m ISWT could not be detected as true change. This study showed the SDC to be 47 m, which approximates the lower 95% confidence limits of the MIC_{pred} of 45 and 43 m and provides further support of the clinical relevance and interpretability of the 10 m ISWT in this population (Scholtes et al., 2011).

An unexpected finding was that cardiac rehabilitation participants did not perceive small improvements in the 10 m ISWT distance as meaningful improvements in their physical fitness and functional capacity. Small improvements may even be associated with a self-rated deterioration in physical fitness and functional capacity. The MIC for patients to detect a deterioration in physical fitness and functional capacity was different to that for improvement. The MIC_{pred} for deterioration and using binary logistic regression was 42 m, 95% CI [-793, 90] when one test was performed, and 32 m, 95% CI [-19, 71] when a second test was performed. This is an important consideration in clinical practice as an increased score of 30 or 40 m may not reflect a meaningful improvement, and patients may still rate their physical fitness and functional capacity as not being improved. This may be a result of the expectations of patients attending cardiac rehabilitation.

Overall, the interpretation of the MIC for a patient to rate themselves as unchanged and deteriorated, in terms of physical fitness and functional capacity, is less clear. The threshold MIC_{pred} for a patient to report a deterioration in physical fitness and functional capacity or being unchanged, was 42 m when one test was performed and 32 m when two tests were performed in a single session. The wide confidence intervals reflect the large variability in the small sample size of the group and makes it difficult to interpret, especially when only one test in each session was performed. Further investigation, with a larger sample size, is required to determine the minimal

important change for those patients who reported a deterioration or unchanged on the global rating of change.

This study was the first to present a framework for interpreting individual change scores using probabilities, odd ratios and likelihood ratios of the 10 m ISWT in cardiac rehabilitation. This information provides another tool for the clinician when interpreting change scores after cardiac rehabilitation. When one test is performed, a patient with a change score of 40 m increase in 10 m ISWT distance has a likelihood ratio of 1.12 for deterioration and a probability of self-reporting a deterioration in physical fitness and functional capacity of 53%, 95% CI [26, 78] and a likelihood ratio of improvement of 0.11 and a probability of reporting an improvement in physical fitness and functional capacity of 10%, 95% CI [1, 40]. In comparison, with a change of 80 m increase in 10 m ISWT distance, the patient has a likelihood ratio of 0.1 and probability of being deteriorated of 9%, 95% CI [.1, 51]; and for improvement a likelihood ratio of 1.9, 95% CI [38, 85] and a probability of being improved of 89%, 95% CI [62, 98].

Retrospective self-rating of change as an anchor is a commonly used method for determining the MIC. However, retrospective ratings, particularly over an extended period, are susceptible to recall bias and in this case the reliability and validity of the global rating of change question was unknown.

Future research could focus on the effect of the severity of cardiac disease on the MIC as well as the effect of the initial performance on the 10 m ISWT. Using predictive modelling, it is possible to investigate the effect of severity of disease or different diseases on the minimum important change with larger sample sizes. In previous

studies and different populations, MIC values have varied using alternative anchors. Future research that uses multiple anchors will be important.

6.5.1 Implications.

The MIC for patients who report any improvement in physical fitness and functional capacity is 70 m. These results suggest that over an eight-week program, patients would need to improve by at least seven shuttles to perceive an improvement in their physical fitness and functional capacity. Patients with small increases in the 10 m ISWT distance may still report a deterioration in their physical fitness and functional capacity.

Chapter 7: Do Field Exercise Tests Accurately Measure Physical Fitness and Functional Capacity in the Cardiac Rehabilitation Assessment? A Systematic Review of Measurement Properties

7.1 Chapter Aims

The purpose of this study was to address the secondary research aim 10, specifically to synthesise the available evidence of commonly used field exercise tests in cardiac rehabilitation.

7.2 Introduction

In cardiac rehabilitation, exercise tests are used to evaluate physical fitness and functional capacity, mobility and balance, and response to and recovery from exercise. Serial testing can provide further information on monitoring change in physical fitness and functional capacity and response to the program for both an individual and a group. The gold standard for measurement of physical fitness and functional capacity is the symptom-limited maximal exercise test, preferably with measurement of peak oxygen uptake while the patient is on his or her usual medication regimen. Compared with a field exercise test, this requires additional equipment and staffing, as well as time to complete (Reeves, Gupta, & Forman, 2016). These factors have the potential to increase cardiac rehabilitation service-delivery costs, delay the commencement of cardiac rehabilitation and this may affect successful adherence and patient outcomes (Pack et al., 2013; Reeves et al., 2016; Reeves & Whellan, 2010).

Staff within the cardiac rehabilitation program with minimal additional costs can complete field exercises tests. The two most commonly used field tests are the 6MWT and the 10 m ISWT. The results presented in Chapters 2 and 3 suggest that both the 6MWT and the 10 m ISWT have a high level of relative reliability when measured for consistence. The systematic error and the large measurement error in the 6MWT may be unacceptable in patients commencing an outpatient cardiac rehabilitation program. The systematic error and measurement error of the 10 m ISWT was small and acceptable in this group. In a small sample, the evidence supporting the concurrent criterion validity of the 10 m ISWT was demonstrated when the distance walked was compared with the time achieved in the Bruce protocol in cardiac patients following their usual medication regimen. Despite demonstrating adequate concurrent criterion validity, the predictive criterion validity of the 10 m ISWT was not supported, seen by a wide variability when the 10 m ISWT distance was used to predict the Bruce test protocol time with 95% confidence. In a sample of 52 patients commencing a cardiac rehabilitation program, evidence supporting the construct validity and the responsiveness of the 10 m ISWT was demonstrated (Chapter 5). The evidence for the validity of the 6MWT was not investigated in this thesis.

The aim of this systematic review was to examine the measurement properties of the field exercise tests used in adults eligible for cardiac rehabilitation and to report which field exercise test is the most suitable test to measure physical fitness and functional capacity in cardiac rehabilitation. Specific questions addressed in this systematic review were: 1) Do the field tests used in cardiac rehabilitation show adequate reliability, and evidence of validity?; 2) Can field exercise tests with adequate measurement properties identify important change in people eligible for cardiac

rehabilitation (interpretability)?; 3) Is one test more suitable than others at measuring physical fitness and functional capacity in cardiac rehabilitation?

7.3 Methods

A study protocol was developed *a-priori* to define the objectives, eligibility criteria, and outcomes of interest. This review was registered prospectively at PROSPERO (registration number CRD42016030092). The systematic review has been reported consistent with the Preferred Reporting Items for Systematic reviews and Meta-Analysis (PRISMA) guidelines (Moher, Liberati, Tetzlaff, Altman, & The PRISMA Group, 2009).

7.3.1 Search strategy.

Studies that reported measurement properties of field exercise tests in cardiac rehabilitation were identified by searching eight electronic databases (Allied and Complementary Medicine (AMED), Cumulative Index to Nursing and Allied Health Literature (CINAHL), Cochrane, Embase, Medline, PsycINFO, PubMed, SPORTDiscus) from the earliest time available until February 2016. There were three concepts in the search strategy: cardiac rehabilitation, field exercise tests and measurement properties. For each concept MeSH terms and free text that related to the concept were combined with OR (Appendix 6). The search terms for the concept of measurement properties were developed from recommendations for a highly sensitive search for retrieving studies on measurement properties (Terwee, Jansma, Riphagen, & de Vet, 2009) and terminology used in the Consensus-based Standards for the selection of health measurement Instruments (COSMIN) checklist (Mokkink et

al., 2010c). The three concepts were combined using 'AND'. Comprehensive search strategies were developed specific for each database. The search terms used for the Medline (Ovid) and Cinahl (EBSCO) databases are shown in Appendix 6. In addition to the electronic searches, reference lists of the included studies were screened for any potentially relevant studies, and authors were screened for additional relevant studies. Restrictions were English language, humans and adults.

7.3.2 Study selection.

Selection was based on predetermined inclusion and exclusion criteria (Table 7.1). The selection criteria aimed to include studies that evaluated the measurement properties of the field exercise tests in adult patients eligible for cardiac rehabilitation. After the removal of duplicates, the search yields were reviewed first on title and abstract, and second on full text of all remaining articles against the selection criteria. Inter-rater reviewer agreement was calculated by Kappa (κ) with 95% confidence intervals (*CI*). The strength of agreement was determined using the following scale 0.01-0.20 slight agreement; 0.21-0.40 fair agreement; 0.41-0.60 moderate agreement; 0.61-0.80 substantial agreement; 0.81-0.99 almost perfect agreement (Landis & Koch 1977). For the title and abstracts, two reviewers independently reviewed batches of 200 studies until agreement was at least substantial (i.e., $\kappa \geq .61$), after which the remaining titles and abstracts were reviewed by a single reviewer. Two reviewers independently reviewed the full text of all remaining articles. For both title and abstract and full text selection, the reviewers discussed any selection disagreements, if disagreement persisted after discussion a third reviewer was consulted.

Table 7.1

Study Selection Criteria

Inclusion	Exclusion
Population	Population
Human adults (≥18 years or older) with diagnosed and managed heart disease eligible for cardiac rehabilitation	Healthy subjects with no diagnosis of heart disease
	Principal diagnosis of pulmonary hypertension
Indicator	Congenital heart disease
Any field exercise test	Risk factors for heart disease
Retest or repeated testing	Indicators
CPET or symptom-limited stress test	Diagnostic testing
Field exercise test	Drug trial
Any measure of physical fitness and functional capacity, exercise or health-related quality of life used in cardiac rehabilitation	Device testing (e.g., external heart, ECG device)
	Non-field exercise test
Outcome	Language
Reporting of psychometric properties of functional outcome measures of field exercise tests according to COSMIN	Other than English
	Design
Design	Review
Any clinical study	Commentary
	Conference proceeding/poster/abstract,
	Diagnostic study
	Prognostic study,
	Case report study

Note. CPET – cardiopulmonary exercise test; ECG – electrocardiogram; COSMIN = COnsensus-based Standards for the selection of health Measurement INstruments.

7.3.3 Data extraction.

A data extraction form was developed by the reviewers and included items related to publication details, population, field test, methods including the sample size, psychometric properties assessed and statistical analysis, results, and any key points

of concerns (Appendix 7). The field test protocols were compared with standard administration. For studies reporting measurement error, the upper limit of the 95% limits of agreement and the SEM were used to estimate the SDC. The SDC was calculated from the SEM using the formula $1.96 \times \sqrt{2} \times SEM$ (Mokkink et al., 2010b). Two reviewers extracted data independently. The two reviewers discussed each data sheet for consistency and accuracy of data extraction.

7.3.4 Methodological quality assessment.

Two reviewers, independently using a modified COSMIN checklist (Appendix 7), assessed the assessment of methodological quality. The checklist is a standardised tool and designed to evaluate the measurement properties of health-related patient-reported outcomes but has been used for other outcome measures. The checklist was chosen for its comprehensiveness and was considered the best assessment tool for this type of review. It has been used by other systematic reviews on the measurement properties of field exercise tests in a range of health conditions (Bartels et al., 2013; Kroman, Roos, Bennell, Hinman, & Dobson, 2014). The checklist was adapted for exercise tests by substituting the term *health-related patient-reported outcomes* with *field exercise test* (see Appendix 7).

The COSMIN checklist contained 12 boxes, with 10 relating to methodological quality of measurement properties and 2 boxes relating to general requirements. Seven boxes were considered relevant to assessing the measurement properties of field exercise tests. These included methodological assessment of relative reliability, measurement error, validity including content, hypothesis testing, cross-cultural, criterion and responsiveness. Internal consistency and structural validity were not

relevant measurement properties of field exercise tests (Bartels et al., 2013). The box titled Interpretability did not generate the same information on the 4-point scale and was not included in the analysis of methodological quality.

For the boxes that related to measurement properties, each item was scored according to the 4-point rating system: excellent, good, fair, poor. For each measurement property, the quality assessment was based on the methodology: including percentage and description of missing data, administration of the test and appropriateness of statistical analysis. For sample size, the COSMIN group recommend a minimum sample size $n = 30$, which was based on evaluating the quality of questionnaires. Previous researchers have argued that psychometric studies on performance measures such as exercise tests generate larger effect sizes and may be appropriately tested using smaller sample sizes (Bartels et al., 2013). Consistent with previous systematic reviews (Bartels et al., 2013; Dobson et al., 2012), the question on sample size was omitted from the scoring of the methodological quality of each study. The sample size was accounted for in the evidence synthesis stage. With the sample size question not included in the quality analysis, the recommended scoring system of worst score counts was followed, so the quality score for each measurement property was the lowest rating given for any item in the box. Depending on the number of measurement properties evaluated in the study, some studies received one quality evaluation whereas others received multiple evaluations. No study was eliminated based on a rating of poor methodological quality.

7.3.5 Quality criteria for assessing the adequacy of measurement properties.

The quality criteria for the adequacy of measurement properties were proposed by Terwee et al. (2007) and were used to rate each measurement property in the studies as positive (+), negative (-) or indeterminate (?) (Table 7.2).

7.3.6 Best evidence synthesis for each measurement property.

A best evidence synthesis was performed for the measurement properties of each field exercise test for specific cardiac patient groups (Table 3). The evidence was assigned either in support of the measurement property (+) or not in support of the measurement property (-), with all the available evidence considered. The synthesised evidence was assigned as high quality (+++ or ---), moderate quality (++ or --), limited (+ or -) or unknown (?) based on the number of studies, their methodological quality, the consistency of findings and the total sample size (Table 7.3). The inclusion of sample size in the best evidence synthesis was consistent with earlier high quality systematic reviews and supported by the developers of the COSMIN (Dobson et al., 2012).

Table 7.2

Criteria for the Adequacy of Measurement Properties adapted from (Bartels et al., 2013)

	Positive (+)	Negative (-)	Indeterminate (?)
Relative reliability	ICC/weighted kappa ≥ .70 or Pearson's *r* ≥.80	ICC/weighted kappa <.70 or Pearson's *r* <.8	Neither ICC/weight kappa nor Pearson's *r* reported
Measurement error	MIC > SDC Or MIC outside LOA Or adequate author justification	MIC ≤ SDC Or MIC equals or inside LOA Or adequate author justification	MIC not determined Or inadequate author justification
Construct validity	≥ 75% of hypotheses supported Or correlation with related constructs > unrelated constructs Or correlation with related construct ≥ .50	< 75% of hypotheses support Or Correlation with related constructs < unrelated constructs Or correlation with related construct < .50	Correlations determined with unrelated constructs
Criterion Validity	Justification of gold standard and correlation ≥ .70	Correlation with gold standard < .70 despite adequate methodology	Gold standard not justified or doubt design or method
Responsiveness	≥ 75% of hypothesis supported Or correlation with change score of related construct ≥ .50 Or AUC ≥ .70 Correlation of change score of related construct > unrelated construct	< 75% hypothesis supported Or AUC < .70 Or correlation of change scores with related constructs < unrelated constructs	Correlations determined with unrelated constructs

Note. ICC = intraclass correlation coefficient; *r* = Pearson's product-moment correlation coefficient; MIC = minimal important change; SDC = smallest detectable change; LOA = limits of agreement; AUC = area under the curve.

Table 7.3

Synthesis criteria for levels of evidence adapted from Terwee et al. (2007)

Quality level of evidence	Rating	Criteria
High	+++ OR ---	Multiple studies with methodological quality rated as at least good OR One study with methodological quality rated as Excellent AND consistent findings ((+) OR (-)) AND combined sample size ≥100
Moderate	++ OR --	Multiple studies with methodological quality rated as Fair OR one study with methodological quality rated as Good AND consistent findings ((+) OR (-)) AND combined sample size 50 - 99
Limited	+ or -	One study with methodological quality rated as Fair AND consistent findings ((+) OR (-)) AND combined sample size 25 - 49
Conflicting	+/-	Multiple studies AND conflicting findings
Unknown (No evidence)	?	Only studies with methodological quality rated as Poor OR only studies with results rated as (?) AND combined sample size < 25

7.3.7 The interpretability of the field exercise tests in cardiac rehabilitation.

The second aim of this systematic review was to describe the interpretability of the field exercise tests that demonstrated adequate measurement properties. A qualitative

synthesis of the available literature on floor and ceiling effects, the amount of change, and the MIC was performed. The range of baseline tests results was collected from the included studies and reported as well as change scores for patient groups presented in the literature. Floor and ceiling effects were considered adequate and given a positive score if there was less than 15% of participants scoring the lowest or highest possible score (McHorney & Tarlov, 1995; Terwee et al., 2007). The MIC was also reported.

7.3.8 Suitability of the field exercise tests in cardiac rehabilitation.

The final aim of this systematic review was to synthesise the available literature to report on the most suitable field exercise test to measure physical fitness and functional capacity in cardiac rehabilitation. The results of the measurement properties, interpretability and feasibility of the field exercise test were collated and presented in a table in a descriptive synthesis.

7.4 Results

7.4.1 Study selection.

The electronic database search resulted in 13,274 studies. After removal of duplicates, 9,513 studies were considered for inclusion in the systematic review. Two reviewers excluded studies that clearly did not meet the inclusion criteria based on title or abstract in batches of 200 before reviewing and discussing any selection disagreements (Round 1: $n = 200$, $\kappa = .72$; Round 2: $n = 200$, $\kappa = .72$; Round 3: $n = 200$, $\kappa = .78$). Following the consistent substantial agreement between the two

reviewers, one reviewer screened the remaining titles and abstracts. There were 106 studies for full text consideration. These were reviewed by two reviewers for inclusion in the systematic review with almost perfect agreement (κ= .93). The final library of studies consisted of 64 retrieved from the electronic databases and 14 retrieved through manual searches such as reference list checks and first author searches (see Appendix 8). This process is summarised in Figure 7.1.

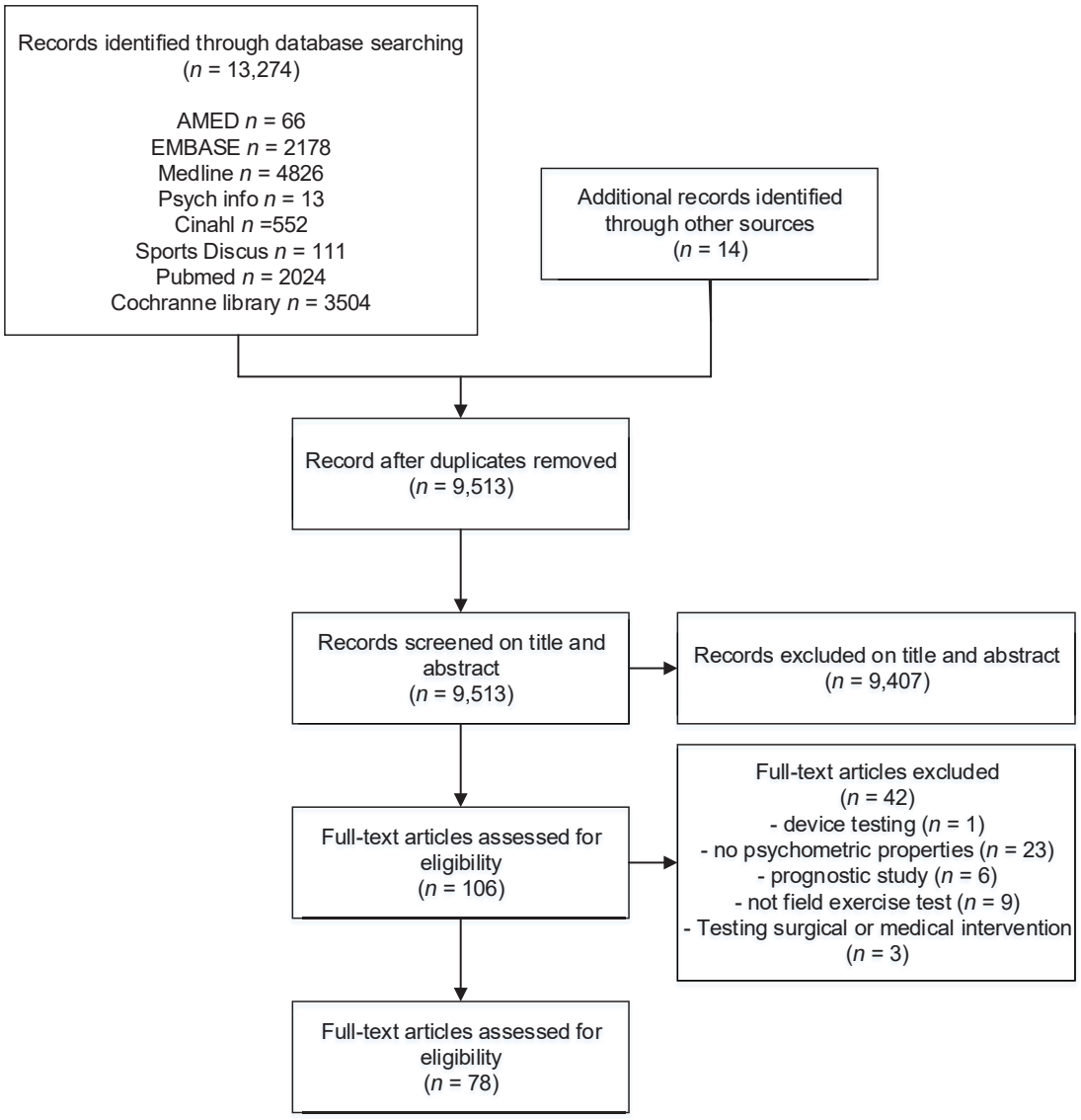


Figure 7.1. Flow chart of literature search and selection process

7.4.2 Study characteristics.

The study characteristics, including population, sample size, cardiac diagnosis, age, gender, and exercise test outcomes for the 6MWT, 10 m ISWT and other field tests are shown in Tables 7.4, 7.5 and 7.6. Sixty-two studies reported measurement properties of the 6MWT, 13 studies reported measurement properties of the 10 m ISWT and 15 of other field exercise tests. Cardiac failure was the most commonly studied condition, with 46 studies using patient groups with a primary diagnosis of heart failure, including five studies of cardiac transplant candidates and one study with patients following cardiac transplant. Thirty-two studies used patients with a primary diagnosis of coronary heart disease, and two studies using patients with permanent pacemakers. The quality review score sheets are shown in Appendix 9. The results of the measurement properties for each study are presented in Appendix 10. Agreement between the reviewers for methodological quality assessment was weighted κ .97, 95% CI [.92, 1.0].

Table 7.4

The 6MWT: Study Characteristics

Study (first author, date)	Diagnosis	Sample size (n)	Age Mean (SD)	Gender M/F (n)	6MWT distance (m) Mean (SD)	Measurement Property and Interpretability
Ades 2003	CHD	33	72 (6)	0/33	Baseline 357 (117); Follow-up-6/12 409 (116)	Construct validity
Allison 2004	CHD	83	72 (4)	57/26	Not reported	Construct validity
Araya-Ramírez 2010	CHD	425	62 (12)	293/132	Baseline-pre-CR 399 (87), range 61 to 608; 3/12-Follow-up 472 (97) M_{diff} 20 ± 16% Group ≤ 24 CR sessions: Baseline 411 (78); Follow-up 483 (90). Change range -54 to 53. Group ≥ 25 CR sessions: Baseline 381 (96); Follow-up 458 (104). Change range -145 to 191	Construct validity Responsiveness
Bajraktari 2011	CHF	77	60 (12)	51/26	Not reported	Construct validity
Baldasseroni 2014	CHD	148	70 (11)	129/19	455 (164)	Construct validity
Baptista 2012	CHD (cardiac surgery)	>350 m: 52 < 350 m: 35	59 (9) 61 (9)	43/9 15/20	248 (86) 436 (78)	Construct validity
Beatty 2012	CHD	556	68	479/77	Median 481 (104)	Criterion validity Construct validity
Bellet 2011	CHD	Baseline 61	62 (11)	45/16	Baseline 507 (85)	Relative reliability

Study (first author, date)	Diagnosis	Sample size (n)	Age Mean (SD)	Gender M/F (n)	6MWT distance (m) Mean (SD)	Measurement Property and Interpretability
		Completed 33	-	-	Baseline-1 521 (62); Baseline-2 533 (71); 6/52 Follow-up-1 554 (63) 6/52 Follow-up-2 570 (74); 6/12 Follow-up-1 564 (79); 6/12 Follow-up-2 581 (89)	Measurement error Responsiveness
Bittner 1993	CHF	833	60 (12)	650/183	374 (117)	Construct validity
Bittner 2000	CHD	156	60 (11)	107/49	Baseline 370 (103) n = 153	Construct validity
					17/52-Follow-up M_{diff} 54 (78)	Responsiveness
Cahalin 1996	CHF	RR 20	-	-	-	Relative reliability
		CV 45	49 (8)	40/5	310 (100)	Criterion validity
Carvalho 2011	CHF	16	58 (10)	12/4	Walk-1 545 (74); Walk-2 550 (66)	Relative reliability Criterion validity
Casillas 2015	CHD (Post CR)	148	59 (9)	133/15	536 (70)	Criterion validity
Cheetham 2005	CHF (Transplant candidates)	16	59 (3)	14/2	Baseline 458 (SE 21); 6/52 Follow-up 470 (SE 30); 12/52 Follow-up 457 (SE 28); 18/52 Follow-up 467 (SE 26)	Criterion validity Responsiveness
Chien 2011	CHF	Exp 24	Exp 57 (16)	Exp 20/4	Exp Baseline 424 (145); 8/52 Follow-up 433 (145)	Construct validity
		Con 27	Con 59 (16)	Con 18/9	Con Baseline 432 (81); 8/52 Follow-up 429 (93)	

Study (first author, date)	Diagnosis	Sample size (n)	Age Mean (SD)	Gender M/F (n)	6MWT distance (m) Mean (SD)	Measurement Property and Interpretability
Cipriano 2010	CHF (Transplant candidates)	12	52 (8)	10/2	399 (123)	Construct validity Criterion validity
Corvera-Tindel 2009	CHF	76 RR 7	63 (11) -	76/0 -	382 (73) -	Relative reliability Construct validity
De Sousa 2008	PPM	24	49 (14)	11/13	557 (95)	Criterion validity
Delahaye 1997	CHF	13	51 (12)	-	448 (118), range 300 to 640	Criterion validity Construct validity
Demers 2001	CHF	Baseline 768	63 (11)	637/41	Baseline 381 (84); 18/52 Follow-up 387 (94); 43/52 Follow-up 387 (103)	Relative reliability Construct validity Responsiveness
Doutreleau 2009	Heart Transplant	22	56 (2)	22/0	516 (13)	Criterion Validity
Flynn 2009	CHF	2331	59 (13)	1670/661	365 (105)	Construct validity
Forman 2012	CHF	CV 2030 HT 2054	- Median 59	- 1459/595	- Median 372 interquartile range 300 to 434	Criterion validity Construct validity
Gary 2004	CHF	Exp 16 Con 16	67 (11) 69 (11)	0/16 0/16	Exp Baseline 256 (112); 12/52-Follow-up 318 (97) Con Baseline 251 (112); 12/52-Follow-up 223 (124)	Responsiveness

Study (first author, date)	Diagnosis	Sample size (n)	Age Mean (SD)	Gender M/F (n)	6MWT distance (m) Mean (SD)	Measurement Property and Interpretability
Gayda 2004	CAD	25	60 (10)	22/3	-	Criterion validity
Green 2001	CHF (Transplant candidate)	7	57 (*SE* 2)	6/1	537 (30)	Criterion validity
Gremeaux 2009	CAD (ACS)	30 Subsample 10	52 (9) 50(9)	28/2 10/0	Baseline 490 (33); 6/52-Follow-up 552 (76) Baseline 489 (36); 6/52-Follow-up 560 (55)	Responsiveness
Gremeaux 2012	CAD (ACS)	30	52 (9)	28/2	Baseline 490 (33); >6/52-Follow-up 552 (76)	Construct validity Responsiveness
Gremeaux 2011	CAD (ACS)	81	58 (9)	77/4	-	Responsiveness Interpretability
Guazzi 2009	CHF	253	62 (10)	199/54	351 (93)	Criterion validity
Guyatt 1985	CHF	18	-	-	Encouraged: Walk-1 452; Walk-2 474; Walk-3 482; Walk-4 495; Walk-5 468; Walk-6 476 Unencouraged: Walk-1 400; Walk-2 408; Walk-3 415; Walk-4 408; Walk-5 428; Walk-6 415	Criterion validity
Hamilton 2000	CHD	94	63 (10)	61/33	Walk-1 521 (112); Walk-2 539 (114); Walk-3 550 (122)	Relative reliability Criterion validity Construct validity
Hanson 2012	CHD	30	63 (8)	24/6	Walk-1 444 (75); Walk-2 496 (68); Walk-3 529 (71)	Relative Reliability Measurement error

Study (first author, date)	Diagnosis	Sample size (n)	Age Mean (SD)	Gender M/F (n)	6MWT distance (m) Mean (SD)	Measurement Property and Interpretability
Ingle 2006	CHF	571	71 (10)	451/120	337 (103)	Construct validity
Ingle 2005	CHF	74	72 (7)	52/22	Baseline 285 (122); 1-year-Follow-up 276 (118)	Relative Reliability Measurement Error Responsiveness
Jehn 2009	CHF	40	61 (14)	29/7	523 (133)	Criterion Validity
Juenger 2002	CHF	205	54 (11)	173/32	470 (100)	Construct validity
Karapolat 2008	CHF (Transplant candidate)	123	42 (12)	100/23	321 (109)	Construct validity
Kavanaugh 1996	CHF	21	62 (6)	17/4	Not reported	Responsiveness
Kervio 2004	CHF	CHFD 12	64 (6)	9/3	Walk-1 445 (73); Walk-2 448 (71)	Relative reliability
		CHFP 12	66 (6)	10/2	Walk-1 407 (55); Walk-2 408 (59)	Measurement error Criterion validity
Kristjansdottir 2004	CHD (cardiac surgery)	10	76 (6)	-	541 (141) range 315 to 762	Criterion validity
Langenfeld 1990	PPM	97	67 (13)	55/42	461 (111) m	Criterion validity
		CV 42	-	-	-	
Lipkin 1986	CHF	10	54 (16)	7/3	Walk-3 506 (82)	Criterion validity
	CHD	12	59 (4)	10/2	Walk-3 386 (137)	

Study (first author, date)	Diagnosis	Sample size (n)	Age Mean (SD)	Gender M/F (n)	6MWT distance (m) Mean (SD)	Measurement Property and Interpretability
Lucas 1999	CHF	307	52 (13)	239/68	393 (104)	Criterion validity
Maldonado-Martin 2006	CHF	97	69 (6)	39/58	439 (111)	Criterion validity
		SHF 50	69 (5)	33/17	432 (115)	
		DHF 47	69 (6)	6/41	447 (109)	
Mandic	CHD	58	72 (6)	38/20	Walk-1 591 (131); Walk-2 611 (120)	Relative reliability
Walker 2013					Best walk 616 (129), range 377 to 1073	Criterion validity
Morales 1999	CHF	46	53 (10)	37/9	485 (91), range 250 t 670	Relative reliability
						Criterion validity
		RR 17			Walk-1 475; Walk-2 496; Walk-3 505	
Morales 2000	CHF	46	53 (10)	37/9	485 (91)	Criterion validity
Nogueira 2006	CHD (AMI)	25	54 (8)	23/2	Walk-1 521 (64); Walk-2 539 (70); Walk-3 545 (70)	Relative reliability Measurement error
O'Keefe 1998	CHF	60	81	22/38	239 (52)	Relative reliability Construct validity Responsiveness
Olper. 2011	CHD (cardiac surgery)	26 pre-CR 20 pre- & post- CR	54 (11)	19/7	Baseline Walk-1 460; Walk-2 500; Walk-3 505. 2/52-Follow-up Walk-1 525; Walk-2 550; Walk-3 545.	Relative reliability Measurement error Responsiveness
Opasich 2004	CHD (cardiac surgery)	2555	65 (10)	1881/674	296 (111), range 15 to 630	Construct validity

Study (first author, date)	Diagnosis	Sample size (n)	Age Mean (SD)	Gender M/F (n)	6MWT distance (m) Mean (SD)	Measurement Property and Interpretability
Opasich 1998	CHF	233	-	197/36	-	Measurement error
		Group A	52 (10)	-	Walk-1 401 (69); Walk -2 420 (71)	
		202	54 (7)	-	Walk -1 400 (70); Walk -2 415 (71)	
		Group B 31				
Opasich 2001	CHF	315	53 (9)	274/41	Walk -1 390 (88); Walk -2 408 (91)	Criterion validity
					mean scores range 134 to 686	
Peeters 1996	CHF	37	81 (6)	15/22	-	Criterion validity
		NYHA II:	84 (4)	8/8	353 (91)	
		16	79 (7)	3/7	172 (89)	
		NYHA III:	80 (6)	4/7	345 (94)	
		10				
		Controls:				
		11				
Pinna 2000	CHF	233	54 (0.6)	197/36	-	Relative reliability
		Group A	-	-	Baseline-Walk-1 394 (*SE* 6); Walk-2 413 (*SE* 6); 3/12-	Measurement error
		194			Follow-up-1 425 (*SE* 10) 3/12 Follow-up-2 438 (*SE*	
					11)	
					Walk-1 400 (*SE* 13) Walk-2 415 (*SE* 13)	
		Group B 30				
Pulz 2008	CHF	63	51 (10)	44/19	491 (94)	Measurement error
						Criterion validity

Study (first author, date)	Diagnosis	Sample size (n)	Age Mean (SD)	Gender M/F (n)	6MWT distance (m) Mean (SD)	Measurement Property and Interpretability
						Construct validity
Riley 1992	CHF	16	65	14/2	Walk-1 352 (81); Walk-2 375 (84); Walk-3 372 (87)	Measurement error
						Criterion validity
Verrill 2003	CHD	630	61 (10)	424/206	♀: Baseline 379 (92); 10-12/52-Follow-up 438 (91)	Construct validity
					♂: Baseline 446 (103); 10-12/52-Follow-up 513 (106)	Responsiveness
Westlake 2005	CHF	200	57 (12)	168/32	410 (92)	Construct validity
Wright 2001	CHD	Group-A	62 (9)	103/56	Baseline 315 (76); 6/52-Follow-up 377 (79) M_{diff} 62	Responsiveness
		(CR) 159			(53)	
		Group-B	61 (10)	22/8	Baseline 330 (92); 6/52-Follow-up 326 (92) M_{diff} -4	
		(no CR) 30			(19)	
Zugck 2000	CHF	113	54 (12)	90/23	466 (107), range 182 to 692	Relative reliability
						Criterion validity

Note. CHD = coronary heart disease; CR = cardiac rehabilitation; CHF = chronic heart failure; RR = relative reliability; CV = criterion validity; SE = standard error; Exp = experimental group; Con = control groups; PPM = permanent pacemaker; HT = hypothesis testing (as part of construct validity); IQR = interquartile range; ACS = acute coronary syndrome; CHFD = patients with chronic heart failure on optimal pharmacological management; CHFP = patients with chronic heart failure on optimal drug management plus pacing; SHF = systolic heart failure; DHF = diastolic heart failure; AMI = acute myocardial infarction; NYHA = New York Heart Association classification; ♀ = women; ♂ = men.

Table 7.5

The 10 m ISWT: Study Characteristics

First author, Date	Diagnosis	Sample size	Age mean (*SD*)	Gender	ISWT (m) Mean (*SD*) or [range]	Measurement Properties
Fowler 2005	CHD (cardiac surgery)	39	61 (9)	34/5	Walk-1 444 (135); Walk-2 486 (147); Walk-3 478 (141); range Walk-1-3 180 to 830; 8/52-Follow-up 568 M_{diff} 82, 95% CI [53, 110]	Relative reliability Measurement error Criterion validity Responsiveness
Green 2001	CHF (transplant candidates)	7 RR 7 CV	52 (8) 57 (5)	7/0 6/1	Walk-1 517 (*SE* 61); Walk-2 513 (*SE* 60)	Relative reliability Criterion validity
Hanson 2016	CHD	62	68 (10)	45/17	Walk-1 378 (173), range 30 to 760; Walk-2 395 (176), range 30 to 790	Relative reliability Measurement error
		10	67 (10)	5/5	Walk-1 378 (224), range, 30 to 740; Walk-2 393 (239) range 30-790; Walk-3 398 (229), range 40 to 770	
Houchen-Wolloff 2015	CHD	220	65 (111)	70/50	Walk-1 391 (173) (♀ 308, ♂ 415) 6/52-Follow-up 456 (187). M_{diff} 65, 95% CI [55, 75] For subsamples: Better (n = 105) 85, 95% CI [71, 99); Slightly Better (n = 52) 70, 95% CI [52, 89]; About the same (n = 54) 29, 95% CI [12, 47); Slightly worse (n = 6) -5; Worse (n = 3) –50	Responsiveness Interpretability

First author, Date	Diagnosis	Sample size	Age mean (*SD*)	Gender	ISWT (m) Mean (*SD*) or [range]	Measurement Properties
Jolly 2008	CHD	353	62 (10)	282 / 71	Walk-1 385 (152); Walk-2 415 (158)	Relative reliability Measurement error
Lewis 2001	CHF (Transplant candidates)	25	53 (8)	84% male	Walk-1 347 (140); Walk-2 400 (146); Walk-3 401 (147)	Relative reliability Criterion validity
Mandic Hodge 2013	CHD	34	70 (8)	27/7	Baseline 9 mins (1.9) (630 m), 1.5-year-Follow-up 8.8 mins (1.8) (610 m)	Construct validity
Mandic Walker 2013	CHD	58	72 (6)	38/20	Peak speed 6.5 (1.2) km/hr (520 < Distance ≤ 630)	Criterion validity
Morales 1999	CHF	46	53 (10)	80/20	503 (150), range 180 to 890	Relative reliability
		RR 17	-	-	Walk-1 445; Walk-2 478; Walk-3 485	Criterion validity
Morales 2000	CHF	46	53 (10)	37/9	503 (150)	Criterion validity
Pepera 2013	CHD	Group A 16	69 (9)	9/7	Walk-1 479 (139); Walk-2 499 (138), range 210 to 750	Construct validity
		Group B 113	69 (9)	82/31	-	
Pulz 2008	CHF	63	51 (10)	44/19	422 (119)	Measurement error Criterion validity

Note. CHD = coronary heart disease; CHF = chronic heart failure; RR = relative reliability; CV = criterion validity; SE = standard error; IQR = interquartile range; ♀ = women; ♂ = men.

Table 7.6

Other Field Exercise Tests: Study Characteristics

First author, date	Test administration	Diagnosis	Sample size (n)	Age M (*SD*)	Gender M/F (*n*)	Outcome	Measurement Properties
Alosco 2012	2 minute step test	CHF	145	69 (9)	89/56	59 (23) steps	Construct validity
Casillas 2015	200 m Fast walk test	CHD (Post CR)	148	59 (9)	133/15	109 (12) sec	Criterion validity
Cervi 2012	Treadmill 6MWT	CHD	42	60 (12)	32/10	Test-1 351 m (109) Test-2 382 m (102) Test-3 400 m (104) Mean Test 2 & 3 391 m (101) 2/52-Follow-up Test-1 469 m (104) Test-2 497 m (98) Test-3 506 m (101) Mean test 2 & 3 500 m (98)	Relative reliability Measurement error Interpretability
Chiaranda 2012	1 km treadmill walk	CHD	178	Range 38 to 83	178/0	Not receiving beta blockers 11 mins 15 secs (2:15) Receiving beta blockers 11 mins 10 secs (2:15)	Criterion validity
De Greef 2005	Modified GFE protocol	CHD (excl NYHA III/IV)	RR 46 CV 24	66 (7) 62 (9)	28/18 17/7	48 (23)	Relative reliability Measurement error Criterion validity

First author, date	Test administration	Diagnosis	Sample size (n)	Age M (*SD*)	Gender M/F (*n*)	Outcome	Measurement Properties
Delahaye 1997	Stair climbing (time to climb 46 steps)	CHF	13	51 (12)		40 sec (17), range 20 to 75	Criterion validity Hypothesis testing
Gayda 2003	20 m shuttle walk test	CHD	17	51 (9)	17/0	Relative $\dot{V}O_2$ 22.56 (6.29) Maximal ventilation (L/Min) 60.19 (15.34) Maximal heart rate 117 (22) Maximal speed 6.78 km/hr (1.23)	Criterion validity
Gremeaux 2009	200 m fast walk test	CHD (ACS)	30 7	52 (9) 50 (9)	28/2 10/0	Pre: 117 sec (3) Post: 107 sec (4)	Responsiveness
Gremeaux 2012	200 m fast walk test	CHD (ACS)	30	52 (9)	28/2	Test 1: 121 (15) Test 2: 117 (3) Post 107 (4)	Relative reliability Criterion validity Construct validity Responsiveness
Gremeaux 2011	200 m fast walk test	CHD (ACS)	81	58 (9)	77/4	Not reported	Interpretability
Guimarães 2008	Treadmill 6MWT	CHF	23	50 (9)	23/0	Walk-1 322 (48); Walk-2 322 (64)	Relative reliability
Houghton 2002	100 m fast walk test	CHF	36 CV 20	69	31/5	90 (*SEM* 4) sec	Criterion validity Construct validity
Meyer 2003	Treadmill 6MWT	CHF	51	59(11)	43/8	Baseline 262 (105); 12/52 Follow-up 409 (114)	Criterion validity Responsiveness

First author, date	Test administration	Diagnosis	Sample size (n)	Age M (*SD*)	Gender M/F (*n*)	Outcome	Measurement Properties
Nogueira 2010	Treadmill 6MWT	CHF	46	52 (9)	22/24	524 (73), range, 349 to 672	Construct validity ?others
Olper. 2011	Treadmill 6MWT	CAD (Cardiac Surgery)	26 pre-CR 20 pre- & post-CR	54 (11)	19/7	≈ Baseline-1 390; Baseline-2 420; Baseline-3 450 m. Follow-up-1 525; Follow-up-2 550; Follow-up-3 445	Relative reliability Measurement error Responsiveness Interpretability

Note. CHF = chronic heart failure; CHD = coronary heart disease; CR = cardiac rehabilitation; GFE: Groningen Fitness for the Elderly Test; NYHA = New York Heart Association; RR = relative reliability; CV = criterion validity; ACS = acute coronary syndrome; *SEM* = standard error of mean.

7.4.3 Best evidence synthesis of measurement properties of field exercise tests used in cardiac rehabilitation.

7.4.3.1 Measurement properties of the 6MWT.

The most commonly evaluated field exercise test was the 6MWT. No study followed standard operating procedures published in 2014 (Holland et al., 2014). The most common standardised protocols were the American Thoracic Association (2002) guidelines or the guidelines published by Guyatt et al. (1985). There was variability in the length of the track, with the shortest reported track length being 13 m (Doutreleau et al., 2009). The instructions and operating procedures for the 6MWT in the retrieved studies are summarised in Appendix 11. Instructions typically advised patients to walk as far as possible or as fast as possible. Two studies limited the intensity of the exercise test with the Borg rate of perceived exertion scale (Allison & Keller, 2004; Guimarães, Carvalho, & Bocchi, 2008). The provision of encouragement varied between studies. The distance walked in six minutes was reported as an outcome measure in all studies, typically measured by markers on the wall or using a tape measure to measure to the nearest metre or foot. One study reported using a pedometer reading to estimate distance walked (Guazzi et al., 2009). The number of tests completed in a single session varied. With the exception of studies reporting relative reliability and measurement error, the majority of studies reported a single 6MWT. When more than one test was performed in a single session, there was variability in the use of test results in the statistical analysis including use of the final test score, the best test score or the mean score. A number of studies did not specify which test result was reported or used in analysis (see Appendix 11).

7.4.3.1.1 Best evidence synthesis of measurement property: relative reliability.

There was moderate-level evidence synthesised from 16 studies supporting the relative reliability of the 6MWT in coronary heart disease and in chronic heart failure populations. For coronary heart disease, correlation coefficients ranged from .88 to .96 in the three studies with good methodological quality and from .88 to .97 in the three studies with fair methodological quality (Appendix 10). For chronic heart failure, the correlation coefficients for the two studies of good methodological quality was .96 and ranged from .80 to .99 in the studies with fair methodological quality (Appendix 10).

7.4.3.1.2 Best evidence synthesis of measurement property: measurement error.

There was moderate-level evidence synthesised from four studies that indicated the measurement error was large between the first and second walk and conflicting evidence from three studies for the adequacy of the measurement error between the second and third walk test in coronary heart disease. The synthesised evidence was inconclusive regarding the measurement error of the 6MWT in chronic heart failure. No study that reported on the measurement error also reported the MIC. However, the researchers provided justification of the adequacy or inadequacy of the size of the measurement error in the studies with a coronary heart disease population, but not chronic heart failure.

In coronary heart disease, the SDC between the first and second walk ranged from 71 m (Bellet et al., 2011) to 106 m (Hanson et al., 2012). The SDC ranged from 49 to 51 m between the second and third 6MWT for specific groups of patients with coronary heart disease, including those between four and seven days following a myocardial infarction, and within three weeks of cardiac surgery (Nogueira et al., 2006; Olper et al., 2011). However, when a mixed sample of patients who were commencing cardiac rehabilitation were assessed, the SDC was 95 m (Hanson et al., 2012). Nogueira et al. (2006) argued the SDC was acceptable if the MIC for respiratory patients (50 m) was accepted as the MIC for patients with coronary heart disease. Hanson et al. (2012) argued that 95 m was too high. Moderate-level evidence indicates that the measurement error of the 6MWT between the first and second walk in coronary heart disease populations is unacceptably large, but there is conflicting evidence between the second and third walk.

In chronic heart failure, the SDC ranged from 42 m when outliers were removed from the sample (Opasich et al., 1998; Pinna et al., 2000) to 47 (Pulz et al., 2008), to 145 m (Ingle et al., 2005). The authors did not discuss the size of the measurement error in relation to the MIC. The adequacy of the measurement error of the 6MWT in chronic heart failure populations remains unknown.

7.4.3.1.3 Best evidence synthesis of measurement property: criterion validity.

There was moderate-level evidence from three studies that did not support the criterion validity of the use of a single 6MWT in coronary heart disease. There was limited-level evidence synthesised from two studies that supported the evidence for

the criterion validity of the 6MWT in coronary heart disease when more than one 6MWT was performed, and the result of the best or final 6MWT was used in analysis. In chronic heart failure, there was strong-level evidence synthesised from nine studies that did not support the criterion validity of the 6MWT when a single 6MWT was performed, and conflicting evidence when more than one test was performed. In the post-heart transplant group, the evidence to support the criterion validity is unknown due to the low sample sizes (Doutreleau et al., 2009). There did not appear to be a relationship between treadmill criterion tests and bicycle criterion tests.

In coronary heart disease, three studies, one with good methodological quality (Casillas et al., 2015) and two with fair methodological quality (Beatty, Schiller, & Whooley, 2012; Gayda et al., 2004), demonstrated that the evidence for criterion validity was not supported when the distance walked from a single 6MWT was associated with peak oxygen uptake (r = .56) (Gayda et al., 2004), METS (r = .66) (Beatty et al., 2012) and maximum HR (r = .23) (Casillas et al., 2015). Two studies of good methodological quality provided support for criterion validity when the 6MWT was repeated, and the best distance of two (Mandic, Walker, et al., 2013) of four (Kristjánsdóttir et al., 2004) 6MWTs were associated with peak oxygen uptake (r = .72) (Mandic, Walker, et al., 2013) or maximal workload (r = .93) (Kristjánsdóttir et al., 2004). Kristjánsdóttir et al. (2004) reported that the fourth 6MWT was the best test for all participants in their sample. One study of poor methodological quality rejected support for the criterion validity of the 6MWT when the best test result of three 6MWTS was associated with retrospective results of a recent symptom-limited exercise test (r = .69) (Hamilton & Haennel, 2000). There appears to be a change in the evidence for criterion validity when more than one 6MWT is performed, the

evidence provides support of repeated testing in a single session and using the best test result or the results of a third test.

In chronic heart failure, when one 6MWT was performed, ten studies, one of excellent methodological quality (Forman et al., 2012), five of good methodological quality (Cahalin et al., 1996; Delahaye et al., 1997; Kervio, Ville, Leclercq, Daubert, & Carré, 2004b; Maldonado-Martín et al., 2006; Zugck et al., 2000) and four of fair methodological quality (Doutreleau et al., 2009; Green et al., 2001; Lucas et al., 1999; Peeters & Mets, 1996) demonstrated that the evidence for criterion validity was not supported when the distance walked from a single 6MWT was associated with peak oxygen uptake($.46 \leq r \leq .68$) (Cahalin et al., 1996; Delahaye et al., 1997; Doutreleau et al., 2009; Forman et al., 2012; Green et al., 2001; Kervio et al., 2004b; Lucas et al., 1999; Maldonado-Martín et al., 2006; Zugck et al., 2000), or treadmill distance (Kendall's Tau = .69) (Peeters & Mets, 1996). Two studies with good methodological quality (Jehn et al., 2009; Kervio et al., 2004b) and one of fair methodological quality (Cheetham et al., 2005) reported support for the evidence of the criterion validity when the distance of the first test was associated with peak oxygen uptake ($.77 \leq r \leq .88$). Jehn et al. (2009) reported a difference between patients with mild heart failure and no functional limitations and patients with moderate heat failure. Kervio et al. (2004b) reported support for the criterion validity in patients with chronic heart failure with optimal drug treatment (CHF_D), but not for patients with cardiac resynchronisation therapy (CHF_P). Cheetham et al. (2005) used a modified chronotropic protocol that was less aggressive than typical symptom-limited exercise tests such as the Bruce test. In chronic heart failure, when more than one 6MWT was completed, the results were conflicting. Two studies of good methodological quality

using the same data (Morales et al., 1999; Morales, Montemayor, & Martinez, 2000) and one of fair methodological quality (Guazzi et al., 2009) rejected support for the criterion validity when the second 6MWT distance was compared with peak oxygen uptake ($.68 \leq r \leq .69$). Two studies, one with good methodological quality (Zugck et al., 2000) and one with fair methodological quality (Carvalho et al., 2011) supported the evidence of criterion validity when the results of the second 6MWT were associated with peak oxygen uptake ($.70 \leq r \leq .71$), and two studies of good methodological quality provided support when the distance walked in the third 6MWT was associated with peak oxygen uptake ($.74 \leq r \leq .88$) (Riley, McParland, Stanford, & Nicholls, 1992; Zugck et al., 2000). Three studies, one with good methodological quality (Lipkin et al., 1986) and two with fair methodological quality (Guyatt et al., 1985; Opasich et al., 2001) did not support the evidence for the criterion validity when repeated 6MWTs were used and the mean score was associated with peak oxygen uptake ($.59 \leq r \leq .69$) (Lipkin et al., 1986; Opasich et al., 2001) or results of cycle ergometry ($r = .42$) (Guyatt et al., 1985).

In patients with a permanent pacemaker, two studies of fair methodological quality supported the evidence for the concurrent criterion validity of the 6MWT when using the peak watts ($r = .74$) (Langenfeld et al., 1990) or estimating peak oxygen uptake from an equation ($r = .71$) (de Sousa et al., 2008). There was fair evidence supporting the evidence for the concurrent criterion validity of the 6MWT in patients with permanent pacemaker.

The evidence for the predictive criterion validity of the 6MWT in coronary heart disease and chronic heart failure populations was inconclusive. There were three studies with sample sizes greater than 50 that reported equations using the 6MWT

distance to predict relative maximum oxygen uptake (Maldonado-Martín et al., 2006; Mandic, Walker, et al., 2013) and one study for absolute oxygen consumption (Guazzi et al., 2009). No study provided results for a validation study. The prediction equations differed and were inconsistent with each other.

7.4.3.1.4 Best evidence synthesis of measurement property: construct validity.

There was limited-level evidence synthesised in support of the evidence for the construct validity of the 6MWT in coronary heart disease in eight studies of fair (Allison & Keller, 2004; Hamilton & Haennel, 2000; Olper et al., 2011) or poor quality (Ades, Savage, Cress, Brochu, & Poehlman, 2003; Araya-Ramírez et al., 2010; Baptista et al., 2012; Bittner, Sanderson, Breland, Adams, & Schumann, 2000; Gremeaux et al., 2012). The results of four studies were indeterminate as they did not include measurements of constructs that related to physical fitness and functional capacity (Baldasseroni et al., 2014; Beatty et al., 2012; Opasich, De Feo, Pinna, & al, 2004; Verrill, Barton, Beasley, Lippard, & King, 2003). There was support for the evidence for the construct validity of the 6MWT in chronic heart failure in three studies of fair methodological quality (Corvera-Tindel, Doering, Roper, & Dracup, 2009; Flynn, Lin, et al., 2009; Juenger et al., 2002), and four studies of poor methodological quality (Delahaye et al., 1997; Guyatt et al., 1985; Karapolat et al., 2008; O'Keefe, Lye, Donnellan, & Carmichael, 1998). Three studies of poor methodological quality did not support the evidence for the construct validity of the 6MWT (Cipriano et al., 2010; Nogueira et al., 2010; Pulz et al., 2008) and 10 studies presented associations or differences from unrelated constructs without *a-priori* hypothesis testing and hence, the adequacy of the measurement property was rated

indeterminate (Bajraktari et al., 2011; Bittner et al., 1993; Chien, Lee, Wu, & Wu, 2011; Demers et al., 2001; Forman et al., 2012; Guazzi et al., 2009; Ingle et al., 2006; Riley et al., 1992; Westlake, Dracup, Fonarow, & Hamilton, 2005; Zugck et al., 2000). No study followed the recommendation by COSMIN by reporting the number of *a priori* hypotheses supported.

In coronary heart disease, there were moderate to strong associations between the 6MWT and results of alternative assessments from the same construct. There was a moderate negative association between the second 6MWT distance and the time to complete the 200 m fast walk test (r = -.57) (Gremeaux et al., 2012) and a positive association between the best distance from three 6MWTs and treadmill 6MWTs (r = .72) (Olper et al., 2011). In both studies, the associations were stronger after completion of a cardiac rehabilitation program.

In coronary heart disease, there were weak to moderate associations between the 6MWT distance and outcomes from assessments of somewhat related constructs (Ades et al., 2003; Allison & Keller, 2004; Araya-Ramírez et al., 2010; Baptista et al., 2012; Bittner et al., 2000; Hamilton & Haennel, 2000) including the 6MWT distance and self-reported measures of physical activity (standardised beta = .51, p < .0001; r = .30) (Allison & Keller, 2004; Bittner et al., 2000), components of quality of life (.22 ≤ r ≤ .62) (Ades et al., 2003; Araya-Ramírez et al., 2010; Baptista et al., 2012; Hamilton & Haennel, 2000) and self-efficacy scores (r = .44) (Allison & Keller, 2004). The association between the 6MWT distance and self-reported measures physical activity as well as self-efficacy strengthened after completion of a cardiac rehabilitation program (Allison & Keller, 2004).

In chronic heart failure, two studies of poor methodological quality reported associations and differences between the 6MWT and other field exercise tests (Delahaye et al., 1997; Pulz et al., 2008). The association between the 6MWT and 10 m ISWT was not reported, however, there were statistically significant differences between the distance walked in the 6MWT and the 10 m ISWT, and there were greater between-test differences for patients with more impairments (Pulz et al., 2008), suggesting that the 6MWT and 10 m ISWT may measure different constructs in cardiac rehabilitation. There was a strong association between the 6MWT and a stair climbing field test of physical fitness (r = -.82) (Delahaye et al., 1997). Unlike the 10 m ISWT, the stair-climbing test was not an incremental or progressive test.

In chronic heart failure, there were weak to moderate associations between the 6MWT distance and outcomes from assessments of somewhat related constructs of self-reported physical limitations and functional abilities (Corvera-Tindel et al., 2009; Flynn, Lin, et al., 2009; Guyatt et al., 1985; Juenger et al., 2002; O'Keefe et al., 1998). While the associations between the 6MWT and measures of physical fitness and functional capacity were weak to moderate, they were stronger than associations with unrelated constructs in three studies of fair methodological quality (Corvera-Tindel et al., 2009; Flynn, Lin, et al., 2009; Juenger et al., 2002). One study of poor methodological quality did not observe this finding (Nogueira et al., 2010). The results of all the associations, including those with indeterminate findings, are presented in Appendix 10.

In the group of patients with chronic heart failure awaiting heart transplant, two studies of poor methodological quality had conflicting results (Cipriano et al., 2010; Karapolat et al., 2008). Cipriano et al. (2010) reported ‘no association’ between

6MWT and a quality of life questionnaire. Karapolat et al. (2008) reported a moderate association between MOS SF36 scale score for Physical Function and the 6MWT distance (r = .48), these associations were stronger than others that did not relate to physical function (.07 ≤ r ≤ .35).

7.4.3.1.5 Best evidence synthesis of measurement property: responsiveness.

There was limited-level evidence synthesised from one study of fair methodological quality (Gremeaux et al., 2012) and three of poor methodological quality (Araya-Ramírez et al., 2010; Bittner et al., 2000; Wright et al., 2001), supporting the evidence for the responsiveness of the 6MWT in coronary heart disease populations. The results of one study of poor methodological quality did not support the responsiveness of the 6MWT (Verrill et al., 2003). In chronic heart failure, there was limited-level evidence synthesised from two studies of fair methodological quality (Gary et al., 2004; Ingle et al., 2005) and two of poor methodological quality (Kavanaugh, Meyers, Baigrie, & Al, 1996; O'Keefe et al., 1998) supporting the evidence for the responsiveness of the 6MWT. Due to the small sample size, the evidence for those waiting cardiac transplantation is unknown. No studies were found assessing responsiveness of the 6MWT in other specific cardiac populations.

In coronary heart disease, support for the responsiveness of the 6MWT was demonstrated with a positive association between the change in 6MWT distance and the change in the METs of a criterion measure (r = .59) (Gremeaux et al., 2011). The change in 6MWT distance was more strongly associated with self-reported measures of quality of life that related to physical function (r = .22) than other measures of

quality of life not directly related physical function (Araya-Ramírez et al., 2010; Bittner et al., 2000). One study of poor methodological quality reported the change in 6MWT was poorly associated with the change in quality of life measures ($r \leq .18$) (Verrill et al., 2003). Significant between-group differences in the 6MWT distance in those that attended a six-week cardiac rehabilitation program compared with those who did not attend were reported (Wright et al., 2001).

In chronic heart failure one study of fair methodological quality used a criterion approach to responsiveness and reported a weak association ($r = .08$) between the change in 6MWT and change in oxygen uptake in heart transplant candidates (Cheetham et al., 2005). In addition, associations between the change in 6MWT distance and changes in other somewhat related constructs were reported, including symptom severity ($r = -.75$) (Ingle et al., 2005) and measures of fatigue, dyspnoea, emotional function and mastery (Kavanaugh et al., 1996). Significant between-group differences in the 6MWT distance in those that attended a 12-week cardiac rehabilitation-walking program compared with those who attended education only sessions were reported (Gary et al., 2004)

Four studies reported only effect sizes of standardised response means in the absence of *a-priori* hypothesis testing and the results were indeterminate for responsiveness (Bellet et al., 2011; Demers et al., 2001; Gremeaux et al., 2009; Gremeaux et al., 2012). Effect sizes ranged from .30 to .65 (Bellet et al., 2011; Olper et al., 2011) in coronary heart disease and from -.02 (O'Keefe et al., 1998) to 2.13 (Demers et al., 2001) in chronic heart failure. A standardised response mean of 1.1 was reported in two studies using the same data (Gremeaux et al., 2009; Gremeaux et al., 2012). These findings were not linked with *a-priori* hypothesis testing.

7.4.3.2 Measurement properties of the 10 m ISWT.

The measurement properties of the 10 m ISWT were evaluated in 12 studies. The majority of studies followed the protocol described by Singh et al. (1992), with two (Mandic, Hodge, et al., 2013; Mandic, Walker, et al., 2013) studies following the protocol described by Tobin and Thow (1999). One study extended the test with the addition of three one-minute levels (Levels 13-15) using the same increments in speed per minute (Pepera, Cardoso, Taylor, Peristeropoulos, & Sandercock, 2013). The most common outcome of the test was the distance walked. One study reported the time at the end of the final completed shuttle as the main outcome measure (Mandic, Hodge, et al., 2013). With the exception of studies reporting relative reliability and measurement error, the majority of studies completed a single test. Two studies reported completing three tests in a single session (Fowler et al. 2005; Hanson et al. 2016).

7.4.3.2.1 Best evidence synthesis of measurement property: relative reliability.

In coronary heart disease, the relative reliability was supported by high-level evidence between the first and second walk and moderate-level evidence between the second and third walk. In chronic heart failure, there was no evidence between the first and second walk and limited-level evidence between the second and third walk.

For coronary heart disease, the correlation coefficients ranged from .94 (Fowler et al., 2005; Jolly et al., 2008) to .99 (Hanson et al., 2016) between the first and second walk and was .99 between the second and third walk (Fowler et al., 2005; Hanson et al.,

2016). For chronic heart failure, the correlation coefficient was .98 between the first and second 10 m ISWT in a group of seven patients with chronic heart failure awaiting transplant (Green et al., 2001). Two studies reported correlation coefficients of .90 (Lewis et al., 2001) and .99 (Morales et al., 1999).

7.4.3.2.2 Best evidence synthesis of measurement property: measurement error.

There was conflicting evidence for the adequacy of the measurement error between the first and second walk and moderate evidence in support of the adequacy of the measurement error between the second and third 10 m ISWT in coronary heart disease populations. There was no evidence in chronic heart failure populations to support or reject the adequacy of the measurement error. No study reported the MID. In the coronary heart disease population, the researchers provided justification on the adequacy or inadequacy of the size of the measurement error.

In coronary heart disease, the SDC when a single 10 m ISWT was completed ranged from 36 m (Jolly et al., 2008), to 54 (Hanson et al., 2016) to 122 m (Fowler et al., 2005). When the SDC was calculated from the *SEM*, it ranged from 33 to 47 m depending on whether the *SEM* was calculated for consistency or absolute agreement (Hanson et al., 2016). When a practice test was included, the SDC ranged from 21 m (Fowler et al., 2005) to 33 m (Hanson et al., 2016). Hanson et al. (2016) argued that the measurement error was acceptable if one walk was completed but acknowledged that the measurement error was smaller when two walks were completed and the first regarded as a practice test. Both Fowler et al. (2005) and Jolly et al. (2008) argued that two walks were required, with the first regarded as a practice test, these

arguments were based on statistically significant differences between the first and second walk ($p < .05$) and improvements in the repeatability coefficient when two tests were completed in a single session.

In chronic heart failure, the SDC was 53 m (Pulz et al., 2008). However, the researchers did not provide any further justification on the adequacy of the measurement error. The evidence for the measurement property of the 10 m ISWT in chronic heart failure remains unknown.

7.4.3.2.3 Best evidence synthesis of measurement property: criterion validity.

There was moderate-level evidence supporting the concurrent criterion validity of the 10 m ISWT in coronary heart disease and in chronic heart failure. In coronary heart disease, one study of excellent methodological quality and a sample size of 39 and one study of good methodological quality and a sample size of 58 provided support for the concurrent criterion validity of the 10 m ISWT when the distance walked was associated with relative peak oxygen uptake ($.72 \leq r \leq .87$) (Fowler et al., 2005; Mandic, Walker, et al., 2013). In chronic heart failure, four studies with good methodological quality (Lewis et al., 2001; Morales et al., 1999; Morales et al., 2000; Pulz et al., 2008), and two with fair methodological quality (Green et al., 2001), supported the concurrent criterion validity of the 10 m ISWT when the distance or number of shuttles was associated with relative peak oxygen uptake ($.78 \leq r \leq .79$).

The predictive criterion validity of the 10 m ISWT remains unknown. Fowler et al. (2005) reported that the relative peak oxygen uptake could be predicted from the 10 m

ISWT distance in coronary heart disease using the formula: $0.03 \times ISWT\ distance\ (m) + 7.81$. In chronic heart failure, the relative peak oxygen uptake being predicted by the distance walked in the 10 m ISWT with the following formulas: $0.283 \times ISWT\ distance\ (m) + 4.2355$ (Green et al., 2001), and $0.023 \times ISWT\ distance\ (m) + 5.9$ (Morales et al., 1999). Cut-off values to predict a relative peak oxygen uptake less than 14 ranged from 380 m (sensitivity 90%, specificity 87%) (Pulz et al., 2008) to 450 m (sensitivity 100%, specificity 89%) (Morales et al., 1999). Due to the small sample sizes and the absence of validation populations, the evidence to support the predictive criterion validity of the 10 m ISWT is unknown.

7.4.3.2.4 Best evidence synthesis of measurement property: construct validity.

There was no evidence to support the construct validity of the 10 m ISWT in cardiac rehabilitation populations. Three studies of poor methodological quality reported associations between the 10 m ISWT and other assessment outcome measures. In coronary heart disease, the distance walked in the 10 m ISWT was associated with self-reported physical activity at a cardiac rehabilitation follow-up assessment (r = .52) (Mandic, Hodge, et al., 2013) and step length and height (Pepera et al., 2013). However, given the poor methodological quality in these studies the evidence for the construct validity of the 10 m ISWT remains unknown.

7.4.3.2.5 Measurement property: responsiveness.

There was limited-level evidence supporting the responsiveness of the 10 m ISWT in coronary heart disease populations. One study demonstrated statistically significant

differences in the distance walked according to patient perception of change. Those who perceived themselves as better or slightly better improved more than those who perceived themselves as about the same, and these differences were significant. The difference between those who perceived themselves as better and slightly better was not significant. Two studies presented the effect sizes after a six-week cardiac rehabilitation program ranging from small (.38) (Houchen-Wolloff et al., 2015) to moderate (.55) (Fowler et al., 2005), however, these changes were not related to *a-priori* hypotheses. No studies were found investigating the responsiveness of the 10 m ISWT in chronic heart failure.

7.4.3.3 Measurement properties of alternative field exercise tests.

The measurement properties of less common alternative field exercise tests were retrieved. Self-paced tests included the treadmill 6MWT (Cervie, Olper, De Santi, & Pierini, 2012; Meyer & Laederach-Hofmann, 2003; Nogueira et al., 2010; Olper et al., 2011), the 200 m fast walk test (Casillas et al., 2015; Gremeaux et al., 2009; Gremeaux et al., 2012; Gremeaux et al., 2011), the 100 m fast walk test (Houghton, Harrison, Cowley, & al, 2002), the 1 km treadmill walk (Chiaranda et al., 2012), the modified Groningen Fitness for the elderly (de Greef et al., 2005) and the 2 minute step test (Alosco et al., 2012). Externally paced field tests included the 20 m shuttle walk test (Gayda, Choquet, Temfemo, & Ahmadi, 2003) and stair climb test (Delahaye et al., 1997).

7.4.3.3.1 Measurement properties of the treadmill 6MWT.

The measurement properties of the treadmill 6MWT were described in four studies; two studies using a coronary heart disease population (Cervie et al., 2012; Olper et al., 2011), and two using a chronic heart failure population (Guimarães et al., 2008; Meyer & Laederach-Hofmann, 2003). The test protocol was described in three of the four studies; the aim of the test was to walk as far as possible in six minutes on a treadmill. Participants were asked to hold the treadmill handrails and were allowed to increase or decrease the treadmill speed as desired (Cervie et al., 2012; Guimarães et al., 2008; Olper et al., 2011). In two studies, it was specified that participants were not to jog (Cervie et al., 2012; Olper et al., 2011). One study limited the intensity of the test to relatively easy to slightly tiring, or 11 to 13 on the Borg rate of perceived exertion scale (Guimarães et al., 2008). In coronary heart disease, the measurement properties investigated were relative reliability, measurement error, construct validity and responsiveness. In chronic heart failure, the measurement properties investigated were relative reliability, measurement error and criterion validity.

The relative reliability of the treadmill 6MWT in coronary heart disease was supported with limited-level evidence between the second and third test, however, between the first and second and in chronic heart failure populations, it remains unknown. In coronary heart disease, the reported *ICC*s between the second and third walk ranged from .93 (Cervie et al., 2012) to .97 (Olper et al., 2011). In chronic heart failure, one study with a sample size of 23 reported an *ICC* of .88 (Guimarães et al., 2008).

There was moderate-level evidence rejecting support for the measurement error between the second and third treadmill 6MWT at baseline in coronary heart disease. The measurement error of the treadmill 6MWT in coronary heart disease between the first and second test in coronary heart disease and in chronic heart failure populations remains unknown. Two studies reported the measurement error of the treadmill 6MWT in coronary heart disease. At baseline, between the second and third walk the SDC ranged from 64 m (Olper et al., 2011) to 77 m (Cervie et al., 2012). At the completion of cardiac rehabilitation, in a sample of 20 patients the SDC between the second and third walk was 50 m (Olper et al., 2011). The MIC was reported to be 54 m (Cervie et al., 2012). The MIC was less than the SDC between the second and third walk at baseline, but greater that the SDC between the second and third walk after a cardiac rehabilitation intervention. The measurement error of the 6MWT in chronic heart failure is unknown; one study presented Bland Altman plots but did not provide author justification of the size of the measurement error (Guimarães et al., 2008).

The evidence for both the criterion and construct validity of the treadmill 6MWT in coronary heart disease remains unknown. In chronic heart failure, there was limited-level evidence rejecting support for the criterion validity of the treadmill 6MWT and no evidence for the construct validity. In coronary heart disease, one study reported an association between the distance walked in the treadmill 6MWT and the corridor 6MWT at baseline (r = .72) and after a cardiac rehabilitation program (r = .67). In chronic heart failure, when the distance walked in the treadmill 6MWT was compared with the peak oxygen uptake expressed as a percentage of the predicted maximum oxygen uptake the association was .53 (Meyer & Laederach-Hofmann, 2003). No

study was retrieved that investigated the construct validity of the treadmill 6MWT in chronic heart failure.

The evidence for the responsiveness of the treadmill 6MWT in cardiac populations remains unknown. Three studies of poor methodological quality were retrieved. In coronary heart disease, there was no difference in the change in 6MWT according to the global rating of change anchors after a two-week intervention (Cervie et al., 2012). The effect size reported after a two-week intervention was .9 (Olper et al., 2011). In chronic heart failure, one study reported a positive association between change in distance and change in oxygen consumption at ventilatory threshold (Meyer & Laederach-Hofmann, 2003).

7.4.3.3.2 Measurement properties of the 200 m and 100 m fast walk test.

The measurement properties of the 200 m fast walk test were described in three studies using a coronary heart disease population (Casillas et al., 2015; Gremeaux et al., 2009; Gremeaux et al., 2012) and the measurement properties of the 100 m fast walk test were described in one study (Houghton et al., 2002). In the 200 m fast walk test patients were required to walk up and down a 50 m corridor as fast as possible without running with standard encouragement at mid distance, rest breaks were permitted and the time taken in seconds was the main outcome measure (Casillas et al., 2015; Gremeaux et al., 2009; Gremeaux et al., 2012). The protocol for the 100 m fast walk test was not described, but the main outcome measure was time taken (Houghton et al., 2002). An improvement in the test performance for both the 200 m fast walk test and the 100 m fast walk test was seen by a reduction in the time taken to

complete the test. The measurement properties that have been assessed in the 200 m fast walk test were relative reliability (Gremeaux et al., 2012), criterion validity (Casillas et al., 2015; Gremeaux et al., 2012), construct validity (Gremeaux et al., 2012) and responsiveness (Gremeaux et al., 2009; Gremeaux et al., 2012). The measurement properties that have been assessed for the 100 m fast walk test in chronic heart failure were criterion validity and construct validity (Houghton et al., 2002).

There was limited-level evidence from one study supporting the relative reliability of the 200 m fast walk test in coronary heart disease. The reported *ICC* between the first and second walk was .97. No studies were retrieved that investigated the measurement error of the 200 m fast walk test.

There was moderate-level evidence that rejected support for the criterion validity of the 200 m fast walk test in coronary heart disease and no evidence for the construct validity or the responsiveness of the 200 m fast walk test. There was a negative moderate association between the 200 m fast walk test time and the highest HR achieved in the final minute of the treadmill symptom-limited exercise test (Casillas et al., 2015). In one study of poor methodological quality, there was a moderate negative association between the 200 m fast walk test and the 6MWT (r = -.417) at the start of cardiac rehabilitation and a strong negative association between the 200 m fast walk test and the 6MWT following cardiac rehabilitation (r = -.566) as well as associations between the 200 m fast walk test and the MOS SF-36 physical component score (r = -.77) that were stronger than between the 200 m fast walk test and the MOS SF-36 mental component score (Gremeaux et al., 2012). For responsiveness, no study used the preferred method of using a criterion or setting *a-priori* hypotheses. Two studies,

using the same data, report a standardised response mean of 1.11 (Gremeaux et al., 2009; Gremeaux et al., 2012).

One study was retrieved that reported measurement properties of the 100 m fast walk test in a chronic heart failure population (Houghton et al., 2002). The criterion and construct validity were evaluated, however, due to the small sample size (n = 20) the evidence is unknown. The association between the treadmill time and the 100 m fast walk test time was -.64 and less than the recommended .70 required to support the measurement property. There was a weak negative association between the 100 m fast walk test and pedometer readings but a moderate negative association between the 100 m fast walk test and quality of life. The measurement properties of the 100 m fast walk test in cardiac populations remains unknown due to insufficient evidence.

7.4.3.3.3 Measurement properties of the 1 km treadmill test.

The measurement properties of the 1 km treadmill test were described in one study (Chiaranda et al., 2012). The test required patients to walk on the treadmill at a pace they believed they could maintain for 10 to 20 minutes at a moderate intensity. The test started when the pace was selected. Rate of perceived exertion was monitored every two minutes and walking speed adjusted accordingly. The time to complete 1 km was recorded as the outcome measure and an improvement in the test was seen by a reduction in the time to complete. The measurement property evaluated in the sample of 178 male patients with coronary heart disease was criterion validity.

There was limited evidence from one study supporting the evidence for the criterion validity in patients who received beta-blocker therapy but not patients who did not

receive beta-blocker therapy (Chiaranda et al., 2012). The average speed and maximum HR during the 1 km treadmill test along with BMI and age were used to predict peak oxygen uptake on the treadmill symptom-limited exercise test. The sample sizes were small. However, the testing of the equations on independent samples showed an adequate association ($r > .70$) for the group receiving beta blocker therapy but not for the group who did not receive beta blocker therapy ($r < .70$). There is limited evidence supporting the criterion validity of the 1 km treadmill test in patients with beta-blocker therapy, and limited evidence to reject support for the criterion validity of the 1 km treadmill test in patients not receiving beta-blocker therapy.

7.4.3.3.4 Measurement properties of the 2 minute step test.

The measurement properties of the 2 minute step test was described in one study using a chronic heart failure population (Alosco et al., 2012). For the step test, patients marched on the spot for two minutes lifting their knees to a marked target on the wall. Patients were allowed to use the back of a chair or wall for balance support. The main outcome measure was the number of times the right knee met the marker on the wall, an increase in this number represented an improvement in test performance (Alosco et al., 2012). The measurement property assessed was construct validity.

There was limited quality evidence from one study providing evidence to support the construct validity of the 2 minute step test in chronic heart failure (Alosco et al., 2012). There were weak negative associations between the 2 minute step test outcome and age ($r = -.20$) and depression score ($r = -.18$) as well as weak associations

between the 2 minute step test outcome and measures of cognitive function (.19 ≤ *r* ≤ .29). The results of the associations are summarised in Appendix 10.

7.4.3.3.5 Measurement properties of the stair climb test.

The measurement properties of a stair climbing test was described in one study using a coronary heart disease population (Delahaye et al., 1997). The test required patients to follow an investigator and ascend 46 steps. The time to climb the steps was the main outcome measure. The measurement properties assessed were criterion validity and construct validity. However, the evidence supporting the construct or criterion validity of the stair climbing test remains unknown due to the low sample size (*n* = 13).

7.4.3.3.6 Measurement properties of the modified Groningen Fitness for the Elderly walking test for cardiac disease.

The measurement properties of a modified Groningen Fitness for the Elderly walking test for cardiac disease was reported in one study retrieved on patients with coronary heart disease (de Greef et al., 2005). The original Groningen Fitness for the Elderly walking test protocol was modified to accommodate patients with coronary heart disease. A walking track that measured 16.7 by 8.3 m was set up, and participants were paced around the track by a series of beeps according to the following speeds: 1 minute at 4.5 km/hr, 1 minute at 5.0 km/hr, 1 minute at 5.5 km/hr, 2 minutes at 6 km/hr, 3 minutes at 6.5 km/hr and 4 minutes at 7km/hr (de Greef et al., 2005). Participants scored 1 point for each 16.67 m walked. The measurement properties assessed were relative reliability, measurement error and criterion validity.

There was limited evidence from one study supporting relative reliability. However, the evidence for measurement error and criterion validity remain unknown (de Greef et al., 2005). Relative reliability between the first and the second test was high, the *ICC* was .98 when all participants were included in the analysis, and .92 when the participants who completed the test (a 35% ceiling effect) were excluded from analysis. The 95% limits of agreement were presented; however, the authors did not provide justification on the size or adequacy of the measurement error. There was a positive association between the modified Groningen Fitness for the Elderly walking test outcome and the relative peak oxygen uptake of the symptom-limited bicycle test (r = .77), however, due to the low sample size the evidence supporting the criterion validity remains unknown.

7.4.3.3.7 Measurement properties of the 20 m shuttle walk test.

The measurement properties of a 20 m shuttle walk test was described in one study using a coronary heart disease population (Gayda et al., 2003). The participants were required to walk or run a 20 m course, the initial speed was 3 km/hr and there were 1 km/hr increments each minute until the patient could no longer maintain the required speed. The outcome measures assessed were maximum oxygen uptake, maximum ventilation, maximal heart rate and maximal speed. The measurement property assessed was criterion validity. The associations between the relative maximum oxygen uptake, maximal heart rate and speed were strong (r = .91, .80 and .89, respectively). The association between maximal ventilation during the shuttle and symptom-limited exercise test was moderate (r = .61). The evidence supporting the criterion validity of the 20 m shuttle walk test remains unknown due to the low sample size (n = 10).

7.4.3.4 Synthesis of results.

A best evidence synthesis was performed for each field exercise test according to patient group. Table 7.7 summarises the methodological quality of each study and the quality of the measurement properties assessed in the studies. The studies from the same patient group for each field exercise test were pooled and a level of evidence was applied as strong, moderate, limited, conflicting or synthesised and whether the quality of the measurement property was in support or not, or unknown, as can be viewed in the shaded rows titled *Best evidence synthesis*.

7.4.4 The interpretability of the field exercise tests.

Only tests with sufficient support for measurement properties were reviewed for interpretability, specifically to determine if the evidence supported the capacity of the field exercise test to detect important changes.

.

Table 7.7

Best Evidence Synthesis Using Methodological Quality and Quality Criteria

Study	Relative Reliability	Measurement Error	Criterion validity	Construct Validity	Responsiveness
6MWT: Coronary heart disease					
Ades 2003				Test-1 Poor (+)	
Allison 2004				Test-1 Fair (+)	
Araya-Ramírez 2010				Test-1 Poor (+)	Poor (+)
Baldasseroni 2014				Test-1 Fair (?)	
Baptista 2012				Test-1 Poor (+)	
Beatty 2012			Test-1 Fair (-)	Test-1 Fair (?)	
Bellet 2011	Test-1-2 Fair (+)	Test-1-2 Fair (-)			Poor (?)
Bittner 2000				Test-1 Poor (+)	Poor (+)
Casillas 2015			Test-1 Good (-)		
Gayda 2004			Test-1 Fair (-)		
Gremeaux 2009					Poor (?)
Gremeaux 2012				Test-2 Poor (+)	Fair (?)
Gremeaux 2011					Fair (+)
Hamilton 2000	Test-1-2-3 Fair (+)		Best Test (3) Poor (-)	Best Test (3) Fair (+)	
Hanson 2011	Test-1-2-3 Good (+)	Test-1-2 Good (-)			
		Test-2-3 Good (-)			
Kristjansdottir 2004			Test-4 (Best test)		
			Good (+)		

Study	Relative Reliability	Measurement Error	Criterion validity	Construct Validity	Responsiveness
Mandic Walker 2013	Test-1-2 Fair (+)		Best test (2) Good (+)		
Nogueira 2006	Test-1-2-3 Good (+)	Test-1-2 Good (-); Test-2-3 Good (+)			
Olper 2011	Test-2-3 Good (+)	Test-2-3 Good (+)		Best test (3) Fair (+)	Poor (?)
Opasich 2004				Test-1 Poor (?)	
Verrill 2003				Test-1 Poor (?)	Poor (-)
Wright 2001					Poor (+)
Best evidence synthesis	Test-1-2 ++ Test-2-3 ++	Test-1-2 -- Test-2-3 +/-	Test-1 -- Repeat testing +	+	+
6MWT: Cardiac heart failure					
Bajraktari 2011				Test-1 Fair (?)	
Bittner 1993				Test-1 Poor (?)	
Cahalin 1996	Test-1-2 Good (+)		Test-1 Good (-)		
Carvalho 2011	Test-1-2 Fair (+)		Test-2 Fair (+)		
Cheetham 2005			Test-1 Fair (+)		Fair (-)
Chien 2011				Test-1 Poor (?)	
Cipriano 2010				Test-1 Poor (-)	
Corvea-Tindel 2009	Test-1-2 Fair (+)			Test-1 Fair (+)	
Delahaye 1997			Test-1 Good (-)	Test-1 Poor (+)	
Demers 2001	Test-1-2 Fair (+)			Mean Test (2) Fair (?)	Poor (?)

Study	Relative Reliability	Measurement Error	Criterion validity	Construct Validity	Responsiveness
Doutreleau 2009			Test 1 Fair (-)		
Flynn 2009				Test-1 Fair (+)	
Forman 2012			Test-1 Excellent (-)	Test-1 Poor (?)	
Gary 2004					Fair (+)
Green 2001			Test-1 Fair (-)		
Guazzi 2009			Test-2 Fair (-)	Test-2 Poor (?)	
Guyatt 1985			Mean Test (6) Fair (-)	Mean Test (6) Poor (+)	
Ingle 2006				Test-1 Poor (?)	
Ingle 2005	Test-1-2 Fair (+)	Test-1-2 Fair (?)			Fair (+)
Jehn 2009			Test-1 Good (+)		
Juenger 2002				Test-1 Fair (+)	
Karapolat 2008				Test-1 Poor (+)	
Kavanaugh 1996					Poor (+)
Kervio 2004	Test-1-2 Fair (+)	Test-1-2 Poor (?)	Test-1 CHF_D Good (+) Test-1 CHF_P Good (-)		
Lipkin 1986			Test-3 Good (-)		
Lucas 1999			Test-1 Fair (-)		
Maldonado-Martin 2006			Test-1 Good (-)		

Study	Relative Reliability	Measurement Error	Criterion validity	Construct Validity	Responsiveness
Morales 1999	Test-1-2 Fair (+)		Test (? of 2) Good (-)^		
Morales 2000			Test-2 Good (-)^		
Nogueira 2010				Test-2 Poor (-)	
O'Keefe 1998	Test-1-2 Fair (+)			Test-1 Poor (+)	Poor (-)
Opasich 1998		Test-1-2 Fair (?)			
Opasich 2001			Mean Test (2) Fair (-)		
Peeters 1996			Test-1 Fair (-)		
Pinna 2000	Test-1-2 Fair (+)	Test-1-2 Fair (?)			
Pulz 2008		Test-1-2 Good (?)	Test (? of 2) Good (+)	Test-1-and-2 Poor (+)	
Riley 1992		Test-1-2: Poor (?)	Test-3 Good (+)	Test-3 Poor (?)	
Westlake 2005				Test-1 Fair (?)	
Zugck 2000	Test-1-2-3 Good (+)		Test-1 Good (-); Test-2 Good (+); Test-3 Good (+)	Test-1 Fair (?)	
Best evidence synthesis	Test-1-2 ++ Test-2-3 ?	Test-1-2 ? Test-2-3 ?	Test-1--- Repeat testing +/-	+	+
6MWT: Permanent pacemakers					
De Sousa 2008			Test-1 Fair (+)		
Langenfeld 1990			Test-1 Fair (+)		

Study	Relative Reliability	Measurement Error	Criterion validity	Construct Validity	Responsiveness
Best evidence synthesis			Test-1 +	+	
10 m ISWT: Coronary heart disease					
Fowler 2005	Test-1-2 Good (+)	Test-1-2 Good (-)	Test-1 Excellent (+)		Poor (?)
	Test-2-3 Good (+)	Test-2-3 Good (+)	Test-2 Excellent (+)		
			Test-3 Excellent (+)		
Hanson 2016	Test-1-2 Good (+)	Test-1-2 Good (+)			
	Test-2-3 Good (+)	Test-2-3 Good (+)			
Houchen-Wolloff 2015					Fair (+)
Jolly 2008	Test-1-2 Good (+)	Test-1-2 Good (-)			
Mandic Hodge 2013				Test-1 Poor (+)	
Mandic Walker 2013			Test-1 Good (+)		
Pepera 2013				Test-1 Poor (+)	
Best evidence synthesis	Test-1-2 +++	Test-1-2 +/-	Test-1 +++	?	+
	Test-2-3 ++	Test-2-3 ++	Test-2 +++		
			Test-3 +++		
10 m ISWT: Chronic heart failure					
Green 2001	Test-1-2 Fair (+)		Test-1 Fair (+)		
Lewis 2001	Test-2-3 Fair (+)		Test-2 Good (+)		
Morales 1999	Test-2-3 Fair (+)		Test-2 Good (+)		
Morales 2000			Test-2 Good (+)		
Pulz 2008		Test-1-2 Good (?)	Good (+)	Poor (+)	

Study	Relative Reliability	Measurement Error	Criterion validity	Construct Validity	Responsiveness
Best evidence synthesis	Test-1-2 ? Test-2-3 +	Test-1-2 ? Test-2-3 ?	Test-1+ Test-2 ++	?	?
Treadmill 6MWT: Coronary heart disease					
Cervie 2012	Test-2-3 Fair (+)	Test-2-3 Fair (-)			Poor (-)
Olper 2011	Test-2-3Good (+)	Baseline: Test-2-3 Good (-) Follow-up: Test-2-3 Good (+)		Poor (+)	Poor (?)
Best evidence synthesis	Test-1-2 ? Test-2-3 +	Baseline Test-1-2 ? Baseline Test-2-3 – Follow-up Test-1-2 ? Follow-up Test-2-3 ?	?	?	?
Treadmill 6MWT: Chronic heart failure					
Meyer 2003			Fair (-)		Poor (+)
Guimarães 2008	Test-1-2 Fair (+)	Test-1-2 Fair (?)			
Best evidence synthesis	Test-1-2 ?	Test-1-2 ?	Test-1-2 –	?	?
200 m fast walk test: Coronary heart disease					
Casillas 2015			Good (-)		
Gremeaux 2009					Poor (?)
Gremeaux 2012	Fair (+)		Fair (-)	Poor (+)	Fair (?)
Best evidence synthesis	+	?	--	?	?
100 m fast walk test: Chronic heart failure					

Study	Relative Reliability	Measurement Error	Criterion validity	Construct Validity	Responsiveness
Houghton 2002			Fair (?)	Poor (?)	
Best evidence synthesis	?	?	?	?	?
1 km treadmill walk: Coronary heart disease					
Chiaranda 2012			Beta blocked: Fair (+) Not beta blocked: Fair (-)		
Best evidence synthesis	?	?	?	?	?
2 min Step test: Chronic heart failure					
Alosco 2012				Fair (?)	
Best evidence synthesis	?	?	?	?	?
Stair climb: Chronic heart failure					
Delahaye 1997			Good (-)	Poor (+)	
Best evidence synthesis	?	?	?	?	?
Modified Groningen Fitness for the Elderly walking test: Coronary heart disease					
De Greef 2005	Test-1-2 Fair (+)	Test-1-2 Poor (?)	Fair (+)		
Best evidence synthesis	Test-1-2 +	?	?	?	?
20 m shuttle walk test: Coronary heart disease					
Gayda 2004			Fair (+)		
Best evidence synthesis	?	?	?	?	?

Note. +++ or --- high quality evidence to support or reject the measurement property; ++ or -- = moderate quality evidence to support or reject the measurement property; + or – limited quality evidence to support or reject the measurement property; +/- conflicting evidence; ? = indeterminate.

^ same data set

7.4.4.1 Interpretability of the 6MWT.

For coronary heart disease, 6MWT distances ranged from 15 m (Opasich et al., 2004) to 762 m (Kristjánsdóttir et al., 2004), and when allowed to run, 1073 m (Mandic, Walker, et al., 2013). For chronic heart failure, walk distances ranged from 134 (Opasich et al., 2001) to 692 m (Zugck et al., 2000). The greatest mean change in 6MWT distance occurred when there was at least 12 weeks of exercise intervention between the walk tests (Araya-Ramírez et al., 2010; Bittner et al., 2000; Verrill et al., 2003). One study reported a mean deterioration in walk test scores of 4 m when there was a six-week delay in the commencement of cardiac rehabilitation (Wright et al., 2001).

Floor and ceiling effects were not specifically discussed in the studies that were retrieved for the systematic review. One study reported 13 out of 58 elderly patients with CHD elected to jog during the 6MWT (Mandic, Walker, et al., 2013). These participants had a mean age of 72 years and had a higher level of aerobic fitness as measured by peak oxygen consumption (ml/kg/min) compared with the rest of the group (24.7 (4.5) vs 18.3 (4.4) $p < .001$). It is possible that the walk test represents insufficient change, and that this reflects a ceiling effect in the traditional 6MWT where participants are instructed not to jog.

One study reported the MIC of the 6MWT in patients with coronary heart disease who attended cardiac rehabilitation (Gremeaux et al., 2011). The results from an anchor-based approach using patients to rate their change and a distribution-based approach using the $SEM_{consistency}$ were consistent at 23 m when the global rating of change was measured during the cardiac rehabilitation program and the before and after cardiac

rehabilitation 6MWT distances were used. The authors also recommended that 25 m be considered the MIC to signify important change following a ROC analysis, at this level the positive predictive value was .9. When the therapists assessed the patient for meaningful change, the results varied, the MIC using an anchor-based approach was 15 m and when using a distribution-based approach was 36 m. The agreement between patient and physiotherapist was κ=.17.

7.4.4.2 Interpretability of the 10 m ISWT.

All participants were able to commence the 10 m ISWT. For the CHD patients, when a range was provided, the distance completed ranged from 30 m (Hanson et al., 2016) to 830 m (Fowler et al., 2005). The range of scores were not provided for the studies with participants with chronic heart failure. For patients who rated themselves as unchanged after a cardiac rehabilitation program, the mean improvement in the 10 m ISWT walk distance was 29 m, 95% CI [12, 47] (Houchen-Wolloff et al., 2015).

One study reported the MIC in patients with CHD across a six-week cardiac rehabilitation exercise intervention (Houchen-Wolloff et al., 2015). The results from an anchor-based approach, to rate themselves at least slightly better, participants needed to improve their walk distance by 70 m and to rate their improvement as better participants needed to improve their walk distance by 85 m. One hundred and seven of the 220 patients achieved an increase of 70 m, and this included 16 who rated themselves as about the same and three people who rated themselves as slightly worse or worse. Using a distribution-based approach, the MIC was 37 m with a small effect size of .38. Of the 220 patients, 147 had an effect size of at least .2. There was poor

agreement between the anchor-based and distribution-based methods (kappa κ = .002).

7.4.5 The preferred field exercise test for measurement of physical fitness and functional capacity in cardiac rehabilitation.

In general, the 6MWT and the 10 m ISWT were well-tolerated. There were no serious adverse responses to testing, with symptoms at the end of test including shortness of breath and fatigue, which typically resolved quickly. Most studies specified exclusion of patients who were unable to walk for any neurological or musculoskeletal reason and the use of a gait aid was not a reason for exclusion. Floor and ceiling effects were not commonly noted.

The 10 m ISWT appears to be a more suitable field exercise test for cardiac rehabilitation than other tests that have been evaluated in this systematic review. There is also evidence supporting some of the measurement properties for the 6MWT in specific cardiac populations. A summary of the measurement properties and interpretability of the 6MWT and the 10 m ISWT are shown in Tables 7.8 and 7.9.

Table 7.8

Summary of the Measurement Properties and Interpretability of the 6MWT and 10 m ISWT in Coronary Heart Disease

Measurement Properties	6MWT	10 m ISWT
Relative Reliability		
Test-1-2	Moderate Support	High Support
Test-2-3	Moderate Support	Moderate Support
Measurement Error		
Test-1-2	Moderate Reject	Conflicting
Test-2-3	Conflicting	Moderate Support
Criterion Validity		
Test-1	Moderate Reject	High Support
Repeat testing	Limited Support	High Support
Construct Validity	Limited Support	Unknown
Responsiveness	Limited Support	Limited Support
Interpretability	Floor: No	Floor: No
	Ceiling: Possibly	Ceiling: No
	MIC (ROC): 25 m	MIC (Distribution): 37 m
		MIC (Anchor): 70 m

Note. MIC = minimum important change; ROC = receiver operating curve.

For the 10 m ISWT, there is emerging support for the measurement properties when a single test is completed and stronger support for the measurement properties when two tests are completed. For the 6MWT, the measurement properties are not supported when a single test is completed but there is emerging support when two or more tests are completed and either the best test score or the final test score is used.

Table 7.9

Summary of the Measurement Properties and Interpretability of the 6MWT and 10 m ISWT in Chronic Heart Failure

Measurement Properties	6MWT	10 m ISWT
Relative Reliability		
Test-1-2	Moderate Support	Unknown
Test-2-3	Unknown	Limited Support
Measurement Error		
Test-1-2	Unknown	Unknown
Test-1-2	Unknown	Unknown
Criterion Validity		
Test-1-2	High Reject	Limited Support
Repeat Testing	Conflicting	Moderate Support
Construct Validity	Limited Support	Unknown
Responsiveness	Limited Support	Unknown
Interpretability	Floor: No	Floor: No
	Ceiling: Unknown	Ceiling: No
	MIC: Unknown	MIC: Unknown

Note. MIC = minimum important change.

7.5 Discussion

The 6MWT and 10 m ISWT have been widely used in cardiac rehabilitation to measure physical fitness and functional capacity. Both tests were completed safely in clinical or field environments with little equipment. No major adverse responses were recorded in any of the included studies. The 6MWT has been studied more extensively than other field exercise tests. The relatively recent development of the 10 m ISWT is a possible reason for the fewer studies retrieved in this systematic review, and this is consistent with findings in a respiratory population (Singh et al., 2014). The current systematic review also retrieved studies on other field exercise

tests, such as the treadmill 6MWT and 200 m fast walk test, but there was limited information on the measurement properties of these alternate field tests. The findings of this systematic review demonstrate there is sufficient evidence to support the 10 m ISWT as a field test in cardiac rehabilitation as well as the 6MWT in some cases and that there is no need to continue to develop alternative field exercise tests for this group of patients.

The measurement properties of the 6MWT and the 10 m ISWT appear to be reasonable, with stronger support emerging for the 10 m ISWT. The results demonstrate support for the measurement properties of the 10 m ISWT when one test is performed and stronger support when the test is repeated and either the best test or the results of the second test used. For the 6MWT, the results of the systematic review show that a minimum of two tests need to be completed and the results of the best test or the final test recorded. The results of this systematic review did not support recording the mean test score of the results. In respiratory disease, the current guidelines recommend that when more than one test is completed the best test score is used (Holland et al., 2014). The results of this systematic review supports the adoption of this practice in cardiac rehabilitation when more than one test is performed.

The evidence for the concurrent criterion validity was stronger for the 10 m ISWT than the 6MWT, which may reflect the similarities between the incremental nature of the 10 m ISWT and the symptom-limited exercise tests. In cardiac populations, the 10 m ISWT may be a better measure of physical fitness and functional capacity than the 6MWT, which may be limited by not allowing ambulation faster than walking. The submaximal self-paced nature of the 6MWT may limit the measurement of

physical fitness and functional capacity especially in the coronary heart disease population.

This systematic review retrieved two studies reporting the MIC in CHD populations, one for the 6MWT and one for the 10 m ISWT. The MIC and the interpretability of the two field exercise tests is not well understood. The studies provided information on the improvement that one group of patients perceives as important. For the 6MWT, 25 m was proposed as the MIC in patients with CHD. For the 10 m ISWT, the MIC ranged from 37 m using a distribution-based approach and 70 using an anchor-based approach for minimum change. A MIC of 25 m for the 6MWT has implications for the acceptable size of the measurement error, with the SDC of all studies retrieved in this systematic review greater than 25 m. This means that the smallest unit of real change that can be detected is greater than the important change identified.

This systematic review did not retrieve any studies reporting on the MIC for detecting worsening health as perceived by the patient or therapist. An important finding is that patients do not seem to rate small improvements in field exercise tests as improvement in their global rating of change (Houchen-Wolloff et al., 2015). This suggests that improvement needs to be relatively large before patients are satisfied that they have made meaningful change to their physical fitness and functional capacity.

The standardisation of the walking tests makes it more likely that test conditions will be similar when different staff, in longitudinal data collection, or across departments or organisations, repeat the 10 m ISWT. There was less variation in the administration of the 10 m ISWT compared with the 6MWT. All studies of the 10 m ISWT used the

recommended 10 m walking track, standard explanations and a pre-recording with beeps to control walking speed. However, in the 6MWT, there was considerable variation in track length, instructions varied regarding whether the patient was instructed to walk as far as possible or as fast as possible, the provision of encouragement was variable and there were attempts to control the intensity of the test in some cases. The large variation in the 6MWT procedures makes it more difficult to be confident of consistency in the test conditions. There was variation in the instructions of the 6MWT, whether patients were to walk as fast as they could or as far as they could as well as variations in the length of the walking track. In a respiratory population, the track design and length, provision of encouragement and whether the instructions advised patients to walk as fast or as far as possible affect test outcomes (Holland et al., 2014). The standardised test administration may in part explain the superior measurement properties of the 10 m ISWT compared with the 6MWT in this population.

The shorter walking track that is required for the 10 m ISWT, compared with the 6MWT, is a clinical advantage. A 10 m flat indoor walking track may be more practical and more easily accessed than a quiet 30 m walking track for the 6MWT. The ease of standardisation makes it more likely that the test conditions will be the same when the 10 m ISWT is repeated by different staff or for longitudinal data collection.

The potential for a ceiling effect was not specifically studied in the studies retrieved for this systematic review, however evidence of a ceiling effect was demonstrated in a group of patients who were permitted to jog during the 6MWT (Mandic, Walker, et

al., 2013), suggesting that a ceiling effect may be a problem in cardiac rehabilitation patients who are not limited by symptoms during walking.

The self-paced 6MWT walk test may be more appropriate and may perform as well as other field exercise tests in some groups but not others. It is possible that the test performs better in patients who are limited by symptoms at a consistent work rate, such as breathlessness in chronic heart failure populations. Evidence that those with greater disease severity perform less well in the 10 m ISWT than the 6MWT and the idea that the 10 m ISWT may underestimate functional capacity in these patients (Pulz et al., 2008) requires needs further investigation.

Field exercise tests offer a quicker and cheaper alternative to a symptom-limited exercise test conducted in a laboratory for assessment and monitoring of physical fitness and functional capacity. If patients are required to wait for a laboratory-based symptom-limited exercise test prior to commencing cardiac rehabilitation the cost of cardiac rehabilitation will increase and start times may be delayed. In clinical practice, it may be possible to complete only one 10 m ISWT. Despite an improvement in the measurement properties on repeated testing, the measurement properties from the studies retrieved in this systematic literature review for the 10 m ISWT were acceptable after one test and this is important consideration for the feasibility of the test.

7.5.1 Strengths and limitations.

The strengths of this systematic review are around the rigour in methods and the application. The systematic review was prospectively registered, and reported

according to PRISMA. It used rigorous methods based on COSMIN to evaluate measurement properties. There was good to excellent agreement between researchers on inclusion and exclusion, data extraction and quality assessment scoring. The results of this review were based on 16,529 participants in 78 studies, which increases the confidence in findings.

Limitations of this review include the restriction of abstracts and studies published in English language and limiting the review to adults with any form of cardiac disease. However, the large number of studies in included in this review with 78 studies evaluating 16,529 patients increases confidence in the findings. The measurement properties reported cannot be extended to populations beyond those included in this systematic review or to children. In addition, the role of the field exercise tests as diagnostic or prognostic indicators was not considered. An underlying assumption to this systematic review was that the measurement properties of the symptom-limited exercise test are acceptable.

Future research in this area should continue to assess the measurement properties of the 6MWT and the 10 m ISWT, in their standardised form, in specific cardiac rehabilitation populations, and using the recommendations from COSMIN. Research that uses the MIC to interpret the size of the measurement error will be important to decide if one test or two tests are required. The construct validity and responsiveness of both the tests needs further assessment with study designs that use *a-priori* hypothesis testing.

In conclusion, there is sufficient evidence to continue to support the use of the 6MWT and 10 m ISWT in cardiac rehabilitation with a preference for the 10 m ISWT. There

is also sufficient evidence to continue to support research investigating the measurement properties and interpretability of these tests as measures of physical fitness and functional capacity and there is no need to develop alternative field exercise tests for this population.

Chapter 8: Discussion and Future Directions

8.1 Chapter Aims

For low-risk patients entering a low-moderate intensity cardiac rehabilitation program, the 6MWT and the 10 m ISWT are two common field exercise tests proposed as alternatives to symptom-limited exercise tests (Goble & Worcester, 1999, p. 162; NSW Department of Health, 2006). The primary aim of this thesis was to determine the best field exercise test for assessing physical fitness and functional capacity in cardiac rehabilitation. In this final chapter, the results of the research are summarised followed by a discussion of the key issues arising from the main findings. Finally, the strengths and limitations of the research are outlined, and recommendations made for future research.

8.2 Summary of Main Findings

This thesis describes a series of studies that evaluated the measurement properties of the 6MWT and 10 m ISWT as a measure of physical fitness and functional capacity in cardiac rehabilitation, according to methodological recommendations from the COSMIN group (Chapters 2 to 7). This was followed by a systematic review and synthesis of the existing literature for the evidence for measurement properties of any field exercise test used in cardiac rehabilitation in people with cardiac disease (Chapter 8).

The results presented in Chapter 2 indicated that the 6MWT might not be an appropriate test to assess physical fitness and functional capacity in a cardiac rehabilitation population due to high systematic error and insufficient test retest reliability. The $ICC_{consistency}$ across three walk tests during a baseline cardiac rehabilitation assessment was high (.94), indicating that the between-patient variation remained stable. The systematic error, even after three walks remained large, as demonstrated by the moderate $ICC_{agreement}$ value (.66), and the large $SEM_{agreement}$ and wide 95% limits of agreement. The research presented in Chapter 2 of this thesis showed, for the first time in a cardiac rehabilitation population, the difference between the $ICC_{consistency}$ and $ICC_{agreement}$ values, and that interpretation of an *ICC* model based on the relative ranking of scores, without consideration of systematic change or absolute values, might be misleading. The additional information from $ICC_{agreement}$ demonstrates that it is an important consideration in relative reliability when the individual absolute scores of the test are used in clinical decision making, such as the use of the distance walked to guide exercise prescription or to measure true change after an intervention.

The results presented in Chapter 3 were in contrast to those presented in Chapter 2, demonstrating high test-retest reliability of the 10 m ISWT with only a small amount of systematic error when used in a mixed cardiac rehabilitation population when one or two tests were performed. The results demonstrated the size of measurement error between the first and second, and the second and third tests were likely to be acceptable, and change scores over cardiac rehabilitation were likely to be greater than measurement error. The results provided evidence to support the use of a single

10 m ISWT or two 10 m ISWTs with the first regarded as a practice test in cardiac rehabilitation.

In light of the support of the retest reliability of the 10 m ISWT, the evidence for the validity of the 10 m ISWT in a mixed cardiac rehabilitation population was investigated. Chapter 4 reported evidence that supports the concurrent criterion validity of the 10 m ISWT in cardiac rehabilitation when one or two tests were performed. The good association between the 10 m ISWT distance and time achieved on the symptom-limited exercise test suggest the 10 m ISWT may be a cost-effective measure of physical fitness and functional capacity in cardiac rehabilitation. The results supported the evidence for the concurrent construct validity when both one 10 m ISWT and two 10 m ISWTs were performed in a single session. Despite the good levels of association, the distance walked in either the first or the second 10 m ISWT was unable to predict, with confidence, the time achieved on the symptom-limited exercise test in an individual, due to relatively wide confidence bands. It is possible that increasing the sample size would improve the utility of the 10 m ISWT to predict the outcomes of the symptom-limited exercise test.

Chapter 5 reported on the construct validity and responsiveness of the 10 m ISWT in cardiac rehabilitation. For construct validity, six of the seven *a priori* hypotheses tested were at least partially accepted. Convergent themes emerged between the distance walked in the 10 m ISWT and alternate measures of physical fitness or functional capacity, including self-reporting of the amount of weekly exercise and physical functioning and reporting of functional activity. Divergent themes emerged between the 10 m ISWT and outcomes that did not relate to physical fitness or functional capacity including the poor association between the 10 m ISWT distance

and the MOS SF-36 scale scores of Social Functioning, Role Emotional and Mental Health. For responsiveness, all five *a-priori* hypotheses tested were at least partially accepted. The *a-priori* hypotheses tested demonstrated support for both the internal and the external responsiveness of the 10 m ISWT over an eight-week cardiac rehabilitation program. Support for responsiveness was present when both one test, and two tests, with the first regarded as a practice test were performed.

The results of the interpretability of the 10 m ISWT over a comprehensive cardiac rehabilitation program were presented in Chapter 6. The minimally important change (MIC) for patients to identify an improvement ranged from 71 m to 92 m when one test was performed, and 70 m to 85 m when two tests were performed, with the first regarded as a practice test. These findings provided further support for the adequacy of the measurement error of the 10 m ISWT, with the MIC being larger than the smallest detectable change when both one test and two tests were completed. This chapter also presented evidence that patients who made small gains in the distance walked in the 10 m ISWT may not perceive any improvement and may report a deterioration in their global rating of change. This was the second study to present information on the MIC of the 10 m ISWT in patients attending cardiac rehabilitation, and the results of this study were consistent with the anchor-based results of the original study (Houchen-Wolloff et al., 2015).

The systematic review reported in Chapter 7 demonstrated that the 6MWT and the 10 m ISWT field exercise tests were the two most commonly used field exercise tests in cardiac rehabilitation and that there was evidence to support at least some of the measurement properties of the two tests. There was stronger support for the 10 m ISWT across all the measurement properties. Differences in support of the

measurement properties between the 10 m ISWT and the 6MWT occurred when one or more than one field test was performed. For example, in chronic heart failure, there was moderate-level evidence to reject concurrent criterion validity when one 6MWT was performed and moderate-level evidence to support concurrent criterion validity when more than one 6MWT was performed. This variability in the summary of measurement properties was not as obvious for the 10 m ISWT when one or more tests were performed. The systematic review also identified further gaps in evidence for the measurement properties of field exercise tests in specific patient groups, such as chronic heart failure. Measurement error, predictive criterion validity and responsiveness of the 10 m ISWT in chronic heart failure remains unknown, as does predictive criterion validity of the 10 m ISWT in coronary heart disease. Therefore, the systematic review supported the conclusions from Chapters 2 through to 7 that the 10 m ISWT is, given the available evidence, the best field exercise test for use in cardiac rehabilitation.

8.3 Key Issues Arising from the Main Findings of the Thesis

8.3.1 The 10 m ISWT, compared with the 6MWT, had superior measurement properties in cardiac rehabilitation.

This thesis has demonstrated the 10 m ISWT has superior measurement properties than the 6MWT in cardiac rehabilitation when used as a measure of physical fitness and functional capacity. From the systematic review, where sufficient evidence was available, the measurement properties of the 10 m ISWT outperformed the 6MWT in both coronary heart disease and chronic heart failure populations. Studies of the measurement properties of alternative corridor or treadmill walk tests were retrieved,

but were much less common. This thesis emphasises that there is no need to continue developing alternative field exercise tests. Focus should be on the 10 m ISWT and perhaps on the 6MWT using a standardised protocol in specific cardiac rehabilitation populations, for example heart failure with moderate impairment (Williams & Naylor, 1992).

The 10 m ISWT demonstrated acceptable measurement properties without the need for practice tests, whereas the 6MWT did not demonstrate sufficient retest reliability when up to three tests were performed. The superior measurement properties of the 10 m ISWT compared with the 6MWT in cardiac rehabilitation may be, in part, due to capacity of the test to measure across the largely heterogeneous cardiac rehabilitation population. The test was suitable for a range of ambulant patients with varying levels of physical fitness and functional capacity without evidence of a floor or ceiling effects. The 6MWT was not a suitable test for all ambulant patients attending cardiac rehabilitation with evidence suggesting a ceiling effect in patients with milder impairment who were able to run during a 6MWT (Mandic, Walker, et al., 2013). It is possible that some patients reached their maximum walk speed and were unable to further increase their distance walked, an effect demonstrated in patients with mild pulmonary artery hypertension (Frost et al., 2005).

The support for the measurement properties of the 10 m ISWT over the 6MWT may have been due to the standardisation of the 10 m ISWT and reduced sources of error affecting reliability (Rothstein, 1985, p. 8). The three sources of error that can threaten test retest reliability of an exercise test in cardiac rehabilitation include inherent characteristics of the tests causing variability, a lack of consistency from the participants, and potential for errors made by the clinician supervising the test

(Rothstein, 1985, p. 8). The 10 m ISWT has limited test and operator variability compared with the 6MWT.

8.3.1.1 Test variability.

The 10 m ISWT is a standardised test, with little scope for test variability and interpretation of how the test should be performed compared with the 6MWT. The standard operating procedures for the 10 m ISWT clearly outline how the test is to be conducted, the instructions given to the patient and reasons for stopping the test (Singh et al., 1992). These standard operating procedures for the 10 m ISWT were adhered to throughout this thesis and each test was conducted under the same conditions. The studies retrieved in the systematic review also showed very little variation in the reported conduct of the 10 m ISWT with the majority of studies following the standard operating procedures of Singh et al. (1992). Two studies reported following the operating procedures of Tobin and Throw (1999) for the 10 m ISWT, however, the procedure described was consistent with that described by Singh et al. (1992). One variation was noted in one study that extended the 10 m ISWT with the addition of three one-minute levels (Levels 13-15) applying the same increments in speed per minute (Pepera et al., 2013).

The current standard operating procedures for the 6MWT were for testing of patients with respiratory conditions (Holland et al., 2014). The research presented in Chapter 2 preceded the release of these standard operating procedures, and followed the protocol described by Guyatt et al. (1985). The studies retrieved for the systematic review highlighted the variations in the conduct of the 6MWT. Variations included the track length or circuit, location of test such as indoor corridor or outdoor walking track,

instructions provided to the patient such as to walk as far as possible or as fast as possible, provision and frequency of encouragement and permission to run or not. The standard operating procedures (Holland et al., 2014) describe track length, and instruction, encouragement as key variables in the 6MWT, known to affect the distance walked.

There has been much variation in the track length of the 6MWT in the literature, despite a minimum of a 30 m track recommended to reduce the number of turns during the 6MWT (Holland et al., 2014). In this thesis, the 6MWT was conducted on a quiet 20 m indoor corridor track with markings on the wall at regular intervals. Space was a constraint at the testing hospital; the hospital did not have a 30 m indoor walking track available for field exercise tests where participants would not be interrupted by other hallway traffic. While there was local standardisation with all the 6MWT conducted on the 20 m track, there may be some variability when comparing to other published literature. To limit variation in 6MWT outcomes due to corridor length and frequency of turning, researchers have recommended that the clinician consider alternative tests such as the 10 m ISWT if a minimum track length of 30 m is not available (Singh, Spruit, Troosters, & Holland, 2015). Strong recommendations were made against amending the 6MWT through variations in track length (Holland et al., 2014). The requirement for a quiet 30 m corridor that allows uninterrupted performance has been described as a cumbersome component (Stevens et al., 1999) and may make this test unsuitable in many clinical environments.

The instructions given prior to the 10 m ISWT were standardised and played to the patient via a recording, this eliminated variability in description and instruction on completing the test (Singh et al., 1992). This means that all participants across all the

research sites received the same instructions in exactly the same manner. In contrast, the clinician supervising the test, often read from a script with the possibility of between-site variations, provided the description and instruction for the 6MWT. In respiratory groups, the instruction to walk as far as possible yields a different result to the instruction to walk as fast as possible (Holland et al., 2014). The instruction, *to cover as much ground as possible*, was used consistently in this thesis and is the preferred instruction (American Thoracic Society, 2002; Guyatt et al., 1985; Singh et al., 1992). It is possible that a pre-recording for the 6MWT similar to the 10 m ISWT would overcome the variability seen in the instructions given.

No encouragement was provided during the 10 m ISWT. Participants were advised of the increase in level of the 10 m ISWT and therefore increase in speed with a triple beep from the audio recording. In contrast, the provision of encouragement and time elapsed in 6MWT protocols varied from no encouragement (Guyatt et al., 1984; Guyatt et al., 1985), to every second minute (American Thoracic Society, 2002), to every minute (Holland et al., 2014) or every 30 seconds (Guyatt et al., 1984; Guyatt et al., 1985). In this thesis, encouragement was provided in a calm and consistent manner with standardised statements at 30-second intervals. Encouragement has been shown to improve walk distance in the 6MWT (Guyatt et al., 1984). The provision of encouragement at 30-second intervals was recommended by Guyatt et al. (1984) and later proposed to limit the variability associated with motivation (Wu et al., 2003). Therefore, it is possible that the standardised operating procedures minimise test and operator variability and therefore measurement error for the 10 m ISWT in cardiac rehabilitation. In contrast, test and operator variability are likely contributors to measurement error in the administration of the 6MWT.

8.3.1.2 Patient variability.

Patients completing the 6MWT showed variations in the distance walked over repeated testing even when testing conditions were standardised within a study. The research in this thesis demonstrated that patients referred to cardiac rehabilitation continued to increase the distance walked on repeated 6MWTs. This variability on repeated testing was not demonstrated in the 10 m ISWT. This difference may be attributable to patient variability. Intrinsic motivation or a learning effect of a submaximal self-paced walk test is a factor in the 6MWT that may not affect externally paced incremental tests, such as the 10 m ISWT. Motivation, patient attitudes and beliefs of exercise tolerance may affect performance in submaximal self-paced exercise testing. A study using the 12 minute walk test in patients with chronic bronchitis demonstrated that physical performance was affected not only by mood such as anxiety, hostility and depressive symptoms, but also by attitudes and beliefs towards oneself, fear of exercise and self-perceived estimation of severity of exercise limitation (Morgan, Peck, Buchanan, & McHardy, 1983). The authors hypothesised that these findings may be similar in conditions like angina pectoris where physiological factors interact with psychological factors (Morgan et al., 1983). A similar variation was observed in healthy adults in an international multicentre study of the 6MWT, where variations in distance walked could not be attributable to anthropometric or physiological factors alone, and the researchers proposed that habitual walk speed, cultural influences, and factors relating to lifestyle, mood, attitude and motivation needed to be considered (Casanova et al., 2011). No studies were retrieved that directly investigated this link in cardiac disease.

A learning effect of the 6MWT could explain the systematic error in repeated testing seen in this thesis, and has been attributed to the variation in other groups of patients with heart disease (Guyatt et al., 1985; Hamilton & Haennel, 2000). This is consistent with the learning effect of the 6MWT demonstrated in healthy individuals (Gibbons, Fruchter, Sloan, & Robert, 2001), "walk naïve" healthy individuals (Wu et al., 2003), and in a 5 minute walk test for people with respiratory conditions (Knox, Morrison, & Muers, 1988). Wu et al. (2003) and Gibbons et al. (2001) demonstrated a learning effect in healthy individuals over three successive 6MWTs and that the learning effect was still present when participants were retested two months later. In respiratory disease, the learning effect was observed during repeated 5 minute self-paced walk tests, with the greatest variation occurring over the first three tests, but variations continued to be observed when up to 12 tests were repeated over three days. In this thesis, systematic change in 6MWT distance was observed when up to three 6MWTs were performed (Chapter 2). Like the participants in the study by Wu et al (2003), it is possible that at commencement of cardiac rehabilitation participants are walk naïve. Furthermore, the patients participating in the research in this thesis did not experience symptoms limiting submaximal exercise test performance during a submaximal walk test, so with each repeated test the participants gained more confidence to cover further ground.

8.3.1.3 Operator variability.

The operator variability in administration of the 6MWT and 10 m ISWT are not well studied. Potential variables from the operator include inaccuracies in test set up such as whether the track length was accurately measured each time; accuracy with timing the test and recording laps or shuttles. There is more opportunity for variation in

instruction and encouragement provided by the operator in the 6MWT. In this thesis, every effort was made to ensure the test was repeated under the same conditions. The track length was measured and the position of the cones was checked prior to every test. Instructions and encouragement for the 6MWT were read from a script in a calm and consistent voice. Therefore, in this thesis it is not thought that operator variability was an important factor contributing to measurement error for the 6MWT and the 10 m ISWT.

8.3.2 The purpose of the exercise tests in cardiac rehabilitation.

Despite some limitations, this thesis has provided evidence to support the 10 m ISWT as a test of physical fitness and functional capacity in patients attending cardiac rehabilitation. From a clinician perspective, the field test outcome of the 10 m ISWT provides information on baseline level of physical fitness and functional capacity; a guide to base treatment decisions and most importantly; a tool to measure and monitor change over time. 10 m ISWT outcomes may be used in goal setting, such setting a goal to improve the distance walked as an absolute value or percentage of the baseline walk distance, or as reassurance to the patient or family about physical fitness and functional capacity at a single point in time or as a measure of responsiveness (American Association of Cardiovascular and Pulmonary Rehabilitation, 2013, p. 61). When used to measure physical fitness and functional capacity the results of this thesis supported the discriminative and evaluative role of the 10 m ISWT but not the predictive role in cardiac rehabilitation (Kirshner & Guyatt, 1985). It is possible that there was insufficient data to determine the predictive utility of the 10 m ISWT in cardiac rehabilitation.

The 10 m ISWT distance walked may have a discriminative function in cardiac rehabilitation, to distinguish levels of physical fitness and functional capacity between patients attending a program (Guyatt et al., 1992b). The results support the discriminative function of the 10 m ISWT with evidence of a high level of test retest reliability when measured for consistency and absolute agreement. Additionally, in patients commencing cardiac rehabilitation, the results of *a-priori* hypothesis testing support the evidence for construct validity. The 10 m ISWT is able to rank patients in order of best test performance to worst test performance (Kirshner & Guyatt, 1985), and this is associated with patient reporting of exercise activities. Clinicians can confidently discriminate between individuals or groups of patients with higher or lower levels of physical fitness and functional capacity using the 10 m ISWT (Keating & Matyas, 1998). This may be useful to describe an individual or group in cardiac rehabilitation in terms of physical fitness and functional capacity at a single time. The results may also guide clinical decision-making for an appropriate exercise intervention, but this thesis did not address this possible clinical application.

It is possible that there was insufficient data to support the 10 m ISWT as a predictive measurement tool. In this thesis, the 10 m ISWT was unable to predict accurately the results of a symptom-limited exercise test. While a strong association was observed between the distances walked in the 10 m ISWT and the symptom-limited exercise test time, the accuracy of an individual 10 m ISWT distance to predict the duration of the symptom-limited exercise test was limited by wide 95% confidence intervals. The systematic review showed variation in the prediction equations that used the 10 m ISWT to predict outcomes of laboratory-based exercise tests including the symptom-limited exercise test and the cardiopulmonary exercise test. The current evidence does

not support the usefulness of the 10 m ISWT as a prediction measurement tool for an individual and the results are inconclusive for a group.

The evaluative purpose of the 10 m ISWT was supported by this thesis for both the group and individual (Kirshner & Guyatt, 1985) and as such, the test is suitable for monitoring the effects of interventions for groups, or for monitoring the progress of an individual. In this thesis, emerging evidence supported both the longitudinal reliability and validity, and the interpretability of the MIC of the 10 m ISWT. The evidence to support the longitudinal validity or responsiveness was demonstrated using *a-priori* hypothesis testing. The *a-priori* hypotheses tested demonstrated support for both the internal and the external responsiveness of the 10 m ISWT over an eight-week cardiac rehabilitation program. These findings were supported by previously reported data demonstrating support for the longitudinal reliability and validity of the test (Pepera et al., 2010). Support for the responsiveness was present when both one test, and two tests, with the first regarded as a practice test, were performed.

8.3.3 Clinical and theoretical implications.

8.3.3.1 Clinical implications for cardiac rehabilitation programs and clinicians.

This thesis demonstrates that the 10 m ISWT should be the field test of choice for measuring physical fitness and functional capacity in low-moderate risk patients commencing cardiac rehabilitation. The 10 m ISWT outcome is suitable to describe patients in cardiac rehabilitation at a single time and is a responsive measure of change in cardiac rehabilitation. The measurement properties of the test support the

discriminative function of the test. This thesis did not investigate the use of test outcome or cut-off values for individual exercise prescription.

The 10 m ISWT is a suitable alternative to the symptom-limited exercise tests as a measure of physical fitness and functional capacity. The results suggest that the two exercise tests measure similar constructs related to physical fitness and functional capacity. The advantage of the 10 m ISWT is that one test can be completed along a 10 m track by staff within the program with results immediately available. The time and cost associated with completing this field test, within the admission and discharge cardiac rehabilitation assessment, is minimal and includes a seated rest prior to the test, the test duration of up to 12 minutes and after-test monitoring until symptoms have returned to normal. In Australia, completion of the symptom-limited exercise test in a cardiology laboratory requires referral from a medical doctor. The time between identification of need, referral, appointment for testing, reporting and the information made available to the cardiac rehabilitation program in Australia is likely to be at least two weeks. These additional processes place increased burden on the patient, and in non-metropolitan settings can involve considerable travel. This thesis supports the use of the 10 m ISWT in assessing physical fitness and functional capacity in low-moderate risk patients attending a mixed cardiac rehabilitation population. For assessment of physical fitness and functional capacity in low-moderate risk patients with a field test such as the 10 m ISWT, there appears to be a reduction in health care associated costs without compromise to patient care (Porter, 2010).

The individual 10 m ISWT distance cannot predict individual results of symptom-limited exercise test or cardiopulmonary exercise test with accuracy. It is likely that

the purpose of measurement of the 10 m ISWT differs from the symptom-limited exercise test. The symptom-limited exercise test has discriminative, predictive and evaluative functions (Kirshner & Guyatt, 1985). The symptom-limited exercise test is used to discriminate between people with and without cardiac disease, level of physical fitness and level of impairment; predict risk for future events and prognosis; and evaluate disease progression and treatment effect (American Thoracic Society & American College of Chest Physicians, 2003; Bruce & Hornsten, 1969; Jelinek & Lown, 1974; Myers et al., 1998; Pichurko, 2012; Stelken et al., 1996). The role of the 10 m ISWT in measuring constructs other than physical fitness and functional capacity was not considered in this thesis. Therefore, if diagnosis and prognosis are required for a high-risk patient with cardiac disease, for example, it is likely that a symptom-limited exercise test or cardiopulmonary exercise test would be required.

It was beyond the scope of this thesis to conduct a feasibility study. However, preliminary evidence supports the feasibility of the 10 m ISWT in cardiac rehabilitation. The results of this thesis support the feasibility of the 10 m ISWT in terms of acceptability, practicality, integration, and limited efficacy testing (Bowen et al., 2009). *Acceptability* was not specifically addressed in this thesis, however, previous research reports that patients prefer corridor walk tests over treadmill tests, largely due to fear of falling and treadmill speed (Peeters & Mets, 1996). The 10 m ISWT is a *practical* test, completed within the clinical environment with minimal equipment and space under the supervision of cardiac rehabilitation staff such as physiotherapists and exercise physiologists. The equipment required is simple and accessible, including chairs, cones for a turning circle, stopwatch, and equipment for assessment such as stethoscope, sphygmomanometer, pulse oximeter, heart rate

monitor, and an audio player with 10 m ISWT recording. The 10 m ISWT requires a 10 m corridor to conduct the test, much less than the minimum 30 m corridor requirement of the 6MWT. The physical requirements for the 10 m ISWT are likely to be suitable for a wide range of cardiac rehabilitation programs, from large metropolitan, well-resourced and staffed programs, to smaller programs conducted in rural communities. It is likely that the completion of a single 10 m ISWT can be *integrated* into existing cardiac rehabilitation assessments with minimal increases in costs. The maximum 10 m ISWT duration is 12 minutes with additional time for rest and monitoring before and after the test. This time is likely to be acceptable to the clinician, the patient and the cardiac rehabilitation coordinators. The research in this thesis has demonstrated the *efficacy* of the 10 m ISWT in cardiac rehabilitation through the adequacy of the measurement properties when a single test is performed, and support for the discriminative and evaluative functions of the test.

8.3.3.2 Recommendations for field exercise tests in clinical practice guidelines for cardiac rehabilitation.

This thesis supports the 10 m ISWT as the preferred field walk test when measuring physical fitness and functional capacity of patients considered low-intermediate risk entering a low-moderate intensity cardiac rehabilitation exercise program. Further, the results support the addition of the 10 m ISWT as an alternative to the symptom-limited exercise test for measurement of physical fitness and functional capacity in cardiac rehabilitation guidelines. The symptom-limited exercise test, supervised by a cardiologist in a laboratory or within cardiac rehabilitation, is reported as the preferred test at the start and end of cardiac rehabilitation (Balady et al., 2007; Dafoe, Huston, Wong, & Stokes, 2003; Japanese Circulation Society Joint Working Group,

2014; Pavy et al., 2012; Pavy, Iliou, Meurin, Tabet, & Corone, 2006; Piepoli et al., 2014; Piepoli et al., 2010; Price et al., 2016). When the symptom-limited exercise test is not available, the 6MWT is a commonly reported alternative (Price et al., 2016). For example, the American Association of Cardiovascular and Pulmonary Rehabilitation (2016) strongly recommend the use of a symptom-limited exercise test prior to commencing cardiac rehabilitation programs and if not possible, then the 6MWT is a recommended alternative. There are a small number of cardiac rehabilitation programs around the world that include the 10 m ISWT as an alternative to symptom-limited exercise testing in cardiac rehabilitation (Audelin, Savage, & Ades, 2008; British Association for Cardiovascular Prevention and Rehabilitation, 2012; McCreery et al., 2013; Scottish Intercollegiate Guidelines Network, 2002). The cumulative evidence provided by this thesis suggests that the use of the 10 m ISWT should be more widely incorporated into practice guidelines as the best field exercise test in cardiac rehabilitation. Additionally, if the purpose of the test is to evaluate physical fitness and functional capacity, the 10 m ISWT can be confidently recommended in guidelines as a cost effective alternative to symptom-limited exercise testing for many patients referred to cardiac rehabilitation.

8.4 Strength and Limitations

A strength of this research is the rigour applied to the methods for the empirical studies and the systematic review. Both the empirical studies and the systematic review followed the framework recommended by COSMIN group (Mokkink et al., 2010b, 2010c). The COSMIN recommendations are a result of a four-round Delphi study and provided an acceptable methodological framework for evaluating measurement properties. Rigour in the statistical methodology was applied, with

recommendations from the COSMIN group used, and where possible, the outcomes were investigated using alternative statistical methods to increase confidence in the findings.

Additional methodological strategies were put in place to ensure adequate rigour of the procedures for testing. This included completion of every exercise test under standard conditions. Every test was conducted in temperature and humidity controlled environments, in quiet, flat, indoor walking tracks. The distance of the track was measured and checked prior to every test. Instructions and encouragement for the 6MWT were read from a script using consistent and calm speech. For the 10 m ISWT, all instructions were played from a recording. The rest breaks before and after the tests were timed and the monitoring equipment was regularly checked by the hospital engineering and maintenance department. All participants were assessed on their usual medications and reassessed at the same time of day.

The methodological design also considered the generalisability of the results. This is seen in the statistical analysis and the procedures of the empirical studies. Appropriate statistics such as the model of ICC were used to ensure generalisability of results beyond this thesis. All ambulant patients who were referred to cardiac rehabilitation were considered for inclusion in this study. This aimed to reflect the heterogeneity of the cardiac rehabilitation population and improve the generalisability of the results beyond specific conditions of patients with heart disease, taking a program perspective. Participants were recruited from both public and private cardiac rehabilitation programs, meaning that there was socioeconomic diversity between participants. It is possible to generalise the results across socioeconomic groups. The cardiac rehabilitation programs in Bendigo follow mainstream cardiac rehabilitation

programs that are typical in Australian cardiac rehabilitation programs and adhere to the practice guideline for cardiac rehabilitation and secondary prevention (Goble & Worcester, 1999). The results are likely to be generalised to the wider cardiac rehabilitation population in Australia and perhaps internationally. The fact that the findings of the systematic review (Chapter 7) reinforced the findings of the empirical studies (Chapters 2 to 6) provide further support for the generalisability of the findings.

The main limitations of this thesis were the relatively small sample sizes in the empirical studies according to the COSMIN recommendations. In each chapter reporting the empirical studies, the sample size was justified according to the type of study (Howell, 2012; Walter et al., 1998). In Chapter 4, the sample size may have been inadequate for stable prediction equations and in Chapter 6, there were not enough participants who reported a deterioration in their global rating of change for physical fitness and functional capacity to measure with confidence the minimum important change for those patients who reported deterioration.

This thesis only investigated the measurement properties of the field exercise tests as a measure of physical fitness and functional capacity, using functional outcomes readily available in cardiac rehabilitation environments. No attempt was made to investigate the role of the field exercise tests as diagnostic or prognostic markers. However, the investigated measurement properties were consistent with the intended purpose of field exercise tests in practice.

One underlying assumption in this thesis was that criterion tests have acceptable measurement properties. However, there is very limited research on the measurement

properties of symptom-limited exercise tests. For example, in as much as the 10 m ISWT was not able to predict accurately time in the criterion measure (Chapter 4), this could have been contributed to by error from the criterion measure as well as error from the 10 m ISWT, or a small sample size.

Finally, one researcher completed all the 6MWTs and 10 m ISWTs carried out in this thesis. While the researcher was blind to the result of the symptom-limited exercise test, the researcher was not blind to the results of the cardiac rehabilitation assessment data in the other research projects (Chapter 2, 3, 5 and 6). All possible attempts were taken to maximise internal consistency of the test performance. The tests were set up in the same manner, at each site using the same corridor for testing, instructions and encouragement where provided was read from a script and reason for stopping the test was also consistent across the tests. The corridor length for the 6MWT was less than the new current recommendations of 30 m (Holland et al., 2014), while this is known to reduce the total distance walked in the test in respiratory patients, it is unknown if it affected the observed systematic error.

8.5 Future Research

This thesis has generated ideas for future research. The broad areas of future research that arise from this thesis are: (1) testing the clinical application of the 10 m ISWT in guiding prescription of an endurance walking program and to guide clinicians in implementing an appropriate level exercise program within cardiac rehabilitation; and (2) continued investigation on the understanding of the MIC for improvement and deterioration in cardiac rehabilitation to improve the interpretability of the test.

Further research is required to investigate the use of the 10 m ISWT outcomes in the implementation of individual exercise programs. Using the individual distance walked in the 10 m ISWT to guide exercise training would be helpful for the clinician. This practical application includes the use of the 10 m ISWT in setting a walking program dose during cardiac rehabilitation. This thesis has demonstrated patient variability in self-selection of walk speed. Selecting a percentage of the peak 10 m ISWT may be a way of overcoming this variability. The role of the 10 m ISWT in assisting prescription of a walking exercise program has been demonstrated in patients with chronic obstructive pulmonary disease as an intervention: the percentage of peak speed calculated from the 10 m ISWT has demonstrated appropriate physiological training responses (Zainuldin, Mackey, & Alison, 2012), and also as a field exercise test of endurance, called the endurance shuttle walk test (Holland et al., 2014; Revill, Morgan, Singh, Williams, & Hardman, 1999; Singh et al., 2014). Further exploring the practical application in cardiac rehabilitation as a means of standardising the exercise intensity and removing the effect of patient variability may improve patient outcomes.

The use of the 10 m ISWT may also have a role of guiding the clinician in prescription of general exercise programs, apart from walking programs, in cardiac rehabilitation, but this has not been investigated. This thesis has demonstrated support for the discriminative function of the 10 m ISWT; specifically that it is able to rank patients in order of low physical fitness and functional capacity to high physical fitness and functional capacity. The hypothesis that the peak results of the 10 m ISWT can effectively guide the implementation of a general exercise program for cardiac rehabilitation has not been tested. It is not known if the individual score or cut-off

values can be used to guide the optimal level of exercise training in cardiac rehabilitation and if this would improve patient outcomes.

The second main area of research that could be generated by this thesis concerns the continued investigation of the interpretability of the test over a cardiac rehabilitation program. The research in this thesis was the second to present information on the MIC of the 10 m ISWT over a cardiac rehabilitation, with similar results to a published study (Houchen-Wolloff et al., 2015). In this thesis, the MIC was calculated three ways; one method used predictive modelling, an innovative approach not previously applied in cardiac rehabilitation. The low numbers in the group of patients reporting a deterioration in both this thesis and the study by Houchen-Wolloff et al. (2015) has made it difficult to determine the MIC for deterioration. This thesis was the first to present results of predictive modelling in understanding the likelihood of important change over the duration of a cardiac rehabilitation program. However, it would be useful to use this model of calculating important change in a larger sample for those patients who report an improvement in physical fitness and for those who report a deterioration in physical fitness and functional capacity. The ability to evaluate 10 m ISWT test performance associated with deterioration would enable clinicians to identify at-risk patients attending cardiac rehabilitation, enabling appropriate referral and investigation.

8.6 Conclusions

The preferred field exercise test for measurement of physical fitness and functional capacity for patients with low-intermediate risk attending a low-moderate intensity exercise program is the 10 m ISWT. The results demonstrated support for the

measurement properties of the 10 m ISWT across a mixed cardiac rehabilitation population when a single test was performed. The discriminative and evaluative functions of the 10 m ISWT in cardiac rehabilitation were also supported. The clinician can use the test to provide information on the physical fitness and functional capacity of an individual or group at a single time, as a tool for ranking level of physical fitness and functional capacity of patients, as a goal-setting tool in treatment planning and as a responsive measure to monitor change over an eight-week cardiac rehabilitation program. The cumulative evidence suggests it is time to stop developing alternative field walk tests for cardiac rehabilitation. Future research should continue to focus on the development of the 10 m ISWT as a measure of physical fitness and functional capacity in cardiac rehabilitation, and how the test can be used in exercise prescription in cardiac rehabilitation.

Appendix 1: Publications

Reprinted here with permission are the three publications that have arisen from research in this thesis.

Hanson, L. C., McBurney, H., & Taylor, N. F. (2012). The retest reliability of the six minute walk test in patients referred to a cardiac rehabilitation programme. *Physiotherapy Research International, 17*, 55-61. doi:10.1002/pri.513

Reprinted from Physiotherapy Research International, 17 , LC Hanson, H McBurney, NF Taylor, The retest reliability of the six minute walk test in patients referred to a cardiac rehabilitation programme, 55-61, Copright (2017), with permission from John Wiley and Sons.

Hanson, L. C., Taylor, N. F., & McBurney, H. (2016). The 10 m incremental shuttle walk test is a highly reliable field exercise test for patients referred to cardiac rehabilitation: a retest reliability study. *Physiotherapy, 102*(3), 243-248. doi:10.1016/j.physio.2015.08.004

Reprinted from *Physiotherapy*, 10(2), LC Hanson, NF Taylor, and H McBurney, The 10 m incremental shuttle walk test is a highly reliable field exercise test for patients referred to cardiac rehabilitation: a retest reliability study, 243-248, Copyright (2017), with permission from Elsevier.

Hanson, L. C., McBurney, H. & Taylor, N. F. (2017). Is the 10 m incremental shuttle walk test a useful test of exercise capacity for patients referred to cardiac rehabilitation? *European Journal of Cardiovascular Nursing,* 1474515117721129.

RESEARCH ARTICLE

The Retest Reliability of the Six-minute Walk Test in Patients Referred to a Cardiac Rehabilitation Programme

Lisa C. Hanson[1,*], Helen McBurney[2] & Nicholas F. Taylor[3]

[1]La Trobe Rural Health School, La Trobe University, Bendigo, Victoria, Australia
[2]School of Primary Health Care, Monash University, Peninsula, Victoria, Australia
[3]Department of Physiotherapy, La Trobe University, Bundoora, Victoria, Australia

Abstract

Background and Purpose. The purpose of this paper was to determine if the Six-minute Walk Test (6MWT) was a reliable exercise test for patients referred to cardiac rehabilitation when up to three tests were performed and to determine if test scores differed according to between-test time interval. **Methods.** Thirty adults aged 63 ± 7.9 years referred to cardiac rehabilitation participated in a repeated measures reliability trial. Participants completed three 6MWTs within a one-week period. Participants were randomly allocated to one of three groups: on the first day, Group A completed three walks, Group B completed two walks and Group C completed one walk. Relative reliability was expressed in a ratio ($ICC_{2,1}$), and absolute reliability was expressed in metres (95% confidence intervals) for group and individuals. **Results.** The 6MWT demonstrated a high level of relative reliability (intraclass correlation coefficients [ICC] = 0.94) across the three walks. There was no statistically significant difference between the test scores of the three groups. However, there was an increase in distance walked from the first to the second to the third 6MWT. Absolute reliability indicated that a change of at least 44 m would be required to be interpreted as true change in a group, and at least 95 m to be interpreted as true change in an individual with 95% confidence. **Conclusion.** Three 6MWTs completed in relatively short timeframes were not sufficient for reliable results as there was an increase in the distance walked, and relatively large increases in distances would be required to be interpreted as change. It did not make any difference whether the tests were all completed on one day or over one week. This study highlighted problems that may arise when relying on reliability coefficients alone to interpret reliability. These results suggest that the 6MWT may not have sufficient reliability to be a suitable test to evaluate exercise tolerance in patients referred to cardiac rehabilitation.

Received 24 November 2010; Revised 9 March 2011; Accepted 23 March 2011

Keywords
exercise test; heart disease; reproducibility of results

***Correspondence**
Lisa Hanson, Department of Allied Health, La Trobe Rural Health School, La Trobe University, PO Box 199, BENDIGO, Victoria, 3552, Australia.
Email: L.Hanson@latrobe.edu.au

Published online 13 July 2011 in Wiley Online Library (wileyonlinelibrary.com) DOI: 10.1002/pri.513

Introduction

The Six-minute Walk Test (6MWT) is a field exercise test that can be used both to assess the exercise tolerance of individual patients who attend a cardiac rehabilitation programme and as a group or programme outcome measure. The test was originally described by Butland et al. (1982) as an alternative measure of exercise

tolerance to the 12-minute walk test for people with pulmonary disease. It has since been used in many other clinical settings including cardiac rehabilitation (Guyatt et al., 1985; Harada et al., 1999; Hamilton & Haennel, 2000; Zugck et al., 2000). The main outcome of the test is the distance walked over six minutes; an improvement in the test score is measured by an increase in the distance walked. When the test is performed more than once, the best score is recorded. It is widely accepted that in a pulmonary population, a practice test is required for a reliable and best test score (Troosters et al., 1999). However, research investigating test–retest reliability of the 6MWT when used in a general cardiac rehabilitation population remains limited, and the results and recommendations vary.

A number of researchers who have assessed the test–retest reliability of the 6MWT have reported increases in the scores during repeated testing of patients with cardiac diseases (Lipkin et al., 1986; Hamilton & Haennel, 2000; Gayda et al., 2004; Kristjánsdóttir et al., 2004; Patrick, 2005; Nogueira et al., 2006). The time elapsed between testing ranged from same-day testing (Cahalin et al., 1996; Kervio et al., 2004; Kristjánsdóttir et al., 2004; Patrick, 2005; Nogueira et al., 2006) to up to eight weeks (O'Keefe et al., 1998; Patrick, 2005). Four studies reported that when tests are repeated there are significant increases in the distances walked between the first and second walks (Hamilton & Haennel, 2000; Gayda et al., 2004; Patrick, 2005; Nogueira et al., 2006). Furthermore, Hamilton and Haennel (2000) reported a significant increase in the distance between the second and third tests. Patrick (2005), who studied a population with chronic heart failure, and Nogueira et al. (2006), who studied patients one week after myocardial infarction, reported changes between the first and second tests but no significant difference between the distances walked in the second and third tests. Kristjánsdóttir et al. (2004) and Lipkin et al. (1986) also reported the final test score as the best but did not provide between-test comparisons. Wu et al. (2003) attributed changes in test scores to a learning effect.

In contrast, three studies have reported it unnecessary to repeat the 6MWT in specific cardiac populations including chronic heart failure and advanced symptomatic heart disease (Cahalin et al., 1996; O'Keefe et al., 1998; Kervio et al., 2004). Kervio et al. (2004) reported no significant difference in the mean distance walked in two 6MWTs. Participants in this study completed an additional practice 6MWT. Cahalin et al. (1996) and O'Keefe et al. (1998) each reported that 6MWT scores were reliable after one test. Cahalin et al. (1996), in studying the reliability of the 6MWT in 20 participants with advanced heart failure, had participants complete two 6MWTs on the same day. O'Keefe et al. (1998), in studying elderly patients with heart failure, reported that 24 patients experienced no change in cardiac status over a period of three to eight weeks. They completed two 6MWTs, one at the start and one at the end of this period. In both studies, reporting was limited to measures of relative reliability, with Cahalin et al. (1996) reporting an intraclass correlation coefficient (ICC) of 0.96 and O'Keefe et al. (1998) an ICC of 0.91

Patrick (2005) discussed the time constraints of repeated testing in a clinical setting. The 6MWT has been shown to be well tolerated in a variety of cardiac populations during repeated same-day testing (Cahalin et al., 1996; Kervio et al., 2004; Kristjánsdóttir et al., 2004; Patrick, 2005; Nogueira et al., 2006), with the test being safely completed three or four times in one day (Kristjánsdóttir et al., 2004; Patrick, 2005; Nogueira et al., 2006). Patrick (2005) suggested that the time required to complete three or four tests may not be practical in the clinical setting.

It remains unclear how many 6MWTs are required to yield the best test score and whether the between-test time affects the test score. A further unknown is whether there are differences in test scores according to whether test repetition occurs on the same day or over days within a one-week period. The aims of this research were the following: first, to determine if the 6MWT is a reliable field exercise test in the general cardiac rehabilitation population when up to three tests are performed and, second, to determine if the time between testing affects test scores.

Methods

Participants

All adults with coronary artery disease referred to cardiac rehabilitation irrespective of severity or duration of the condition were eligible to participate in this study. Participants were excluded if they had any condition where exercise would be contraindicated or if they were unable to walk for any neurological or musculoskeletal reason. In addition, participants were excluded if limitations in English language production or comprehension skills precluded them from understanding the consent form.

Procedure

University and Hospital Human Ethics Committees approved this study, and written informed consent was obtained from each participant. Testing took place in the Physiotherapy Department at the participating hospital. Thirty participants were randomly allocated by an independent staff member to one of three groups, labelled Group A, Group B and Group C. All participants completed three 6MWTs on admission to a cardiac rehabilitation programme and prior to the commencement of an exercise intervention. Group A participants ($n = 10$) completed three 6MWTs on the same day; Group B participants ($n = 10$) completed two 6MWTs on the first assessment and a third test within one week prior to commencement of cardiac rehabilitation; Group C participants ($n = 10$) completed one 6MWT on the first assessment and a further two tests within a one-week period, prior to commencement of cardiac rehabilitation. No change in management or exercise intervention occurred between testing.

The 6MWT was completed in an indoor flat 20-m corridor. Cones were placed 0.5 m in from both ends of the 20-m circuit to allow for a turning circle. Instructions and encouragement for the tests were standardized according to Guyatt et al. (1984). Participants were allowed a 20-min rest break prior to commencing the first test of the session, and when more than one 6MWT was completed on the same day, participants were given a 20-min rest break between tests (Steele, 1996). The distances walked in the three tests were recorded to the nearest metre.

For safety, heart rate and oxygen saturation were monitored throughout the test, and blood pressure was monitored before and after each test. Participants were not able to commence the 6MWT if resting systolic blood pressure was greater than or equal to 200 mmHg or resting diastolic blood pressure was greater than or equal to 110 mmHg (American College of Sports Medicine, 2010). Criteria for early termination of the test included patient distress, dizziness, angina or onset of severe musculoskeletal pain (Cahalin et al., 1996).

Statistical methods

Participant characteristics were recorded (age, gender, cardiac intervention and other relevant history) to allow description of the study sample. Where relevant, means were expressed followed by standard deviation. Retest reliability was expressed in two ways: (a) as relative reliability, as a ratio in the form of the intraclass correlation ($ICC_{2,1}$) (Shrout & Fleiss, 1979) and (b) as absolute reliability, in metres, the unit of measurement using 95% confidence intervals.

The ICC model (2,1) gives a relative index of the variance between subjects to the variance between subjects plus error variance, as well as accounting for variance between test and retest scores. The ICC was interpreted with the following guidelines: good reliability was a score greater than 0.75, moderate reliability between 0.50 and 0.75 and poor reliability less than 0.50 (Portney & Watkins, 1993).

Retest reliability was also calculated as absolute reliability in terms of distance walked using 95% confidence intervals for groups and individuals (Taylor et al., 2004). Absolute reliability is a calculation of how much a group or individual would need to change to reflect true change over measurement error. Confidence intervals for the group mean scores were calculated using the following equation:

$$95\%\mathrm{CI(group)} = \mathrm{Mdiff} \pm \frac{t_{0.975} \times \mathrm{SDdiff}}{\sqrt{N}}$$

where Mdiff is the mean difference of retest minus test scores, SDdiff is the standard deviation of the difference between retest and test scores, N is the number of participants, and $t_{0.975}$ is the critical value for t with a two-tailed test at that sample size. To determine if there were changes between scores for first, second and third walks, a repeated measures analysis of variance (ANOVA) was performed. Post hoc comparisons (least significant difference) were performed to test for statistically significant differences between the walk tests (i.e. Walk 1 and Walk 2, Walk 2 and Walk 3, Walk 1 and Walk 3).

To determine the degree of change required in an individual, otherwise known as the limits of agreement (Altman, 1999) or least significant difference (Tillotson & Burton, 1991), the 95% confidence intervals were recalculated substituting $N = 1$:

$$95\%\mathrm{CI(individual)} = \mathrm{Mdiff} \pm t_{0.975} \times \mathrm{SDdiff}$$

A one-way ANOVA was applied to determine if there were any statistically significant differences between the 6MWT scores in Group A, Group B and Group C. Further, a mixed-plot two-way ANOVA was applied to determine if there were any interaction

effects between the groups (independent measure) and each walk (repeated measure).

Results

Thirty participants (24 men, 6 women) with an average age of 63 ± 7.9 years were recruited for this study. The average height of participants was 174 ± 7.0 cm, and the average weight was 84 ± 12.5 kg. Of the participants, 11 (37%) were referred to cardiac rehabilitation following a percutaneous intervention, 10 (33%) after coronary artery bypass graft surgery, six (20%) following commencement of medical management for ischaemic heart disease, two (7%) following mitral valve replacement and one (3%) following an aortic valve replacement. The average time elapsed since an acute event was 27 ± 11 days. All participants were able to mobilise independently without a gait aid. No participant had a history of falls. All participants completed the three 6MWTs safely without complications, and no test was prematurely ceased.

For the combined sample ($n = 30$), good relative retest reliability was shown for the three walks with the ICC (2,1) = 0.94 (see Table 1). For the individual groups, ICC (2,1) ranged from 0.83 to 0.84 (see Table 1).

For absolute reliability, the 95% CI were calculated for both the group and the individual (Table 2). For group results, a change of at least 45 m would be required as an indication of real change, over and above measurement error, in a cardiac rehabilitation group between the second and third 6MWTs and 99 m between the first and third 6MWT with 95% confidence.

Table 1. The mean distances walked during the tests

	Walk 1 (m)	Walk 2 (m)	Walk 3 (m)	ANOVA	ICC (2,1)
Group A ($n = 10$)	437 ± 81	505 ± 71	528 ± 63	$F(2,18) = 67.4$ $p < 0.001$	0.84
Group B ($n = 10$)	440 ± 69	487 ± 61	532 ± 59	$F(2,18) = 31.5$ $p < 0.001$	0.84
Group C ($n = 10$)	455 ± 80	495 ± 78	526 ± 93	$F(2,18) = 29.9$ $p < 0.001$	0.83
Total ($n = 30$)	444 ± 75	496 ± 68	529 ± 71	$F(2,58) = 109.5$ $p < 0.001$	0.94
Between groups ANOVA	$F(1,9) = 0.154$ $p = 0.86$	$F(1,9) = 0.160$ $p = 0.85$	$F(1,9) = 0.021$ $p = 0.98$	$F(4,54) = 1.67$[a] $p = 0.17$	

Group A completed three 6MWTs during the first session; Group B completed two 6MWTs during the first session; Group C completed one 6MWT during the first session.

[a]Mixed-plot two-way ANOVA.

Table 2. Group and individual confidence intervals

	Walk	Mdiff (m)	95% CI (group) (m)	95% CI (individual) (m)
Group A ($n = 10$)	2 and 1	68.1	49.4–86.8	8.9–127.3
	3 and 2	23.0	6.8–39.2	−28.3–74.3
	3 and 1	91.1	70.9–111.3	27.1–155.1
Group B ($n = 10$)	2 and 1	47.7	28.8–66.6	−12.1–107.5
	3 and 2	45.1	19.2–71.0	−36.9–127.1
	3 and 1	92.8	60.1–125.5	−10.6–196.2
Group C ($n = 10$)	2 and 1	39.8	26.1–53.5	−3.4–83.0
	3 and 2	31.4	10.5–52.3	−34.8–97.6
	3 and 1	71.2	45.1–97.3	−11.5–153.9
Combined group ($n = 30$)	2 and 1	51.9	42.1–61.7	−1.8–105.6
	3 and 2	33.2	21.9–44.5	−28.8–95.1
	3 and 1	85.0	71.0–99.1	8.2–161.9

Mdiff = mean difference; CI = confidence intervals; Group A completed three 6MWTs during the first session; Group B completed two 6MWTs during the first session; Group C completed one 6MWT during the first session.

There was a statistically significant increase in the mean distance walked over the three walks for the sample ($n = 30$) and for each group (see Table 1). Post hoc analysis showed that this difference was significant for all walk test score comparisons.

If the 6MWT was used as an individual outcome measure, a change in test score of at least 95 m would be required as an indication of real change, over and above measurement error, in an individual attending cardiac rehabilitation between the second and third 6MWTs and 162 m between the first and third 6MWT with 95% confidence.

There were no statistically significant differences in the walk test scores between the three groups (Walk 1 $F(1,9) = 0.154$, $p = 0.86$; Walk 2 $F(1,9) = 0.160$ $p = 0.85$; Walk 3 $F(1,9) = 0.021$, $p = 0.98$) (see Table 1). There were no differences between each walk for each group, and there were no interaction effects ($F(4,54) = 1.67$, $p = 0.17$).

Discussion

The purpose of this study was to assess the evidence for retest reliability of the 6MWT as a measure of exercise tolerance in a general cardiac rehabilitation population when up to three tests were performed and to determine if the time between tests affected test scores. The results indicated that in this cardiac rehabilitation population, three 6MWTs repeated in relatively short time frames were insufficient to yield reliable scores. Although results supported the relative reliability of the 6MWT in this population, they did not support the absolute reliability of the 6MWT.

The high ICCs in this study indicated good relative reliability with repeated testing. This finding is consistent with earlier reports by Cahalin et al. (1996) and O'Keefe et al. (1998). However, the use of reliability coefficients alone can be misleading (Keating & Matyas, 1998; Costa-Santos et al., 2011). Although participants generally kept their relative test score order to other participants in the group, the group showed significant improvements in the distance walked across the three walks, and relatively large increases in distance would be required to be interpreted as change with 95% confidence.

The performance of participants continued to improve over the three tests, and these improvements were significant. A potential learning effect (Wu et al., 2003) provides one explanation for this improvement. The results of this study are consistent with findings by Gayda et al. (2004), who studied elderly patients with cardiac disease, and Hamilton and Haennel (2000), who studied cardiac rehabilitation patients with mild disease. However, the results contrast with the findings of Patrick (2005), who studied a chronic heart failure population. It is possible that this specific population responded differently from the general cardiac rehabilitation population and, hence, may not be directly comparable. The 6MWT is a self-paced test; the participants in the current study were not limited by symptoms, whereas it is possible that patients with chronic heart failure experience symptoms such as breathlessness that limit performance.

This study demonstrated considerable increases in the 6MWT distances of participants referred to cardiac rehabilitation. Absolute reliability expressed as confidence intervals, even between the second and third walks, were large. The results of this study demonstrated that between the second and third 6MWTs, the individual would need to increase the distance walked by 95 m to be confident of real change over and above measurement error. In contrast, Patrick (2008) reported that an individual with chronic heart failure would need to improve by 64 m to be confident of real change. The improved score demonstrated by Patrick (2008) suggests that this test may be appropriate in people with chronic heart failure.

Some researchers have measured 6MWT results before and after the intervention of cardiac rehabilitation (Tallaj et al., 2001; Wright et al., 2001; Roberts et al., 2006). The change in 6MWT scores ranged from 57 ± 73.0 m in cardiac rehabilitation patients with left ventricular ejection fraction of equal to or greater than 40% studied by Tallaj et al. (2001) to 85.7 ± 80.7 m in cardiac rehabilitation patients with left ventricular ejection fraction of less than 40% in the same study. Wright et al. (2001) and Roberts et al. (2006) reported improvements of 62 ± 53 m and 67 m, respectively. The results of the current study suggest that the 6MWT is unlikely to have sufficient reliability to detect changes in individuals who complete cardiac rehabilitation because the distance they would need to increase over and above measurement error exceeds the increase in distance observed after cardiac rehabilitation programmes.

It is possible that adding further testing sessions, that is having a fourth or even a fifth testing session, may yield reliable results in terms of absolute reliability. However, in a clinical setting it is unlikely that clinical staff and patients would have the time to complete more

than three tests without causing delays to care and delay to the start of cardiac rehabilitation.

This study showed that the time between testing did not affect outcome. No difference was found between the group that completed three tests in one session, the group that completed two tests in the first session and a follow-up test within one week and the group that completed one test in the first session and two follow-up tests within one week. This finding suggests that, provided the test was reliable, testing could be fitted in with the schedule of the clinician and patient without affecting test results as long as three tests are completed.

When used as a measure of exercise tolerance in the general cardiac rehabilitation population, the 6MWT was not reliable even after three tests. The relatively high ICC demonstrated in our study is potentially misleading, as it obscures the fact that absolute reliability was not as good. The systematic measurement error was large and would likely obscure any true change in patients. These results suggest that the 6MWT is not a practical test to use in this population because it is not practical for a clinician or patient to complete more than three 6MWTs (Patrick, 2005). A test that is not reliable cannot be valid (Streiner & Norman, 1995). These results suggest that, in the general cardiac population, if it is not feasible to complete more than three tests, the 6MWT is not a reliable test. Further research is necessary to investigate other field exercise tests that are practical, reliable and valid tests for patients receiving cardiac rehabilitation.

Implications

Despite a good relative retest reliability expressed by the ICC, the large changes in distance over three tests and relatively low levels of absolute reliability suggest that the 6MWT is not a reliable test for the general cardiac rehabilitation population. When used in this population, three tests were not sufficient to produce a reliable test without systematic change. This test requires time and effort by the patient and the clinician and after three tests still records large systematic variability in scores. Although it did not make any difference whether the three tests were completed on one day or over a week, it may not be feasible for the cardiac rehabilitation patient and the clinician to continue repeating the 6MWT. Future research is needed to investigate other field exercise tests that are reliable, valid and feasible in the general cardiac rehabilitation population.

Potential conflict of interest

None.

REFERENCES

Altman D. Practical Statistics for Medical Research. London: Chapman & Hall, 1999.

American College of Sports Medicine. ACSM's Guidelines for Exercise Testing and Prescription (eighth edition). Baltimore, MD: Williams & Wilkins, 2010.

Butland RJ, Pang J, Gross ER, Woodcook AA, Gedds DM. Two-, six-, and 12-minute walking tests in respiratory disease. British Medical Journal (Clinical Research Education) 1982; 284: 1607–1608. DOI: 10.1136/bmj.284.6329.1607.

Cahalin LP, Mathier MA, Semigran MJ, Dec GW, Disalve TG. The six-minute walk test predicts peak oxygen uptake and survival in patients with advanced heart failure. Chest 1996; 110: 325–332. DOI: 10.1378/chest.110.2.325.

Costa-Santos C, Bernardes J, Ayres-de-Campos D, Costa A, Costa C. The limits of agreement and the intraclass correlation coefficient may be inconsistent in the interpretation of agreement. Journal of Clinical Epidemiology 2011; 64: 264–269. DOI: 10.1016/j.jclinepi.2009.11.010.

Gayda M, Temfemo A, Choquet D, Ahmaidi S. Cardiorespiratory requirements and reproducibility of the six-minute walk test in elderly patients with coronary artery disease. Archives of Physical Medicine and Rehabilitation 2004; 85: 1538–1543. DOI: 10.1016/j.apmr.2003.11.037.

Guyatt GH, Pugsley SO, Sullivan MJ, Thompson PJ, Berman LB, Jones NL, Fallen EL, Taylor DW. Effect of encouragement on walking test performance. Thorax 1984; 39: 818–822.

Guyatt GH, Sullivan MJ, Thompson PJ, Fallen EL, Pugsley SO, Taylor DW, Berman LB. The 6-minute walk: a new measure of exercise capacity in patient with chronic heart failure. Canadian Medical Association Journal 1985; 132: 919–923.

Hamilton DM, Haennel RG. Validity and reliability of the 6-minute walk test in a cardiac rehabilitation population. Journal of Cardiopulmonary Rehabilitation 2000; 20: 156–165. DOI: 10.1097/00008483-200005000-00003.

Harada N, Chiu V, Steward A. Mobility-related function in older adults: assessment with a 6-minute walk test. Archives of Physical Medicine and Rehabilitation 1999; 80: 837–841. DOI: 10.1016/S0003-9993(99)90236-8.

Keating JL, Matyas TA. Unreliable inferences from reliable measurements. The Australian Journal of Physiotherapy 1998; 44: 5–10.

Kervio G, Ville NS, Leclercq C, Daubert J, Carré F. Intensity and daily reliability of the six-minute walk test in moderate chronic heart failure patients. Archives

of Physical Medicine and Rehabilitation 2004; 85: 1513–1518. DOI: 10.1016/j.apmr.2003.09.035.

Kristjánsdóttir A, Ragnarsdóttir M, Einarsson MB, Torfason B. A comparison of the 6-minute walk test and symptom limited graded exercise test for phase II cardiac rehabilitation of older adults. Journal of Geriatric Physical Therapy 2004; 27: 65–68.

Lipkin DP, Scriven AJ, Crake T, Poole-Wilson A. Six minute walking test for assessing exercise capacity in chronic heart failure. British Medical Journal (Clinical Research Education) 1986; 292: 653–655.

Nogueira PA, Leal AC, Pulz C, Nogueira ID, Filho JA. Clinical reliability of the 6 minute corridor walk test performed within a week of a myocardial infarction. International Heart Journal 2006; 47: 533–540. DOI: 10.1536/ihj.47.533.

O'Keefe ST, Lye M, Donnellan C, Carmichael DN. Reproducibility and responsiveness of quality of life assessment and six minute walk test in elderly heart failure patients. Heart 1998; 80: 377–382.

Patrick JM. Effects of long versus short bouts of exercise in heart failure patients. Masters degree thesis, La Trobe University, Melbourne, 2005.

Patrick JM. When is one six minute walk test enough? Paper presented to the Australian Cardiac Rehabilitation Association Conference, Alice Springs, Australia, August, 2008.

Portney LG, Watkins MP. Foundations of Clinical Research: Applications to Practice. Norwalk, CT: Appleton & Lange, 1993.

Roberts E, Li FKW, Sykes, K. Validity of the 6-minute walk test for assessing heart rate recovery after an exercise-based cardiac rehabilitation programme. Physiotherapy 2006; 92 (2): 116–121. DOI: 10.1016/j.physio.2005.06.005.

Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. Psychological Bulletin 1979; 86: 420–428.

Steele B. Timed walking tests of exercise capacity in chronic cardiopulmonary illness. Journal of Cardiopulmonary Rehabilitation 1996; 16: 25–33.

Streiner DL, Norman GR. Health Measurement Scales. Oxford: Oxford University Press, 1995.

Tallaj TA, Sanderson B, Breland J, Adams C, Schumann C, Bittner V. Assessment of functional outcomes using the 6-minute walk test in cardiac rehabilitation: comparisons of patients with and without left ventricular dysfunction. Journal of Cardiopulmonary Rehabilitation 2001; 21: 221–224.

Taylor NF, Dodd KJ, Graham HK. Test–retest reliability of hand-held dynamometric strength testing in young people with cerebral palsy. Archives of Physical Medicine and Rehabilitation 2004; 85: 77–80. DOI: 10.1016/S0003-9993(03)00379-4.

Tillotson KM, Burton AK. Noninvasive measurement of lumbar sagittal mobility an assessment of the flexicurve technique. Spine 1991; 16: 29–33.

Troosters T, Gosselink R, Decramer M. Six minute walking distance in healthy elderly subjects. The European Respiratory Journal 1999; 14: 270–274. DOI: 10.1034/j.1399-3003.1999.14b06.x.

Wright DJ, Kahn KM, Gossage EM, Saltissi S. Assessment of a low intensity cardiac rehabilitation programme using the six-minute walk test. Clinical Rehabilitation 2001; 15: 119–124.

Wu G, Sanderson B, Bittner C. The 6-minute walk test: how important is the learning effect? American Heart Journal 2003; 146: 129–133. DOI: 10.1016/S0002-8703(03)00119-4.

Zugck C, Kruger C, Durr S, Gerber D, Haunstetter A, Horning K, Kubler W, Haass M. Is the 6-minute walk test a reliable substitute for peak oxygen uptake in patients with dilated cardiomyopathy? European Heart Journal 2000; 21: 540–549. DOI: 10.1053/euhj.1999.1861.

ELSEVIER

Physiotherapy

Physiotherapy 102 (2016) 243–248

The 10 m incremental shuttle walk test is a highly reliable field exercise test for patients referred to cardiac rehabilitation: a retest reliability study

Lisa C. Hanson [a,*], Nicholas F. Taylor [b], Helen McBurney [c]

[a] *La Trobe Rural Health School, La Trobe University, Bendigo, Victoria, Australia*
[b] *School of Allied Health, La Trobe University, Bundoora, Victoria, Australia*
[c] *School of Primary Health Care, Monash University, Victoria, Australia*

Abstract

Objectives To determine the retest reliability of the 10 m incremental shuttle walk test (ISWT) in a mixed cardiac rehabilitation population.
Design Participants completed two 10 m ISWTs in a single session in a repeated measures study. Ten participants completed a third 10 m ISWT as part of a pilot study.
Setting Hospital physiotherapy department.
Participants 62 adults aged a mean of 68 years (*SD* 10) referred to a cardiac rehabilitation program.
Main outcome measures Retest reliability of the 10 m ISWT expressed as relative reliability and measurement error. Relative reliability was expressed in a ratio in the form of an intraclass correlation coefficient (*ICC*) and measurement error in the form of the standard error of measurement (*SEM*) and 95% confidence intervals for the group and individual.
Results There was a high level of relative reliability over the two walks with an *ICC* of .99. The $SEM_{agreement}$ was 17 m, and a change of at least 23 m for the group and 54 m for the individual would be required to be 95% confident of exceeding measurement error.
Conclusions The 10 m ISWT demonstrated good retest reliability and is sufficiently reliable to be applied in practice in this population without the use of a practice test.

Keywords: Heart disease; Exercise test; Reproducibility of results

Introduction

The 10 m ISWT is an externally paced and incremental field exercise test that may provide clinicians with information on an individual's exercise capacity and limitations, change over time, and individual or group response to an intervention. The test was originally developed to assess the functional capacity of patients with chronic airways obstruction [1] but has since been shown to be well tolerated in cardiac populations [2].

Previous research supports the relative reliability of the 10 m ISWT in the cardiac rehabilitation population, with reliability coefficients between the first and second walk ranging from 0.80 [3] to 0.94 [2,4] to 0.98 [5]. Relative reliability is further improved when the test is repeated, with reliability coefficients of .99 between the second and third walk [4,6].

Less information is known about the measurement error of the 10 m ISWT in cardiac rehabilitation populations and the number of tests required to achieve scores that minimise measurement error. Measurement error can be estimated for an individual score or around a change score for either a group, such as a cardiac rehabilitation group or for an individual within a program. Previous research has demonstrated that for change scores within a group the minimum amount of change required to be interpreted as real change over and

* Correspondence: La Trobe Rural Health School, PO Box 199, Bendigo, Victoria 3552, Australia. Tel.: +61 3 5448 9102; fax: +61 3 5448 9199.
E-mail address: L.Hanson@latrobe.edu.au (L.C. Hanson).

http://dx.doi.org/10.1016/j.physio.2015.08.004

above measurement error after the completion of one 10 m ISWT ranged from 36 m in a mixed cardiac rehabilitation group in a single session [2], to 44 m in a group who were attending a community based maintenance cardiac rehabilitation program and completed the testing over a minimum of eight weeks [3] to 56 m for a group of patients following coronary artery bypass surgery who completed the tests within one week [4]. For an individual, after one test an improvement in test score required to be interpreted as more than measurement error ranged from 53 m in individuals with heart failure in a single session [7] to 122 m [4] to 203 m [3]. Additionally, Fowler *et al.* [4] reported measurement error was minimised if a practice walk was included. They demonstrated if a second walk was completed within one week, an increase in walk distance of more than 5 m for groups and 21 m for individuals exceeded measurement error with 95% confidence, in patients following coronary artery bypass graft surgery [4].

There remains uncertainty about whether there is systematic error when the 10 m ISWT is repeated in a cardiac rehabilitation population. Three papers reported significant improvement in test scores for the 10 m ISWT between the first and second walk [2,4,8], whereas four papers reported no significant change in test scores [5,7,9,10]. Two papers reported using a practice walk but did not provide any further information in their data analysis [5,6]. It is possible that a practice walk may be required when using the 10 m ISWT in a general outpatient cardiac rehabilitation population but little information is available about the absolute reliability of the test when assessed in a single session. The aim of the current study was to determine the relative reliability and measurement error of the 10 m ISWT in a mixed cardiac rehabilitation population.

Method

Design

A repeated measures design was used in the physiotherapy department at the two participating hospitals. All participants completed two 10 m ISWTs in a single session prior to commencing a cardiac rehabilitation program. The first 10 participants participated in a pilot study and completed a third 10 m ISWT in the same session. All walk tests were completed under the same conditions and by the same assessor.

Participants

Eligible participants were all adults with coronary artery disease referred to cardiac rehabilitation irrespective of severity or duration of the condition. Exclusion criteria included any medical condition where exercise would be contraindicated; unable to walk for any neurological or musculoskeletal reason; presentation to cardiac rehabilitation for risk factor reduction or with congenital heart disease; children or pregnancy. Participants were excluded if they had previously completed a 10 m ISWT or cardiac rehabilitation or if limitations in English language production or comprehension skills precluded them from understanding the consent form.

University and Hospital Human Ethics Committees approved this research. All participants volunteered to participant in the study and provided written informed consent. No patient who met the inclusion criteria refused to participate in the study.

A sample size calculation was completed according to the method described by Walter *et al.* [11] and based on a priori set levels of optimal and minimal acceptable limits of reliability for clinical measurement. For two tests, a minimum of 46 people would be needed if a minimum *ICC* level of .8 (P_0) was accepted and the hypothesis that findings from this study would be consistent with the current literature at an *ICC* of .9 (P_1) [2,4,5], at a level of significance (α) of .05 and power of .8 ($\beta = 0.2$).

The baseline descriptive characteristics of the participants are summarised in Table 1. Of the 62 participants, 39 (63%) were referred following a revascularisation procedure and 16 (26%) were referred following medical management for coronary artery disease and seven (11%) following other cardiac interventions. The mean age of participants was 68 years (*SD* 10) years ranging from 46 to 91 years.

Procedure

The 10 m ISWT protocol was administered according to the description of Singh *et al.* [1]. Participants walked along an indoor flat 10 m course marked by two cones placed 0.5 m in from each end of the course (Fig. 1). A shuttle referred to one 10 m lap. Standardised prerecorded instructions for the test were played from a digital recording immediately prior to beginning the test. The test was externally paced, with

Table 1
Baseline demographic characteristics of the sample.

	Main study ($n = 62$)	Pilot study sample ($n = 10$)
Age, years	68 (10)[a]	67 (10)[a]
Gender		
Male:female, *n*	45:17	5:5
Intervention, *n* (%)		
Revascularisation procedure	39 (63)	8 (80)
Medical management	16 (26)	2 (20)
Other	7 (11)	0
Days post most recent cardiac event	29 (18)	28 (12)
Height, cm	170 (9)[a]	161 (11)[a]
Weight, kg	84 (15)[a]	83 (17)[a]
BMI, kg/m^2	29 (5)[a]	32 (5)[a]
Use of a gait aid, *n* (%)		
Single point stick	3 (5%)	2 (20%)

Note: BMI, body mass index.
[a] Mean (*SD*).

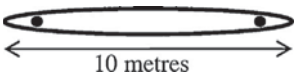


Fig. 1. Schematic representation of the 10 m incremental shuttle walk test.

signal beeps at regular intervals to indicate when the participant should be turning around the cone to commence the next shuttle. A triple beep signalled the next level and an increase in walk speed [1]. Participants commenced the test at a walking speed of 0.5 m/seconds (level 1), allowing the participant 20 seconds to complete each of the three shuttles in level 1. There was a speed increment of 0.17 m/seconds each minute for a maximum of 12 minutes. There were 14 shuttles in level 12, requiring a walking speed of 2.37 m/seconds [1].

The test was stopped when the participant could no longer maintain the required pace or was more than 0.5 m from the cone before the signal beep after one opportunity to catch up or if the test was completed. Additional criteria for early termination of the test included patient distress, dizziness, angina, or onset of severe musculoskeletal pain, failure of the heart rate to increase with exercise, fall in oxygen saturation below 90% [12] or attainment of 85% of the maximum heart rate [1] using the heart rate reserve method. The number of shuttles completed were recorded and at the completion of each test converted to the distance walked.

Participants were allowed a 20 minute rest break prior to commencing the first test and were given a 30 minute rest break between subsequent tests. Prior to commencing the test, with the participant in a seated position, baseline heart rate, blood pressure, respiratory rate, oxygen saturation and rate of perceived exertion using the rate of perceived exertion (RPE) Borg 6-20 scale were recorded. Participants were not able to commence the 10 m ISWT if resting systolic blood pressure was greater than or equal to 200 mmHg or resting diastolic blood pressure was greater than or equal to 110 mmHg [12]. Heart rate and oxygen saturation were monitored throughout each test using a portable pulse oximeter without interruption to the test. Monitoring of heart rate, blood pressure, oxygen saturation, respiratory rate and RPE continued after completion of the test until all values were within 20% of the baseline recordings.

Statistical methods

Where relevant, means were expressed followed by standard deviation. In the case of missing values, data were not imputed. Retest reliability was expressed in two ways: first as relative reliability, and second as measurement error. The relative reliability of the sample was interpreted using the in intraclass correlation coefficient ($ICC_{agreement}$) two-way random effects model with 95% confidence intervals across the two walks. The *ICC* was interpreted with the following guidelines: good reliability was a score greater than .75, moderate reliability between .50 and .75 and poor reliability less than .5 [13].

For the measurement error around an individual score the $SEM_{agreement}$ was derived. The $SEM_{agreement}$ takes into account the variability in repetition, and was derived using the following equation [14]:

$$SEM_{agreement} = \sqrt{(\sigma^2_{observations} + \sigma^2_{residual})}$$

where $\sigma^2_{observations}$ is the variance due to systematic difference between the tests and $\sigma^2_{residual}$ is the variance due to the interaction between subjects and observations. The $SEM_{agreement}$ as a percentage of the grand mean was also expressed. This provided information on the relative size of the $SEM_{agreement}$.

Group and individual 95% confidence intervals were used to calculate the measurement error around the change scores [15]. Confidence intervals for the group mean scores were calculated using the following equation:

$$95\% \quad CI_{(group)} = M_{diff} \pm \frac{t_{0.975} \times SD_{diff}}{\sqrt{n}}$$

where M_{diff} is the mean difference of retest minus test scores, SD_{diff} is the standard deviation of the difference between retest and test scores, n is the number of participants, and $t_{0.975}$ is the critical value for t with a two-tailed test at that sample size.

To determine the degree of change required in an individual, otherwise known as the limits of agreement the 95% confidence intervals were recalculated substituting $n = 1$ into the previous equation [15].

Data for absolute agreement for the individual were presented graphically using the technique described by Bland and Altman [16]. The individual mean test and retest scores were plotted against the corresponding individual change scores for all combinations of pairs of walk tests as well as the 95% limits of agreement. Change scores outside the 95% limits of agreement were considered real change and change scores that fall within the limits of agreement cannot be distinguished from measurement error.

Pilot study

For the 10 participants in the pilot study, the mean distance for the first walk was 378 m (*SD* 224, range 30 to 740 m), for the second walk was 393 m (*SD* 239, range 30 to 790 m) and for the third walk was 398 m (*SD* 229, range 40 to 770 m). Whilst there was a trend for an increase in the distance walked from the first to the second to the third walk, this distance was not significant ($F_{(2,18)} = 3.197$, $P = .065$). Good relative retest reliability was obtained (see Table 2), with an $ICC_{agreement}$ of .992 (95% *CI*: .976 to .998) over the three walks. The $SEM_{agreement}$ and the percentage of the grand mean and group and individual confidence intervals for the pilot study are shown in Table 2. Based on the high level of retest reliability demonstrated in the pilot study it was decided to proceed with testing on two walks for the full sample.

Table 2
Summary of the reliability from the pilot study ($n = 10$).

ISWT walk	$ICC_{agreement}$ (95% *CI*)	$SEM_{agreement}$ (m)	M_{diff} (SD_{diff}) (m)	95% *CI* (group)	95% *CI* (individual)
1 and 2	.990 (.961 to .998)	23 (6%)	15 (30)	−6 to 36 m	−53 to 83 m
1 and 3	.989 (.940 to .997)	24 (6%)	20 (29)	−1 to 41 m	−46 to 86 m
2 and 3	.997 (.990 to .999)	12 (3%)	5 (17)	−7 to 17 m	−33 to 43 m

Note: ISWT, 10 m ISWT; *ICC*, intraclass correlation; SEM, standard error of measurement; M_{diff}, mean difference; SD_{diff}, standard deviation; *CI*, confidence intervals.

Results

A consecutive series of 62 patients completed two 10 m ISWTs in a single session. All participants completed the 10 m ISWT without complications. The mean distance walked improved by 17 m (*SD* 18) over the two walks. Participants walked a mean distance of 378 m (*SD* 173, range 30 to 760 m) for ISWT Walk 1 and 395 m (*SD* 176, range 30 to 790 m) for ISWT Walk 2. This difference was statistically significant ($t_{(61)} = 7.613$, $P = <.001$).

Good relative retest reliability was shown between walk 1 and 2, the two-way mixed effects $ICC_{agreement}$ was .990 (95% *CI*: .928 to .997). The $SEM_{agreement}$ was 17 m or 4% of the grand mean. For the group, a change of at least 23 m would be required as an indication of real change, over and above measurement error with 95% confidence, between the first and second test. For the individual in this population, a change of at least 54 m from the first to the second walk test would be required as an indication of real change, over and above measurement error with 95% confidence (Table 3).

A Bland Altman plot for the sample is presented in Fig. 2. Fig. 2 shows one participant with a difference in distance walked outside the upper limit of the limits of agreement. This participant was a 62 year old woman who was 36 days post coronary artery bypass graft surgery with a Walk 1 distance of 480 m and a Walk 2 distance of 560 m completed 20 minutes later.

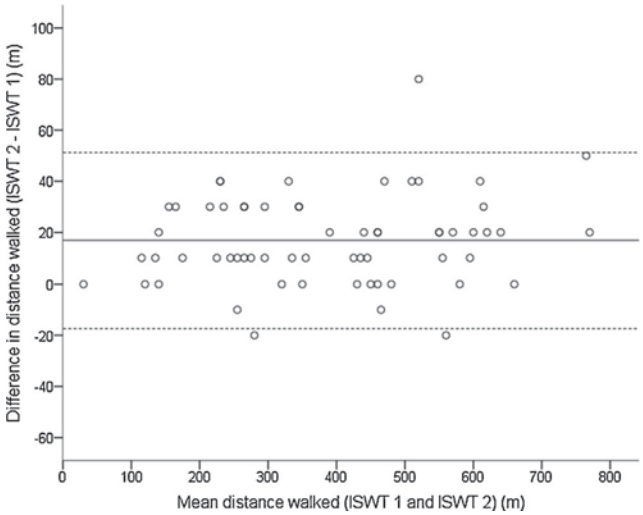


Fig. 2. Bland Altman plots of the mean distance walked in 10 m ISWT score plotted against the difference in walk distance between the retest and test score for walk 1 (ISWT 1) and walk 2 (ISWT 2). The *x*-axis is the mean score, calculated by (ISWT 1 + ISWT 2)/2; and *y*-axis the absolute difference, calculated by ISWT 2 − ISWT 1. Key: unbroken line = mean difference, broken lines = limits of agreement calculated by mean difference ± 1.96 × SD_{diff}.

Discussion

This study investigated the relative reliability and the measurement error of repeated 10 m ISWTs. The study supports the reliability of the 10 m ISWT in cardiac rehabilitation population. The results provide evidence that the addition of the second test may not be clinically relevant and one test may be sufficient for adequate test retest reliability in a cardiac rehabilitation population. The relative reliability expressed as the $ICC_{agreement}$ was reported at least .99 between the first and second walk in the main study. The $SEM_{agreement}$ was 17 m or 4% of the grand mean and the 95% confidence intervals suggest the group would need to improve by 23 m and the individual by 54 m to be confident of true change over and above measurement error.

Despite the high levels of reliability observed, the results indicated there was a statistically significant difference between the distance walked between the first and the second walk. The mean difference was 17 m. Although statistically significant, the magnitude of the difference is unlikely to be clinically significant. The improvement in walk distance maybe attributed to test familiarisation or a learning effect similar to that seen in the self-paced 6MWT [17,18]. It is also possible that the small increase in distance walked was a result of a tiring effect from completing two walk tests in a single session. However, there is evidence that walk distance is not affected by the time between tests ranging from a single session to one week [17]. Therefore, while, in this study, a patient in a cardiac rehabilitation program may be expected to

Table 3
Summary of measurement error for the change scores ($n = 62$).

ISWT walk	M_{diff} (SD_{diff}) (m)	Group 95% *CI* (m)	Individual 95% *CI* (m)
1 and 2	17 (18)	12 to 22	−19 to 53

Note: ISWT, 10 m ISWT; M_{diff}, mean difference; SD_{diff}, standard deviation of the difference; *CI*, confidence intervals.

walk a little further if tested on a second occasion, the increase in distance is small relative to the total distance walked, and small relative to what might be expected to occur after an intervention.

Two studies have reported group improvements of more than 58 m following a six week intervention [4,19]. Both studies reported at least one practice 10 m ISWT in the initial baseline testing, but no practice test at follow-up, after the intervention. The baseline scores reported in these studies, ranging from 444 m (*SD* 135) [4] to 618 m (*SD* 165) [19] were higher than the results of our study. This may have been due to an improved baseline level of fitness, the younger age or higher proportion of male participants in the previous two studies compared with our study. Our baseline walk test distances are consistent with the results in a recent study with a similar cohort in terms of age, body mass index and proportion of male participants which reported a baseline distance of 360 m (*SD* 90) [20]. Our results suggest that a single test of the 10 m ISWT is sufficiently reliable to detect this amount of typical change observed after a short cardiac rehabilitation intervention, and for that change to be interpreted as real change over and above measurement error both for individuals and in the evaluation of group programs.

Our results suggest that the 10 m ISWT has a higher level of retest reliability than another common field exercise test, the 6MWT. Hanson *et al.* [17] reported that after three tests large changes were required to overcome measurement error in both the group and individual with 95 per cent confidence. Our pilot study showed that when a practice walk was included, the group would need to improve by 18 m and the individual by 44 m. If no practice walk was included, the group would need to improve by 23 m and the individual by 54 m. In comparison, the earlier study by Hanson *et al.* [17] showed when the 6WMT was used in a mixed outpatient cardiac rehabilitation population and a practice walk was included an improvement of 45 m for the group and 96 m for the individual was required to overcome measurement error. The greater number of tests required for a reliable retest result in the 6MWT make it less practical in a clinical setting.

Study strengths and limitations

This study adheres to the recommendations for methodological quality proposed in the COSMIN checklist for studies relating to the measurement properties relative reliability and measurement error [21]. In their recommendations on sample size, a good sample size was considered to lie between 50 and 99 inclusive.

Participants in this study were from a mixed cardiac rehabilitation group, and all had stable and treated cardiac disease. It is unlikely that these findings can be generalised beyond this population, for example to include people with untreated cardiac disease. Also, it could be considered a limitation that the pilot study involving three walks was completed on a relatively small sample. However, the positive results with two walks confirm that a third walk adds little value in improving reliability in the ISWT.

Conclusion

The 10 m ISWT is feasible when used as a field exercise test in a mixed outpatient cardiac rehabilitation population. Our results indicate good retest reliability of the 10 m ISWT when one test is completed at baseline screening. Better reliability can be obtained with a practice test but it may not be necessary if the purpose of the test is to evaluate change after completing a cardiac rehabilitation program. The choice of completing one or two tests may be left to the clinician or researcher to balance the reliability with the time taken to complete a second test. The externally paced and incremental 10 m ISWT appears to have better retest reliability than the self-paced 6MWT in this population.

Ethical approval: This study was approved by La Trobe University: University Human Ethics Committee, HEC10-082; La Trobe University: Faculty Human Ethics Committee, FHEC07/99; St John of God Health Care Ethics Committee, Ref. No. 438; and Bendigo Health: Human Research Ethics Committee, Ref. No. 1/2007.

Conflict of interest: None declared.

References

[1] Singh SJ, Morgan MD, Scott S, Walters D, Hardman AE. Development of a shuttle walking test of disability in patients with chronic airways obstruction. Thorax 1992;47:1019–24.
[2] Jolly K, Taylor RS, Lip GYH, Singh S. Reproducibility and safety of the incremental shuttle walking test for cardiac rehabilitation. Int J Cardiol 2008;125:144–5.
[3] Pepera G, McAllister J, Sandercock G. Long-term reliability of the incremental shuttle walking test in clinically stable cardiovascular disease patients. Physiotherapy 2010;96:222–7.
[4] Fowler SJ, Singh SJ, Revill S. Reproducibility and validity of the incremental shuttle walking test in patients following coronary artery bypass surgery. Physiotherapy 2005;91:22–7.
[5] Green DJ, Watts K, Rankin S, Wong P, O'Driscoll JG. A comparison of the shuttle and 6 minute walking tests with measured peak oxygen consumption in patients with heart failure. J Sci Med Sport 2001;(4):292–300.
[6] Lewis ME, Newall C, Townend JN, Hill SL, Bonser RS. Incremental shuttle walk test in the assessment of patients for heart transplantation. Heart 2001;86:183–7.
[7] Pulz C, Diniz RV, Alves AN, Tebexreni AS, Carvalho AC, de Paola AA, *et al.* Incremental shuttle and six minute walking tests in the assessment of functional capacity in chronic heart failure. Can J Cardiol 2008;24:131–5.
[8] Morales FJ, Martinez A, Mendez M, Agarrado A, Ortega LF, Fernandez-Guerra J, *et al.* A shuttle walk test for assessment of functional capacity in chronic heart failure. Am Heart J 1999;138:291–8.
[9] Arnott AS. Assessment of functional capacity in cardiac rehabilitation. Coron Health Care 1997;1:30–6, http://dx.doi.org/10.1016/S1362-3265(97)80037-5.
[10] Payne GE, Skehan JD. Shuttle walking test: a new approach for evaluating patients with pacemakers. Heart 1996;75:414–8.

European Journal of Cardiovascular Nursing

EUROPEAN SOCIETY OF CARDIOLOGY®

Original Article

European Journal of Cardiovascular Nursing
1–11

DOI: 10.1177/1474515117721129
journals.sagepub.com/home/cnu
SAGE

Is the 10 m incremental shuttle walk test a useful test of exercise capacity for patients referred to cardiac rehabilitation?

Lisa C Hanson[1], Helen McBurney[2] and Nicholas F Taylor[3]

Abstract

Background: Field exercise tests are a feasible alternative to the symptom-limited exercise test for measuring exercise capacity in patients attending cardiac rehabilitation.

Aims: To evaluate the criterion validity of the 10 m incremental shuttle walk test (ISWT) as a useful tool for measurement of exercise capacity for patients eligible for cardiac rehabilitation.

Methods: Fifteen patients eligible for cardiac rehabilitation completed a treadmill symptom-limited exercise test and two ISWTs with the order of testing randomised. Associations between the symptom-limited exercise test and the ISWT were explored and the ability of the ISWT to predict symptom-limited exercise test duration estimated.

Results: There was a moderate to high association ($r^2 \geqslant 0.72$) between ISWT duration and distance, and symptom-limited exercise test duration; and a moderate association between ISWT peak heart rate and end of test oxygen saturation and these outcomes on the symptom-limited exercise test ($0.47 \leqslant r^2 \leqslant 0.67$). However, prediction of symptom-limited exercise test duration based on the ISWT produced wide 95% confidence intervals, for example, ranging from 9.1 to 16.3 minutes for an individual who completes the ISWT. Order of testing did not affect the results and the association between the ISWT and symptom-limited exercise test was similar for both the first (ISWT 1) and second test (ISWT 2) ISWT.

Conclusions: The results provide support for the ISWT as a convenient field test of exercise capacity in a cardiac rehabilitation population, but not as a surrogate to predict symptom-limited exercise test duration for individuals. A single ISWT may provide as good an estimate of exercise capacity as repeating the test.

Keywords

10 m ISWT, criterion validity, cardiac rehabilitation, exercise test, walk test, validity of results, coronary heart disease, symptom-limited exercise test

Date received: 11 October 2016; revised: 26 June 2017; accepted: 27 June 2017

Introduction

The gold standard for measuring physical fitness and functional capacity in patients with cardiac disease is the symptom-limited exercise test with electrocardiogram (ECG) monitoring or the cardiopulmonary exercise test.[1–5] However, despite these recommendations, it is not always possible to complete a symptom-limited exercise test at entry and exit to a cardiac rehabilitation exercise programme.[6] Field exercise tests, including the six minute walk test (6MWT) and the 10 m incremental shuttle walk test (ISWT), may be a suitable alternative for entry assessment into a low to moderate intensity exercise-based cardiac rehabilitation programme.[7] These alternatives for exercise testing have been included in national guidelines

[1]La Trobe Rural Health School, La Trobe University, Australia
[2]School of Primary Health Care, Monash University, Australia
[3]School of Allied Health, La Trobe University, Australia

Corresponding author:
Lisa Hanson, La Trobe Rural Health School, La Trobe University, PO Box 199, Bendigo, Victoria 3552, Australia.
Email: L.Hanson@latrobe.edu.au

for cardiac rehabilitation programmes in Australia, New Zealand, throughout the UK and USA.[8–11]

Recent research suggests that the 6MWT may not have sufficient test–retest reliability when used in the entry assessment to cardiac rehabilitation.[12] However, the ISWT demonstrates good test–retest reliability without the use of a practice walk.[13] The ISWT is an externally paced, incremental field exercise test. Test results provide objective measures used to make inferences about exercise capacity.[14,15] The test requires minimal equipment, can be safely completed in a clinical setting and does not need to be supervised by a medical doctor.[16] A recent study showed no major adverse incidents over 1457 ISWTs undertaken in patients with heart disease.[16] The investigation of criterion validity, or the degree to which the outcome measures of the ISWT agree with the outcome measures of a gold standard[17] in cardiac rehabilitation populations has been reported for patients with chronic heart failure,[18–22] coronary artery bypass surgery,[23] and elderly stable coronary artery disease patients.[24] All followed the protocol described by Singh et al.[25] All compared the ISWT distance walked or number of shuttles or peak ISWT speed with the peak oxygen consumption during a symptom-limited exercise test and reported moderate to high correlations ($0.72 \leqslant r \leqslant 0.84$) and 52–71% of the variation in the criterion measure predicted from the 10 m ISWT distance or number of shuttles walked.[19,20,22–24,26]

There is limited published research comparing other functional outcome measures from a symptom-limited exercise test with the ISWT. Measuring oxygen consumption is not always possible in clinical cardiology practices. The associated costs of staff, equipment outlay and equipment maintenance are often prohibitive.[27] Functional outcomes of a symptom-limited exercise test include duration, maximal heart rate and blood pressure, oxygen saturation, symptoms and limiting factors and ECG changes. These functional measures are commonly reported in patient referrals to cardiac rehabilitation and where maximal exertion was attained, these functional outcome measures have been considered as useful as oxygen consumption.[28,29]

The association between the ISWT and functional outcome measures of a symptom-limited exercise test in patients with chronic heart failure has been reported to be moderate, between 29% and 46% of the variation in the maximum heart rate of the gold standard measure could be predicted from variation in peak heart rate in the ISWT.[18,20–22,24] There is no information on the association, specifically the concurrent validity or the predictive validity, of the functional outcomes of the ISWT in a cardiac rehabilitation population.

The aim of this research was to determine the concurrent and predictive criterion validity of the ISWT as an objective measure of exercise capacity in a cardiac rehabilitation population. We aimed to determine first, if those who achieve a high functional level on a symptom-limited exercise test also do well on the ISWT and if functional variables such as peak exercise heart rate measured during the symptom-limited exercise test were correlated with the ISWT, and second, if the ISWT distance could predict the symptom-limited exercise test duration.

Methods

Design

The research design involved a concurrent validation model. Both the predictor test (the ISWT) and the criterion test (symptom-limited exercise test) were administered to a sample of participants with coronary heart disease, with the order of tests randomly allocated. A consecutive series of patients referred to the cardiology clinic in a non-metropolitan regional city in Australia for a symptom-limited exercise test between June 2014 and July 2014 who met the eligibility criteria of the study were invited to participate by the manager of the centre. The clinic serves an area of 59,000 km^2 with a population of 100,000. University and hospital human ethics committees approved this research. All participants volunteered to participate in the study and provided written informed consent. No patient who met the inclusion criteria refused to participate in the study.

Participants

The clinic nurse identified the patients eligible for participation and provided a basic overview of the study and initial consent to be contacted by the researcher (LCH). A detailed description of the project and informed consent was obtained by the researcher.

All adults (>18 years) with stable and treated coronary heart disease, irrespective of severity or duration of the condition, who were referred to complete a treadmill symptom-limited exercise test in the cardiology clinic, were eligible to participate in this study. Participants were excluded if they had any condition for which exercise would be contraindicated, if they were unable to walk for any neurological or musculoskeletal reason, or were required to complete a bicycle or another alternative symptom-limited exercise test. Participants were also excluded if they had previously completed an ISWT, were diagnosed with cardiovascular risk factors in the absence of diagnosed cardiovascular disease, with congenital heart disease or were children or were pregnant, or were required to cease their usual cardiac medications for the symptom-limited exercise test.

Sample size estimates were completed according to the formula for correlation described by Howell.[30] The following formula was used:

$$n = \left(\frac{\delta}{\rho_1} \right)^2$$

where n is the sample size; δ is a constant and for a power of 0.80 equals 2.8; and ρ_1 is an estimate of the correlation

in the population, and was based on the minimum correlation of 0.75 which was defined as the lower limit of a strong correlation[31] and greater than a suggested minimal acceptable correlation of 0.70.[32] A sample of n=15 would be required for a power of 0.80 and a minimum correlation of 0.75. Recommendations for the sample size for prediction equations vary from five to 10 cases per independent variable[35] to 50 cases per independent variable.[34] Based on this, the sample size estimate of 15 was used to explore the evidence for the concurrent and predictive criterion validity of the ISWT in patients with cardiac disease.

Outcome measures

To account for any series effects participants were randomly allocated to complete either the symptom-limited exercise test or ISWT first. Each participant completed both the symptom-limited exercise test and ISWT within one week, with a minimum elapsed time of one day.

Criterion measure: symptom-limited exercise test. Participants completed their symptom-limited exercise test under the supervision of a medical doctor at the cardiology clinic and followed the usual practice guidelines for the clinic. Each participant rested for a minimum of 30 minutes prior to commencing the test. Standardised instructions were given to the patient and prior to commencing the test, and then a resting ECG and non-invasive blood pressure were recorded. The symptom-limited test was administered according to the Bruce protocol.[35] The Bruce protocol has relative reliability of 0.78 in patients with known coronary artery disease.[36] The test terminated if the patient achieved maximal exertion or breached any of the following criteria including significant ECG changes: a reduction in blood pressure with increasing workload; unreasonable hypertension; onset of angina or increasing angina; patient reporting symptoms of distress or dizziness; excessive shortness of breath or claudication; or changes in general appearance. The primary outcome measure was the duration of the test. Secondary outcome measures were peak exercise heart rate and oxygen saturation measured by pulse oximetry.

Predictor test: ISWT. The ISWT protocol was administered according to the description of Singh et al.[25] Participants walked along an indoor flat 10 m course. The test was externally paced, with signal beeps at regular intervals to indicate when the participant should turn around the cone to commence the next shuttle. The test commenced with a walking speed of 0.5 m/second, with speed increments of 0.17 m/second each minute for a maximum of 12 minutes. The ISWT has relative reliability of 0.99 in a cardiac rehabilitation population.[13] The primary outcome measures were distance walked, calculated from the number of completed shuttles and test duration. The secondary outcome measures were peak heart rate and oxygen saturation at test completion.

Participants were allowed a 20-minute rest prior to commencing the first ISWT and were given a 30-minute rest before starting the second ISWT.[37] Prior to commencing the test, baseline heart rate, blood pressure, respiratory rate, oxygen saturation and rate of perceived exertion (RPE) using the RPE Borg 6–20 scale were recorded. Participants were not able to commence the ISWT if resting systolic blood pressure was greater than or equal to 200 mmHg or resting diastolic blood pressure was greater than or equal to 110 mmHg.[1,38] Heart rate and oxygen saturation were monitored throughout each test, using a portable pulse oximeter, without interruption to the test. Monitoring of heart rate, blood pressure, oxygen saturation, respiratory rate and RPE continued after completion of the test until all values returned to within 20% of baseline.

Statistical methods

Data were analysed using the statistical package for the social sciences (IBM SPSS Statistics for Windows, version 22.0; IBM Corp., Armonk, NY, USA). Where relevant, means were expressed followed by standard deviation. In the case of missing values, data were not imputed. After checking for any series effects due to the order of testing using a mixed plot two-way analysis of variance, the groups were combined for testing of criterion validity.

Criterion validity of the ISWT. The strength of the relationship between the symptom-limited exercise test and two trials of the ISWT for all primary and secondary outcome measures was assessed with Pearson's product–moment correlation coefficient (r) with 95% confidence intervals (CIs). The correlation coefficient was interpreted as strong, a correlation greater than 0.75, moderate between 0.50 and 0.75 and weak less than 0.5.[31] The index of reliability, in other words the estimated maximum correlation based on error in the two tests was 0.88. Bland Altman plots were used to display graphically the concurrent criterion-related validity of the ISWT. The variation in the symptom-limited exercise test explained by the ISWT was summarised by the coefficient of determination (r^2).[39]

The predictive validity of the ISWT was examined using the standard error of estimate and linear regression with 95% CIs for the slope (b) for the ISWT distance and the symptom-limited exercise test duration. CIs were calculated around the regression equation, and for the group and individual. The 95% confidence limits for the symptom-limited exercise test duration for the group and individual at specific ISWT scores (X) were calculated.[30,40] Individual confidence limits are also known as the prediction interval.[30] The specific ISWT distances were the distance walked at the completion of each level of the ISWT.

Table 1. Participant characteristics.

Characteristics	Group A (*n*=7)	Group B (*n*=8)	Total (*n*=15)
Age	63 (7)	67 (8)	65(8)
Gender, *n* male:female	5:2	7:1	12:3
Intervention, *n* (%)			
Revascularisation procedures	6 (86)	7 (88)	13 (87)
Medical management	1 (14)	1 (13)	2 (13)
Height, cm	174 (12)	174 (11)	174 (11)
Weight, kg	92 (24)	89 (16)	91 (19)
BMI, kg/m^2	30 (6)	30 (5)	30 (5)
BMI category, *n* (%)			
Underweight	0	0	0
Healthy	1 (14)	2 (25)	3 (20)
Overweight	2 (29)	1 (13)	3 (20)
Obese	3 (43)	5 (63)	8 (53)
Morbidly obese	1 (14)	0	1 (7)
Waist, cm	103 (16)	104 (12)	103 (10)
Hip, cm	106 (10)	103 (12)	104 (10)
Medications, *n* (%)			
Nitrates	0 (0)	1 (13)	1 (7)
β-blockers	5 (71)	5 (63)	10 (67)
ACE inhibitor	2 (29)(	1 (13)	3 (20)
Calcium antagonist	4 (57)	1 (13)	5 (33)
Angiotension II receptor blocker	4 (57)	2 (25)	6 (40)
Lipid-lowering statin	7 (100)	6 (75)	13 (87)
Antiplatelet	7 (100)	8 (100)	15 (100)
Diuretic	0 (0)	1 (13)	1 (7)
Smoking status, *n* (%)			
Never	3 (43)	6 (75)	9 (60)
Ex-smoker	4 (57)	2 (25)	6 (40)
Current smoking	0 (0)	0 (0)	0 (0)
Medical comorbidities, *n* (%)			
Diabetes mellitus (type 2)	1 (14)	3 (38)	4 (27)
Respiratory	1 (14)	1 (13)	2 (13)
Hypertension	2 (29)	5 (63)	7 (47)
Depression/anxiety	1 (14)	0 (0)	1 (7)
Lower limb musculoskeletal	4 (57)	5 (63)	9 (60)
Upper limb musculoskeletal	3 (43)	3 (38)	6 (40)
Other musculoskeletal	3 (43)	1 (13)	4 (27)
Other	4 (57)	7 (88)	11 (73)

BMI: body mass index.

Results

Twelve men and three women participated in this study, with a mean age of 65 years (SD 8) (Table 1). All patients had diagnosed and treated coronary heart disease. Thirteen were referred following a revascularisation procedure for atherosclerosis and included 12 following a percutaneous intervention and one following coronary artery bypass surgery. Two participants with coronary heart disease did not require a revascularisation intervention and were managed medically, one of whom had mild left ventricular dysfunction. There were no statistically significant differences between group A and group B in terms of age, height, weight, body mass index, and waist or hip circumference.

All participants completed a symptom-limited exercise test and two ISWTs on their usual medications within one week, with the elapsed time ranging from one to six days. The results of the exercise tests are shown in Table 2. There were no interaction effects between the exercise test results and the order of the tests ($F(2,26)=1.283$; $P=0.294$). As there were no significant between-group differences or interaction effects between the groups (group A and group B) only data for the combined sample ($n=15$) will be presented.

All correlations between the ISWT and the symptom-limited exercise test ranged from moderate to strong, with the correlation for the primary outcome measures ranging from 0.85 to 0.87 between the ISWT distance and the

Table 2. Symptom-limited exercise test and 10 m ISWT results for the sample and group A (symptom-limited exercise test completed first) and group B (ISWT completed first).

	Total (*n*=15)	Symptom-limited exercise test completed first (*n*=7)	ISWT completed first (*n*=8)	M_{diff} (SD_{diff})	95% CI
ISWT 1					
Distance (m)	522 (216)	566 (229)	484 (211)	82 (312)	−164, 328
Duration (min)	7.8 (2.0)	8.3 (2.0)	7.4 (2.0)	0.9 (2.8)	−1.4, 3.1
Peak HR (bpm)	111 (11)	115 (10)	108 (12)	7 (15)	−4.9, 19.2
SpO_2	97 (1)	97 (1)	97 (1)	−0.1 (1.7)	−1.5, 1.3
ISWT 2					
Distance (m)	535 (219)	577 (229)	499 (219)	78 (317)	−172, 328
Duration (min)	8.0 (1.9)	8.4 (2.0)	7.6 (1.9)	0.7 (2.8)	−1.4, 3.0
Peak HR (bpm)	114 (10)	117 (10)	112 (10)	5 (14)	−6.0, 15.7
SpO_2	97 (1)	97 (1)	98 (1)	−0.4 (1.3)	−1.4, 0.7
Symptom-limited exercise test					
Duration (mins)	7.6 (2.5)	8.4 (2.9)	6.9 (1.9)	1.5 (3.5)	−1.2, 4.2
Peak HR (bpm)	135 (10)	135 (10)	136 (10)	−0.2 (14)	−11.3, 10.9
SpO_2	95 (2)	95 (1)	95 (2)	0.4 (2.7)	−1.7, 1.6

ISWT 1: first 10 m incremental shuttle walk test; ISWT 2: second 10 m incremental shuttle walk test; M_{diff}: mean difference, SD_{diff}: standard deviation of the difference, CI: confidence interval, HR: heart rate, SpO_2: oxygen saturation measured by pulse oximetry.

Table 3. Coefficient of variation for primary and secondary outcome measures for the 10 m ISWT and the symptom-limited exercise test.

Relation between outcome measures	*r* (95% CI)	r^2	M_{diff} (SD_{diff})	s_{yx}	Intercept (*a*)	Slope (*b*) (95% CI)
Primary outcome measures						
SLET duration and ISWT 1 distance	0.87 (0.63, 0.95)	0.75	–	1.3	2.4	0.01 (0.007, 0.01)
SLET test duration and ISWT 1 duration	0.89 (0.62, 0.95)	0.74	0.24 (1.3)	1.3	−0.8	1.1 (0.70, 1.5)
SLET duration and ISWT 2 distance	0.85 (0.59, 0.95)	0.72	–	1.4	2.5	0.01 (0.006, 0.01)
SLET duration and ISWT 2 duration	0.84 (0.57, 0.94)	0.70	0.39 (1.4)	1.5	−0.9	1.1 (0.6, 1.5)
Secondary outcome measures						
SLET HR and ISWT 1 HR	0.70 (0.29, 0.89)	0.49	−24 (8)	7.1	68.0	0.6 (0.2, 1.0)
SLET HR and ISWT 2 HR	0.79 (0.47, 0.93)	0.63	−21 (6)	6.0	46.3	0.8 (0.4, 1.1)
SLET SpO_2 and ISWT 1 SpO_2	0.82 (0.52, 0.94)	0.67	2 (1)	1.1	−30.7	1.3 (0.7, 1.8)
SLET SpO_2 and ISWT 2 SpO_2	0.68 (0.26, 0.89)	0.47	2 (1)	1.4	−42.2	1.4 (0.5, 2.3)

r: Pearson's product-moment correlation coefficient; CI: confidence interval; r^2: coefficient of determination; M_{diff}: mean difference; SD_{diff}: standard deviation of the difference; S_{yx}: standard estimate of error; SLET: symptom-limited exercise test; ISWT 1: first 10 m incremental shuttle walk test; ISWT 2: second 10 m incremental shuttle walk test; HR: heart rate; SpO_2: oxygen saturation measured by pulse oximetry.

symptom-limited exercise test duration (Table 3). The correlation values were similar for both the first and second ISWT. The coefficient of variation (r^2) for all outcome measures showed at least 72% of the variation in the symptom-limited exercise test duration could be explained by the ISWT distance.

For both the first and second ISWT agreement was satisfactory, with all scores falling within the limits of agreement (Figure 1). The relationship between the symptom-limited exercise test duration and the ISWT distance walked is shown in Figure 2.

The predicted symptom-limited exercise test duration with group and individual confidence limits were calculated for the distance walked in the first and second ISWT and are shown in Table 4. For example, an individual who completed all 12 levels of the first ISWT could be predicted, with 95% confidence, to complete between 9.2 and 16.0 minutes of the symptom-limited exercise test when a Bruce protocol is followed. The predicted symptom-limited exercise test duration was similar whether one and two walk tests were completed.

Discussion

The results support the concurrent criterion validity of the ISWT in a cardiac rehabilitation population. This study provided evidence of a strong correlation between the duration of the symptom-limited exercise test and the distance walked in the ISWT. The Bland Altman plots, showing agreement between the duration of the symptom-limited

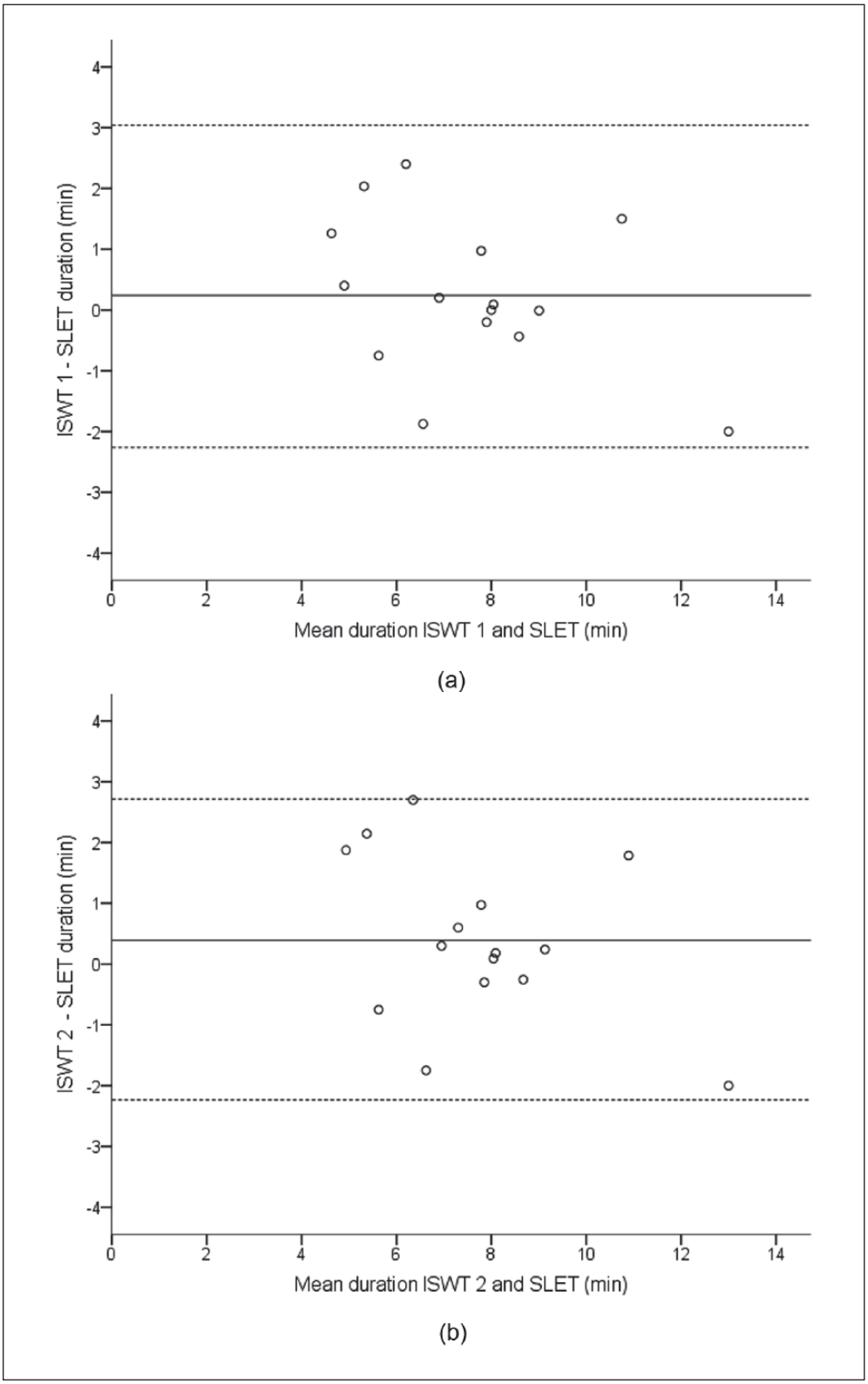


Figure 1. Bland Altman plots of the duration of the 10 m ISWT and the duration of the symptom-limited exercise test. (a) The first 10 m incremental shuttle walk test. (b) The second 10 m incremental shuttle walk test.
Unbroken line: mean difference; broken lines: limits of agreement calculated by $1.96 \times SD_{diff}$; SLET: symptom-limited exercise test; ISWT 1: first 10 m incremental shuttle walk test; ISWT 2: second 10 m incremental shuttle walk test.

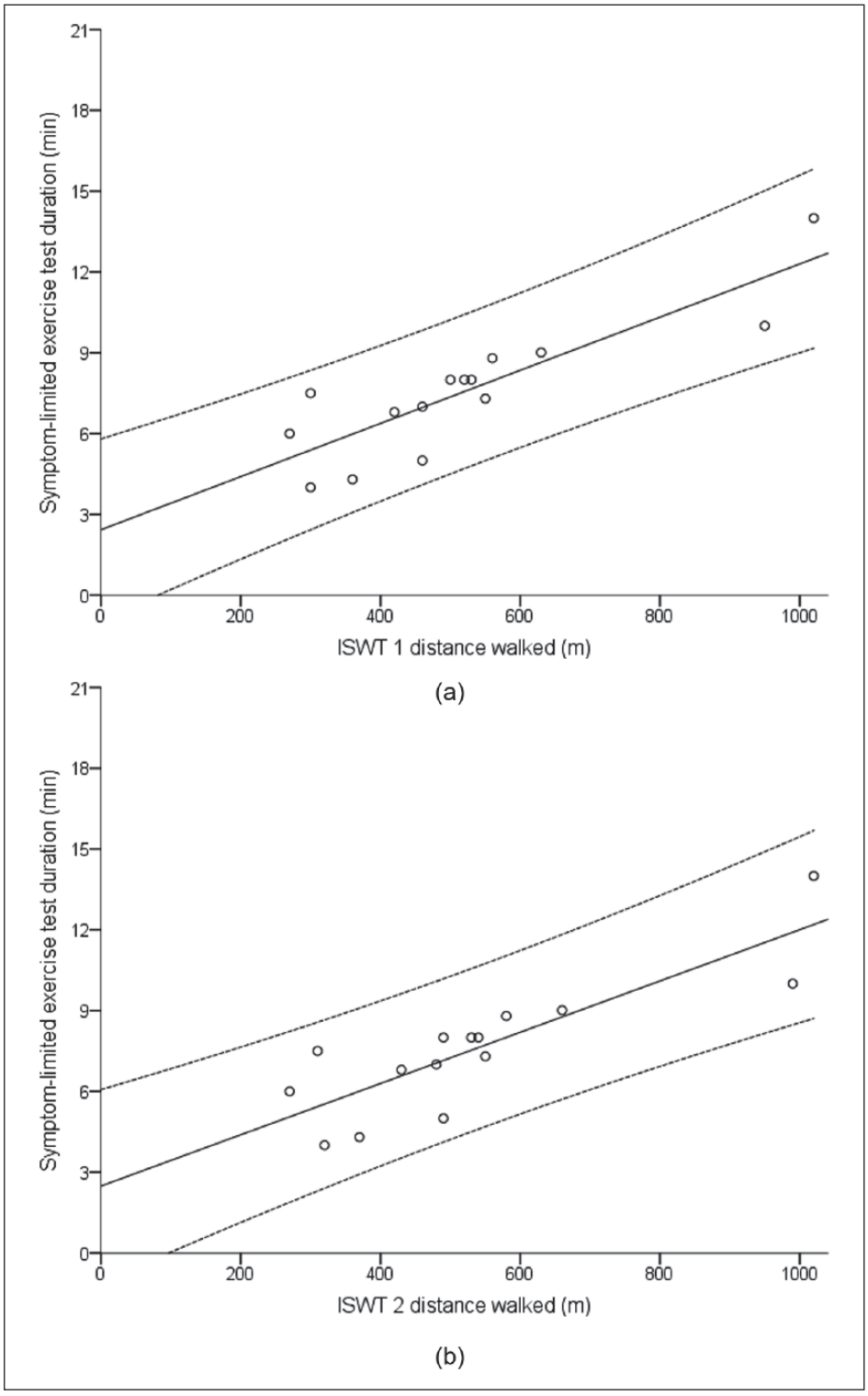


Figure 2. Linear relationship between 10 m ISWT distance walked and the symptom-limited exercise test duration is plotted with individual 95% confidence limits. (a) The linear relationship between the first ISWT and the symptom-limited exercise test was r^2=0.75, the linear equation to predict the duration of the symptom-limited exercise test was:
Symptom – limited exercise test duration (mins) = 2.43 + (0.01 × ISWT1 distance (m)) with *r*=0.87, 95% CI [0.63, 0.95], standard estimate of error of 1.3. (b) The linear relationship between the second ISWT and the symptom-limited exercise test was r^2=0.72, the linear equation to predict the duration of the symptom-limited exercise test was:
Symptom – limited exercise test duration (mins) = 2.48 + (0.01 × ISWT2 distance (m)), *r*=0.85, 95% CI [0.59, 0.95], standard estimate of error of 1.4.
ISWT 1: first 10 m incremental shuttle walk test; ISWT 2: second 10 m incremental shuttle walk test.

Table 4. Prediction of symptom-limited exercise test duration from ISWT distance with 95% CIs for the group and individual.

ISWT 1 distance (m)	Predicted duration symptom-limited exercise test (minutes)		
	Estimated duration	Group 95% CI	Individual 95% CI[a]
0	2.4	0.5, 4.3	0, 5.8
30	2.7	0.9, 4.5	0, 6
70	3.1	1.4, 4.8	0, 6.4
120	3.6	2.0, 5.2	0.4, 6.8
180	4.2	2.8, 5.6	1.1, 7.3
250	4.9	3.7, 6.1	1.9, 7.9
330	5.7	4.7, 6.7	2.7, 8.7
420	6.6	5.8, 7.4	3.7, 9.5
520	7.6	6.9, 8.3	4.7, 10.5
630	8.7	7.9, 9.5	5.8, 11.6
750	9.9	8.9, 11.0	6.9, 12.9
880	11.2	9.8, 12.6	8.1, 14.3
1020	12.6	10.7, 14.5	9.2, 16.0
ISWT 2 distance (m)			
0	2.5	0.4, 4.6	0, 6.2
30	2.8	0.8, 4.8	0, 6.4
70	3.2	1.3, 5.1	0, 6.7
120	3.7	2.0, 5.4	0.2, 7.2
180	4.3	2.8, 5.9	0.9, 7.7
250	5.0	3.7, 6.3	1.7, 8.3
330	5.8	4.7, 6.9	2.6, 9.0
420	6.7	5.8, 7.6	3.6, 9.8
520	7.7	6.9, 8.5	4.6, 10.8
630	8.8	8.0, 9.6	5.7, 11.9
750	10.0	8.9, 11.1	6.8, 13.2
880	11.3	9.8, 12.8	8.0, 14.6
1020	12.7	10.8, 14.6	9.1, 16.3

[a]If the lower band confidence limit was calculated as less than 0 it was denoted as 0.
ISWT 1: first 10 m incremental shuttle walk test; ISWT 2: second 10 m incremental shuttle walk test; CI: confidence interval.

exercise test and the duration of the ISWT, supported these results. The results supported the concurrent criterion validity of the ISWT, and demonstrated that the two exercise tests showed consistency in ranking patients according to physical fitness and functional capacity. The results did not improve when a second ISWT was performed, further supporting previous research that a single ISWT may be sufficient in this population.[13]

The correlations reported in the current study for patients with coronary heart disease were stronger than previously published correlations between symptom-limited exercise test duration and ISWT distance that ranged from 0.54 to 0.68 in patients with chronic heart failure.[18,20,22] In addition, the current study reported strong correlations ranging from 0.69 to 0.79 for peak heart rate achieved in the symptom-limited exercise test and ISWT, compared with earlier reports of moderate correlations.[18,20,21] The current study used the Bruce protocol for the symptom-limited exercise test, whereas the three earlier studies used alternative protocols. Differences in the patient population may also account for differences. Participants in the current study were any ambulant patient with treated and stable coronary heart disease. The earlier studies used more homogenous groups of patients with chronic heart failure. It is possible that patients with chronic heart failure respond differently to the demands of an incremental exercise test compared with patients from a cardiac rehabilitation group. In the current study, the ISWT mean distance walked was greater than previously reported studies,[18,20] and may reflect the stage of recovery.

The results do not support the predictive criterion validity of the ISWT in a cardiac rehabilitation group when functional outcome measures such as distance walked and duration of the symptom-limited exercise test are used. This is the first study to present information on the accuracy of the predictive criterion validity of the ISWT. While a linear regression equation was presented, the 95% CIs around the slope (*b*) were wide and made the equation difficult to interpret. In addition, the clinically relevant 95% confidence limits for individuals, or the prediction interval, remained wide, making it difficult in an individual to predict with accuracy the duration of the

Bruce protocol from the distance walked in the ISWT. The relatively wide individual confidence limits may have been a result of the error contribution from both exercise tests. However, repeating the ISWT did not reduce the CIs, and the effect of repeating the symptom-limited exercise test is unknown. Despite justification of the sample size, it was small compared with the recommendations included in the COSMIN checklist[41] and future research could use the model presented in this paper with an increased in the number of participants. An increase in the number of participants should provide a more stable regression equation and may reduce the 95% CIs and prediction limits.[34] Clinically, the ability to predict the symptom-limited exercise test duration based on the ISWT, or in reverse, to predict the ISWT distance based on the symptom-limited exercise test duration is worth exploring as it provides an important means of translating meaning between the laboratory and clinical environment.

While we used a symptom-limited exercise test protocol considered a gold standard in exercise testing,[42] this study relied on functional outcome measures such as distance ambulated and duration of test, and not the measurement of maximum oxygen consumption. Cardiopulmonary exercise testing, with specific measurement of maximum oxygen consumption and other aerobic measures, is not a practical and available option for many patients with cardiac disease who live outside of metropolitan regions in Australia and other countries.[43,44] In the city where this research was conducted the measurement of oxygen consumption during exercise testing for cardiac patients was not available, and patients would need to travel at least 150 km to access these medical tests. Future research could include monitoring of maximum oxygen consumption using a portable oxygen analyser.

There were no adverse responses in testing in this sample and we were not able to determine cut-offs or thresholds in performance in the ISWT for people with cardiac disease. A larger sample size may be able to identify cut-offs or thresholds in performance in the ISWT for people with cardiac disease. It would be clinically useful for a clinician to have available cut-off points, for example, to identify when a patient needs a review by his or her cardiologist.

Implications

The ISWT when compared with a symptom-limited exercise test is easy to implement, requires fewer resources, does not require expensive equipment, and when used as an estimate for physical fitness and functional capacity does not need a medical doctor to supervise and can be conducted within cardiac rehabilitation without the need for referral. Our results demonstrate that the functional results of the ISWT measure a similar construct as the symptom-limited exercise test, and the two tests show consistency in ranking of participants in terms of physical fitness and functional capacity. However, in absolute terms, there is emerging evidence to suggest that the ISWT cannot replace the symptom-limited exercise test. In particular, using the regression equation generated from this small sample, the distance walked in the ISWT cannot be used to predict accurately individual functional performance on the symptom-limited exercise test. While the ISWT test should not be aimed at replacing the symptom-limited exercise test, it may provide a suitable, cost and time effective clinical alternative of testing physical fitness and functional capacity for patients attending cardiac rehabilitation.

Implications for practice

- The 10 m incremental shuttle walk test is a practical and feasible field exercise test for use in cardiac rehabilitation when a single test is performed.
- The functional results of the 10 m incremental shuttle walk test measure a similar construct as the functional results of a symptom-limited exercise test.
- The 10 m incremental shuttle walk test cannot be used as a surrogate to predict individual symptom-limited exercise test results with accuracy.

Declaration of conflicting interest

The authors declare that there is no conflict of interest.

Funding

This study was supported by an Australian Government Research Training Program Scholarship.

References

1. American College of Sports Medicine. *ACSM's health-related physical fitness assessment manual*, 3rd edn. Baltimore, MD: Lippincott Williams & Wilkins, 2010.
2. Balady GJ, Williams MA, Ades PA, et al. Core components of cardiac rehabilitation/secondary prevention programs: 2007 update. A scientific statement from the American Heart Association Exercise, Cardiac Rehabilitation, and Prevention Committee, the Council on Clinical Cardiology; the Councils on Cardiovascular Nursing, Epidemiology and Prevention, and Nutrition, Physical Activity, and Metabolism; and the American Association of Cardiovascular and Pulmonary Rehabilitation. *Circulation* 2007; 115: 2675–2682.
3. Palange P, Ward SA, Carlsen K-H, et al. Recommendations on the use of exercise testing in clinical practice. *Eur Respir J* 2007; 29: 185–209.
4. Piepoli MF, Corra U, Benzer W, et al. Secondary prevention through cardiac rehabilitation: from knowledge to implementation. A position paper from the Cardiac Rehabilitation

Section of the European Association of Cardiovascular Prevention and Rehabilitation. *Eur J Cardiovasc Prev Rehabil* 2010; 17: 1–17.
5. Thompson PD, Arena R, Riebe D, et al. ACSM's new preparticipation health screening recommendations from ACSM's guidelines for exercise testing and prescription. *Curr Sports Med Rep* 2013; 12: 215–217.
6. Simms K, Myers C, Adams J, et al. Exercise tolerance testing in a cardiac rehabilitation setting: an exploratory study of its safety and practicality for exercise prescription and outcome data collection. *Baylor University Medical Centre Proceedings* 2007; 20: 344–347.
7. Taylor A, Bell J and Lough F. Cardiac rehabilitation and secondary prevention. In: Pryor JA and Prasad SA (eds). *Physiotherapy for respiratory and cardiac problems*, 4th edn. Edinburgh: Churchill Livingstone, 2010, pp. 470–494.
8. American Association of Cardiovascular and Pulmonary Rehabilitation. *Performance measure for improvement in functional capacity at completion of cardiac rehabilitation* [Internet]. Chicago, IL (USA): American Association of Cardiovascular and Pulmonary Rehabilitation; 2016 [cited 2017 April 11]. Available from: http://www.aacvpr.org/Portals/0/Program%20Certification/Performance%20Measures/Measures/CR%20Functional%20Capacity%20Performance%20Measure_10.24.16.pdf.
9. Goble AJ and Worcester MUC. *Best practice guidelines for cardiac rehabilitation and secondary prevention*. Melbourne, Australia: Heart Research Centre, on behalf of Department of Human Services Victoria, 1999.
10. New Zealand Guidelines Group and Heart Foundation. *Evidence-based best practice guideline: Cardiac rehabilitation*. Wellington: New Zealand Guideline Group, 2002.
11. Price KJ, Gordon BA, Bird SR, et al. A review of guidelines for cardiac rehabilitation exercise programmes: is there an international consensus? *Eur J Prev Cardiol* 2016; 23: 1715–1733.
12. Hanson LC, McBurney H and Taylor NF. The retest reliability of the six minute walk test in patients referred to a cardiac rehabilitation programme. *Physiother Res Intl* 2012; 17: 55–61.
13. Hanson LC, Taylor NF and McBurney H. The 10 m incremental shuttle walk test is a highly reliable field exercise test for patients referred to cardiac rehabilitation: a retest reliability study. *Physiotherapy* 2016; 102: 243–248.
14. Singh SJ, Puhan MA, Andrianopoulos V, et al. An official systematic review of the European Respiratory Society/American Thoracic Society: measurement properties of field walking tests in chronic respiratory disease. *Eur Respir J* 2014; 44: 1447–1478.
15. Solway S, Brooks D, Lacasse Y, et al. A qualitative systematic overview of the measurement properties of functional walk tests used in the cardiorespiratory domain. *Chest* 2001; 119: 256–270.
16. Jolly K, Taylor RS, Lip GYH, et al. Reproducibility and safety of the incremental shuttle walking test for cardiac rehabilitation. *Int J Cardiol* 2008; 125: 144–145.
17. Mokkink LB, Terwee CB, Patrick DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *J Clin Epidemiol* 2010; 63: 737–745.
18. Green DJ, Watts K, Rankin S, et al. A comparison of the shuttle and 6 minute walking tests with measured peak oxygen consumption in patients with heart failure. *J Sci Med Sport* 2001; 4: 292–300.
19. Keell SD, Chambers JS, Francis DP, et al. Shuttle-walk test to assess chronic heart failure. *Lancet* 1998; 352: 705.
20. Lewis ME, Newal C, Townend JN, et al. Incremental shuttle walk test in the assessment of patients for heart transplantation. *Heart* 2001; 86: 183–187.
21. Morales FJ, Martinez A, Mendez M, et al. A shuttle walk test for assessment of functional capacity in chronic heart failure. *Am Heart J* 1999; 138: 291–298.
22. Pulz C, Diniz RV, Alves AN, et al. Incremental shuttle and six minute walking tests in the assessment of functional capacity in chronic heart failure. *Can J Cardiol* 2008; 24: 131–135.
23. Fowler SJ, Singh SJ and Revill S. Reproducibility and validity of the incremental shuttle walking test in patients following coronary artery bypass surgery. *Physiotherapy* 2005; 91: 22–27.
24. Mandic S, Walker R, Stevens E, et al. Estimating exercise capacity from walking tests in elderly individuals with stable coronary artery disease. *Disabil Rehabil* 2013; 35: 1853–1858.
25. Singh SJ, Morgan MD, Scott S, et al. Development of a shuttle walking test of disability in patients with chronic airways obstruction. *Thorax* 1992; 47: 1019–1024.
26. Green P, Cohen DJ, Genereux P, et al. Relation between six-minute walk test performance and outcomes after transcatheter aortic valve implantation (from the PARTNER trial). *Am J Cardiol* 2013; 112: 700–706.
27. Chatterjee S, Sengupta S, Nag M, et al. Cardiopulmonary exercise testing: a review of techniques and applications. *Anesth Clin Res* 2013; 4: 340–346.
28. Bruce RA, Blackmon JR, Jones JW, et al. Exercise testing in adult normal subjects and cardiac patients. *Pediatrics* 1963; 32: 742–756.
29. Bruce RA and Hornsten TR. Exercise stress testing in evaluation of patients with ischemic heart disease. *Prog Cardiovasc Dis* 1969; 11: 371–390.
30. Howell DC. *Statistical methods for psychology*, 8th edn. Australia: Wadsworth, Cengage Learning, 2012.
31. Portney LG and Watkins MP. *Foundations of clinical research: Applications to practice*, 2nd edn. Upper Saddle River, New Jersey: Prentice Hall Health, 2000.
32. Terwee CB, Bot SDM, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol* 2007; 60: 34–42.
33. Norman GR and Streiner DL. *Biostatistics: the bare essentials*, 3rd edn. Hamilton, Ontario: B.C. Decker, 2008.
34. Tabachnick BG and Fidell LS. *Using multivariate statistics*, 6th edn. Boston: Pearson Education, 2007.
35. Bruce RA, Kusumi F and Hosmer D. Maximal oxygen intake and nomographic assessment of functional aerobic impairment in cardiovascular disease. *Am Heart J* 1973; 85: 546–562.
36. Benhorin J, Pinsker G, Moriel M, et al. Ischemic threshold during two exercise testing protocols and during ambulatory electrocardiographic monitoring. *J Am Coll Cardiol* 1993; 22: 671–677.

37. Steele B. Timed walking tests of exercise capacity in chronic cardiopulmonary illness. *J Cardiopulm Rehabil Prevent* 1996; 16: 25–33.
38. American College of Sports Medicine. *ACSM's guidelines for exercise testing and prescription*, 8th edn. Baltimore: Williams & Wilkins, 2010.
39. Portney LG and Watkins MP. *Foundations of clinical research: Applications to practice*, 3rd edn. Upper Saddle River, New Jersey: Pearson Prentice Hall, 2009.
40. Altman DG and Gardner MJ. Regression and correlation. In: Altman DG, Machin D,Bryant TN and Gardner MJ (eds). *Statistics with confidence*, 2nd edn. London: British Medical Journal Books, 2000, pp. 73–92.
41. Terwee CB, Mokkink LB, Knol DL, et al. Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. *Quality of Life Research* 2012; 21: 651–657.
42. Noonan V and Dean E. Submaximal exercise testing: clinical application and interpretation. *Phys Ther* 2000; 80: 782–807.
43. Chatterjee S, Sengupta S, Nag M, et al. Cardiopulmonary exercise testing: a review of techniques and applications. *J Anesth Clin Res* 2013; 4:340.
44. National Guideline Centre UK. *Preoperative tests (update): Routine preoperative tests for elective surgery*. London: National Institute for Health and Care Excellence (UK), 2016.

[11] Walter S, Eliasziw M, Donner A. Sample size and optimal designs for reliability studies. Stat Med 1998;17:101–10.
[12] American College of Sports Medicine. ACSM's guidelines for exercise testing and prescription. 8th ed. Baltimore: Williams & Wilkins; 2010.
[13] Portney LG, Watkins MP. Foundations of clinical research: applications to practice. 2nd ed. Upper Saddle River, NJ: Prentice Hall Health; 2000. p. 565.
[14] de Vet HCW, Terwee CB, Mokkink LB, Knol DL. Measurement in Medicine. Cambridge: Cambridge University Press; 2011. p. 111–3.
[15] Taylor NF, Dodd KJ, Graham HK. Test-retest reliability of hand-held dynamometric strength testing in young people with cerebral palsy. Arch Phys Med Rehabil 2004;85(1):77–80.
[16] Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986;327(8476):307–10.
[17] Hanson LC, McBurney H, Taylor NF. The retest reliability of the six minute walk test in patients referred to a cardiac rehabilitation programme. Physiother Res Int 2012;17:55–61.
[18] Wu G, Sanderson B, Bittner C. The 6-minute walk test: how important is the learning effect? Am Heart J 2003;146:119–24.
[19] Robinson HJ, Samani NJ, Singh SJ. Can low risk cardiac patients be 'fast tracked' to Phase IV community exercise schemes for cardiac rehabilitation? A randomised controlled trial. Int J Cardiol 2011;146:159–63.
[20] Pepera G, Cardoso F, Taylor MJD, Peristeropoulos A, Sandercock GRH. Predictors of shuttle walking test performance in patients with cardiovascular disease. Physiotherapy 2013;99:317–22.
[21] Terwee CB, Mokkink LB, Knol DL, Ostelo RWJG, Bouter LM, de Vet HCW. Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. Qual Life Res 2012;21(4):651–717.

Appendix 2: Human Research Ethic Committee Approvals

Ethics approvals were granted from all relevant ethics committees for each study.

Chapter 2

Faculty Human Ethics Committee, La Trobe University: FHEC 6/174

Human Research Ethics Committee, Bendigo Health: HREC 24/2006

Chapter 3

Faculty Human Ethics Committee, La Trobe University: FHEC 07/99

La Trobe University Human Ethics Committee: HEC No. 10-082

Human Research Ethics Committee, Bendigo Health: HREC 1/2007

St John of God Health Care Ethics Committee: Ref No. 438

Chapters 4, 5 and 6

Chapter 4, 5 and 6 were considered in the same ethics approval statement.

Human Research Ethics Committee, Bendigo Health: HREC 1/2007

St John of God Health Care Ethics Committee: Ref No. 438

La Trobe University
Faculty of Health Sciences
MEMORANDUM

TO: Dr Helen McBurney
Ms Deborah Ludeman
School of Physiotherapy

SUBJECT: *Reference:* **FHEC06/174**

Student or Other Investigator: Lisa Hanson

Title: **The reproducibility of the six minute walk test in patients referred to cardiac rehabilitation**

DATE: 5 February 2007

The Faculty Human Ethics Committee (FHEC) has considered and approved the above project. You may now proceed.

Please note that Informed Consent forms need to be retained for a minumum of 7 years. Please ensure that each participant retains a copy of the Informed Consent form. Researchers are also required to retain a copy of all Informed Consent forms separately from the data. The data must be retained for a period of 5 years.

Please note that any modification to the project must be submitted in writing to FHEC for approval. You are required to provide an annual report (where applicable) and/or a final report on completion of the project. A copy of the progress/final report can be downloaded from the following website:
www.latrobe.edu.au/rgso/forms-resources/forms/ethic-prog-final.rtf.

Please return the completed form to Ms Barbara Doherty, Secretary, FHEC, Faculty of Health Sciences Office, La Trobe University, Victoria 3086.

A copy of this memorandum is enclosed for you to forward to the student(s) concerned.

Barbara Doherty
Secretary
Faculty Human Ethics Committee
Faculty of Health Sciences

La Trobe University
Faculty of Health Sciences
MEMORANDUM

TO: Assoc Prof Helen McBurney — School of Physiotherapy
Prof Nicholas Taylor

SUBJECT: *Reference:* **FHEC07/99**

Student or Other Investigator: Lisa Hanson, Deborah Ludeman (Bendigo Health Care Group)

Title: **A comparison of the incremental shuttle walk test and the six minute walk test in patients referred to cardiac rehabilitation**

DATE: 27 September 2007

The Faculty Human Ethics Committee (FHEC) has considered and approved the above project. You may now proceed.

Please note that Informed Consent forms need to be retained for a minimum of 7 years. Please ensure that each participant retains a copy of the Informed Consent form. Researchers are also required to retain a copy of all Informed Consent forms separately from the data. The data must be retained for a period of 5 years.

Please note that any modification to the project must be submitted in writing to FHEC for approval. You are required to provide an annual report (where applicable) and/or a final report on completion of the project. A copy of the progress/final report can be downloaded from the following website:
www.latrobe.edu.au/rgso/forms-resources/forms/ethic-prog-final.rtf.

Please return the completed form to The Secretary, FHEC, Faculty of Health Sciences Office, La Trobe University, Victoria 3086.

A copy of this memorandum is enclosed for you to forward to the student(s) concerned.

Natalie Humphries
Secretary
Faculty Human Ethics Committee
Faculty of Health Sciences

RESEARCH SERVICES

MEMORANDUM

To: **Prof Nicholas Taylor, School of Physiotherapy, Faculty of Health Sciences**
Ms Lisa Hanson, School of Physiotherapy, Faculty of Health Sciences

From: **Secretary, La Trobe University Human Ethics Committee**

Subject: **Review of Human Ethics Committee Application No. 10-082**

Title: **The Measurement Properties of the Modified 10m Incremental Shuttle Walk Test when used in Cardiac Rehabilitation**

Date: **21 December 2010**

Thank you for your recent correspondence in relation to the research project referred to above. The project has been assessed as complying with the *National Statement on Ethical Conduct in Human Research*. I am pleased to advise that your project has been granted ethics approval and you may commence the study.

The project has been approved from the date of this letter until 30 September 2011.

Please note that your application has been reviewed by a sub-committee of the University Human Ethics Committee (UHEC) to facilitate a decision about the study before the next Committee meeting. This decision will require ratification by the full UHEC at its next meeting and the UHEC reserves the right to alter conditions of approval or withdraw approval. You will be notified if the approval status of your project changes.

The following standard conditions apply to your project:

- **Limit of Approval.** Approval is limited strictly to the research proposal as submitted in your application while taking into account any additional conditions advised by the UHEC.

- **Variation to Project.** Any subsequent variations or modifications you wish to make to your project must be formally notified to the UHEC for approval in advance of these modifications being introduced into the project. This can be done using the appropriate form: *Ethics - Application for Modification to Project* which is available on the Research Services website at *http://www.latrobe.edu.au/research-services/ethics/human.htm*. If the UHEC considers that the proposed changes are significant, you may be required to submit a new application form for approval of the revised project.

- **Adverse Events.** If any unforeseen or adverse events occur, including adverse effects on participants, during the course of the project which may affect the ethical acceptability of the project, the Chief Investigator must immediately notify the UHEC Secretary on telephone (03) 9479 1443. Any complaints about the project received by the researchers must also be referred immediately to the UHEC Secretary.

- **Withdrawal of Project.** If you decide to discontinue your research before its planned completion, you must advise the UHEC and clarify the circumstances.

If you have any queries on the information above or require further clarification please contact me through Research Services on telephone (03) 9479-3589, or e-mail at: humanethics@latrobe.edu.au.

Ms. Lynda Boldt

Administrative Officer – Research
Acting Secretariat – University Human Ethics Committee
Research Compliance Unit
Research Services | La Trobe University | Bundoora 3086
T: 03 9479 3589 | F: 03 9479 1464 | E: l.boldt@latrobe.edu.au | http://latrobe.edu.au/research-services/

RESEARCH SERVICES

MEMORANDUM

To: **Professor Nicholas Taylor, School of Physiotherapy, Faculty of Health Sciences**
Ms Lisa Hanson, School of Physiotherapy, Faculty of Health Sciences

From: **Secretary, La Trobe University Human Ethics Committee**

Subject: **Review of Human Ethics Committee Application No. 10-082**

Title: **The Measurement Properties of the Modified 10m Incremental Shuttle Walk Test when used in Cardiac Rehabilitation**

Date: **25 August 2011**

Thank you for submitting your modification request for ethics approval to the La Trobe University Human Ethics Committee (UHEC) for the project referred to above. The UHEC Chair has reviewed and approved the following modification:

- **Extension of ethics approval granted until 31 December 2012.**

Please note that your application has been reviewed by a sub-committee of the UHEC in the interest of facilitating a decision on your application before the next committee meeting. The decision to approve your project will need to be ratified by the full UHEC and consequently approval for your project may be withdrawn or conditions of approval altered. However, your project may commence prior to ratification of the approval decision. You will be notified if the approval status of your project is altered.

The following standard conditions apply to your project:

- **Limit of Approval.** Approval is limited strictly to the research proposal as submitted in your application while taking into account any additional conditions advised by the UHEC.

- **Variation to Project.** Any subsequent variations or modifications you wish to make to your project must be formally notified to the UHEC for approval in advance of these modifications being introduced into the project. This can be done using the appropriate form: *Ethics - Application for Modification to Project* which is available on the Research Services website at *http://www.latrobe.edu.au/research-services/ethics/HEC_human.htm*. If the UHEC considers that the proposed changes are significant, you may be required to submit a new application form for approval of the revised project.

- **Annual Progress Reports.** If your project continues for more than 12 months, you are required to submit an *Ethics - Progress/Final Report Form* annually, on or just prior to 12 February. The form is available on the Research Services website (see above address). Failure to submit a Progress Report will mean approval for this project will lapse. An audit may be conducted by the UHEC at any time.

- **Adverse Events.** If any unforeseen or adverse events occur, including adverse effects on participants, during the course of the project which may affect the ethical acceptability of the project, the Chief Investigator must immediately notify the UHEC Secretary on telephone (03) 9479 1443. Any complaints about the project received by the researchers must also be referred immediately to the UHEC Secretary.

- **Withdrawal of Project.** If you decide to discontinue your research before its planned completion, you must advise the UHEC and clarify the circumstances.

- **Final Report.** A Final Report (see above address) is required within six months of the completion of the project or by **30 June 2013.**

- **Annual Progress Reports.** If your project continues for more than 12 months, you are required to submit an *Ethics - Progress/Final Report Form* annually, **on or just prior to 12 February**. The form is available on the Research Services website (see above address). Failure to submit a Progress Report will mean approval for this project will lapse. An audit may be conducted by the UHEC at any time.

- **Final Report.** A Final Report (see above address) is required within six months of the completion of the project or by **31 March 2012.**

If you have any queries on the information above or require further clarification please contact me through Research Services on telephone (03) 9479-1443, or e-mail at:

humanethics@latrobe.edu.au.

On behalf of the University Human Ethics Committee, best wishes with your research!

Ms Barbara Doherty
Administrative Officer (Research Ethics)
University Human Ethics Committee
Research Compliance Unit / Research Services
La Trobe University Bundoora, Victoria 3086
P: (03) 9479 – 1443 / F: (03) 9479 - 1464
http://www.latrobe.edu.au/research-services/ethics/

RESEARCH SERVICES

MEMORANDUM

To:	**Professor Nicholas Taylor, School of Physiotherapy, Faculty of Health Sciences** **Ms Lisa Hanson, School of Physiotherapy, Faculty of Health Sciences**
From:	**Acting Secretary, La Trobe University Human Ethics Committee**
Subject:	**Review of Human Ethics Committee Application No. 10-082**
Title:	**The Measurement Properties of the Modified 10m Incremental Shuttle Walk Test when used in Cardiac Rehabilitation**
Date:	**21 December 2012**

Thank you for submitting your modification request for ethics approval to the La Trobe University Human Ethics Committee (UHEC) for the project referred to above. The UHEC Chair has reviewed and approved the following modifications:

- **Extension of ethics approval granted until 31 December 2013.**

Please note that your application has been reviewed by a sub-committee of the UHEC in the interest of facilitating a decision on your application before the next committee meeting. The decision to approve your project will need to be ratified by the full UHEC and consequently approval for your project may be withdrawn or conditions of approval altered. However, your project may commence prior to ratification of the approval decision. You will be notified if the approval status of your project is altered.

The following standard conditions apply to your project:

- **Limit of Approval.** Approval is limited strictly to the research proposal as submitted in your application while taking into account any additional conditions advised by the UHEC.

- **Variation to Project.** Any subsequent variations or modifications you wish to make to your project must be formally notified to the UHEC for approval in advance of these modifications being introduced into the project. This can be done using the appropriate form: *Ethics - Application for Modification to Project* which is available on the Research Services website at *http://www.latrobe.edu.au/research-services/ethics/HEC_human.htm*. If the UHEC considers that the proposed changes are significant, you may be required to submit a new application form for approval of the revised project.

- **Annual Progress Reports.** If your project continues for more than 12 months, you are required to submit an *Ethics - Progress/Final Report Form* annually, on or just prior to 12 February. The form is available on the Research Services website (see above address). Failure to submit a Progress Report will mean approval for this project will lapse. An audit may be conducted by the UHEC at any time.

- **Adverse Events.** If any unforeseen or adverse events occur, including adverse effects on participants, during the course of the project which may affect the ethical acceptability of the project, the Chief Investigator must immediately notify the UHEC Secretary on telephone (03) 9479 1443. Any complaints about the project received by the researchers must also be referred immediately to the UHEC Secretary.

- **Withdrawal of Project.** If you decide to discontinue your research before its planned completion, you must advise the UHEC and clarify the circumstances.

- **Final Report.** A Final Report (see above address) is required within six months of the completion of the project or by **30 June 2014.**

If you have any queries on the information above or require further clarification please contact me through Research Services on telephone (03) 9479-1443, or e-mail at: humanethics@latrobe.edu.au.

Ms Barbara Doherty
Administrative Officer (Research Ethics)
University Human Ethics Committee

Research Services La Trobe University Bundoora, Victoria 3086
P: (03) 9479 – 1443 F: (03) 9479 - 1464
http://www.latrobe.edu.au/research-services/ethics/

BENDIGO HEALTH
regional service, local care

CONFIDENTIAL

ABN 26 875 445 912

PO Box 126
Bendigo 3552

Human Research Ethics Committee
Phone: (03) 5454 6407
Fax: (03) 5454 6420

http: *www.bendigohealth.org.au/HREC.shtml*

Tuesday, 19 December 2006

Ms Lisa Hanson
Senior Physiotherapist
Bendigo Health
PO Box 126
Bendigo Vic 3552

Dear Lisa

Re: Study Title: The Reproducibility of the Six Minute Walk Test in Participants Referred to Cardiac Rehabilitation

HREC Reference Number: 24/2006

I am pleased to advise you that the Human Research Ethics Committee of the Bendigo Health Care Group has approved the above project. The project has been approved for the period 15/12/2006-01/11/2007.

Would you please note that the following standard conditions apply:

a. *Limit of Approval*: approval is limited strictly to the research proposal as submitted in your application. In addition, approval by the HREC *does not* guarantee that an individual BHCG unit or service will agree to provide resources or support to your research. Such assistance will need to be negotiated separately.
b. *Variation to Project*: any subsequent variations or modifications you might wish to make to your project must be notified formally to the committee for further consideration and approval. If the committee considers that the proposed changes are significant, you may be required to submit a new application for approval of the revised project.
c. *Incidents of Adverse Effects*: researchers must report immediately to the committee anything which might affect the ethical acceptance of the protocol including adverse effects on subjects or unforeseen events that might affect continued ethical acceptability of the project.
d. *Progress Reporting*: please be aware that the Human Research Ethics Committee requires all researchers to submit a report on each of their projects yearly, or at the conclusion of the project if it continues for less than a year. Failure to submit a progress report may mean approval for this project will lapse. **The first and final progress report for this project is due on 01/11/2007.**
e. Auditing: all projects may be subject to audit by members of the committee.

If you have any further queries on these matters, or require additional information, please contact me on 5454 – 6407, or e-mail: acrombie@bendigohealth.org.au. Human Research Ethics Committee information and application forms are available on the Committee's website, http://www.bendigohealth.org.au/HREC.shtml.

Please quote the HREC reference number and the title of the project in any future correspondence. On behalf of the committee, I wish you well in your research.

Yours Sincerely

Angela Crombie
Secretary
Human Research Ethics Committee
Bendigo Health Care Group

www.bendigohealth.org.au

Bendigo
Castlemaine
Cohuna
Echuca
Kerang
Kyneton
Kyabram
Maryborough
Mildura
Ouyen
Rochester
Sea Lake
Swan Hill
Wedderburn

BENDIGO HEALTH
regional service, local care

CONFIDENTIAL

ABN 26 875 445 912

www.bendigohealth.org.au

PO Box 126
Bendigo 3552

Human Research Ethics Committee
Phone: (03) 5454 6412
Fax: (03) 5454 6420

http: www.bendigohealth.org.au/HREC.

31st August 2007

Ms Lisa Hanson
Senior Physiotherapist
Bendigo Health
PO Box 126
Bendigo VIC 3552

Dear Lisa,

Re: Study Title: A Comparison of the Incremental Shuttle Walk Test and the Six Minute Walk Test in Patients Referred to Cardiac Rehabilitation

HREC Reference Number: 1/2007

I am pleased to advise you that the Human Research Ethics Committee of the Bendigo Health Care Group has approved the above project.

The project has been approved for the period 01/09/2007-01/09/2008.

Would you please note that the following standard conditions apply:

a. *Limit of Approval*: approval is limited strictly to the research proposal as submitted in your application. In addition, approval by the HREC *does not* guarantee that an individual BHCG unit or service will agree to provide resources or support to your research. Such assistance will need to be negotiated separately.
b. *Variation to Project*: any subsequent variations or modifications you might wish to make to your project must be notified formally to the committee for further consideration and approval. If the committee considers that the proposed changes are significant, you may be required to submit a new application for approval of the revised project.
c. *Incidents of Adverse Effects*: researchers must report immediately to the committee anything which might affect the ethical acceptance of the protocol including adverse effects on subjects or unforeseen events that might affect continued ethical acceptability of the project.
d. *Progress Reporting*: please be aware that the Human Research Ethics Committee requires all researchers to submit a report on each of their projects yearly, or at the conclusion of the project if it continues for less than a year. Failure to submit a progress report may mean approval for this project will lapse. **The first and final progress report for this project is due on 01/09/2008.**
e. Auditing: all projects may be subject to audit by members of the committee.

If you have any further queries on these matters, or require additional information, please contact me on 5454 – 6412, or e-mail: SAMcCarthy@bendigohealth.org.au. Human Research Ethics Committee information and application forms are available on the Committee's website, http://www.bendigohealth.org.au/HREC.

On behalf of the committee, I wish you well in your research.

Yours Sincerely

Sally McCarthy
Secretary
Human Research Ethics Committee
Bendigo Health Care Group

ST JOHN OF GOD
HEALTH CARE

HOSPITALS DIAGNOSTICS OUTREACH

5 August 2010

Ethics Committee
Lvl 3 St John of God House
177-179 Cambridge Street
Subiaco WA 6008
PO Box 14
Subiaco WA 6904
Tel: 61 8 9382 6940
Fax: 61 8 9382 6037
www.sjog.org.au

Dr Nicholas Taylor
School of Physiotherapy
Faculty of Health Science
La Trobe University
BUNDOORA VIC 3086

Dear Dr Taylor

Re: The measurement properties of the modified 10m incremental shuttle walk test when used in cardiac rehabilitation *(Our ref No: 438)*

Thank you for forwarding the above study for the review of the St John of God Health Care Ethics Committee ("the Committee").

I am pleased to advise that the Committee at its meeting on 5 August 2010, granted ethical approval of your study as per the National Health and Medical Research Council's National Statement on Ethical Conduct in Human Research (2007) ("the National Statement"). Please find attached a signed and dated Committee membership list.

I now confirm final approval for your study to be conducted at St John of God Hospital in Subiaco ("the participating site").

The Committee is a Human Research Ethics Committee (HREC) that is constituted and operates in accordance with the National Statement. In line with the National Statement requirements, researchers need to keep the Committee and the institution (specifically, the participating site) promptly and regularly informed on the progress of their approved research including:

1. any serious, and suspected, unexpected serious adverse events, any unforeseen events, any *significant* protocol deviations or violations, any withholding or withdrawal of study approval by another HREC/institution, and any allegation or suspicion of research misconduct, that may affect continued ethical approval of the study.
2. any proposed changes to the research/research documentation as previously approved by the Committee, including any proposed study extensions.
3. when the study is completed, abandoned, terminated, suspended or withdrawn.

The Committee would also appreciate receiving *at a minimum* an *annual* study progress report as well as a final report on the study results and/or any subsequent publications.

I wish you well with your study.

Yours sincerely

Professor Con Michael
(as a delegate of St John of God Health Care)

Enc.
cc. Ms Lisa Hanson, School of Physiotherapy, Curtin University of Technology
cc. Ms Penny Limmer, Allied Health & Outpatient Rehabl. Unit Manager, SJG Hospital Bendigo

Appendix 3: Participant Information and Consent Forms

Each participant received a participant information and consent form specific to the enrolled study.

Chapters in this thesis that required a participant information and consent form: Chapter 2, Chapter 3, Chapter 4, Chapter 5 and 6.

Participant Information and Consent Form

Date: 1st October 2006
Site: Cardiac Rehabilitation; Physiotherapy Department, The Bendigo Hospital

Full Project Title: The Reproducibility of the Six Minute Walk Test in Patients Referred to Cardiac Rehabilitation**.**

Principal Researcher: Lisa Hanson (Senior Physiotherapist, PhD Student)

Principal Researcher: Dr Helen McBurney (Associate Professor, Supervisor)

Associate Researcher: Deborah Ludeman (Cardiac Rehabilitation Co-ordinator)

This Participant Information and Consent Form is 7 pages long. Please make sure you have all the pages.

1. Your Consent

You are invited to take part in this research project.

This Participant Information contains detailed information about the research project. Its purpose is to explain to you as openly and clearly as possible all the procedures involved in this project before you decide whether or not to take part in it.

Please read this Participant Information carefully. Feel free to ask questions about any information in the document. You may also wish to discuss the project with a relative or friend or your local health worker. Feel free to do this.

Once you understand what the project is about and if you agree to take part in it, you will be asked to sign the Consent Form. By signing the Consent Form, you indicate that you understand the information and that you give your consent to participate in the research project.

You will be given a copy of the Participant Information and Consent Form to keep as a record.

2. Purpose and Background

The purpose of this project is to investigate the number of six minute walk tests that need to be performed to provide consistent results. The project will also investigate any changes when the six minute walk test is repeated over minutes, days, one week, or many weeks.

A total of 30 people will participate in this project.

Previous research has shown that the distance walked improves with the number of tests performed. It is not known how many walk tests patients with cardiac disease need to complete in order to get consistent results. It is also not known how time can affect the consistency of the results.

You are invited to participate in this research project because you have been referred to cardiac rehabilitation, and as part of your physiotherapy assessment you would be completing an exercise test.

The results of this research may be used to help researcher Lisa to obtain a degree.

3. Procedures

Participation in this project will involve completing three six minute walk tests during your initial assessment. The walk test involves walking up and down a 20 meter corridor for six minutes.

In this project there will be three groups of participants.

- The first group will complete three six minute walk tests during one afternoon. A 20 minute rest break will occur between each test.
- The second group will complete six minute walk tests over two assessments, generally one or two days apart. The second assessment will coincide with the day that you return to begin cardiac rehabilitation.
- The third group will complete six minute walk tests over two assessment, one week apart. This will coincide with the day that you return for cardiac rehabilitation.

It is estimated that your initial physiotherapy assessment will take 15 to 30 minutes longer than the usual assessment. People who are not participating in the research would normally perform two walk tests.

Upon completion of the cardiac rehabilitation program we will complete a discharge assessment. During this assessment, the usual two six minute walk tests will be completed.

4. Possible Benefits

The exercise test is used to guide the exercise program we develop for you when you start the program. The walk test also provides us with important information to develop a home walking program. The more accurate the results of the walk test, the more appropriate the exercise program will be for you.

We cannot guarantee or promise that you will receive any benefits from this project.

5. Possible Risks

The test may provoke symptoms such as shortness of breath or tiredness in the legs. The test is self-paced, you are able to vary your pace, and you are able to sit down at any stage. You may find that you feel puffed while you are walking, this will generally settle within two to five minutes after completing the test. If this is your first attempt at exercise, you may find that your muscles will be sore from walking.

As the test is self-paced, and you may stop to rest at any time, the risk of a serious side effect such as angina is minimal. In the event of a side effect all the usual safety procedures will be followed.

If you find the test distressing in any way, you may suspend or end your participation in the project.

There may be additional unforeseen or unknown risks.

6. Alternatives to Participation

You are not under any obligation to participate. If you chose not to participate you will still receive the usual physiotherapy assessment, which includes an exercise test. You will receive the same care within the program. If you do not participate, the results of your exercise test will not be used for the research.

7. Privacy, Confidentiality and Disclosure of Information

Any information obtained in connection with this project and that can identify you will remain confidential. Any information that can identify you will be stored in your file within Bendigo Health. It will only be disclosed with your permission, except as required by law.

If you give us your permission by signing the Consent Form, we plan to photocopy this research sheet, remove any identifying features such as your name and address and use the information for further analysis. Information that will be retained includes: age, gender, cardiac diagnosis and intervention, any relevant medical history, and medications, and any walking problems. These results along with the results of your walk test will be entered into a statistics computer program and further analysed.

This de-identified information will be stored in a locked filing cabinet in the offices of La Trobe University Physiotherapy School Bendigo, specifically E5 room 505. Only myself and Dr Helen McBurney will have access to this information. The de-identified data will be archived in locked compactus in HS31.05 on Bundoora Campus of LaTrobe. The information will be kept for seven years, and then shredded.

In any publication, information will be provided in such a way that you cannot be identified. Only group data will be reported, no individuals will be able to be identified.

8. New Information Arising During the Project

During the research project, new information about the risks and benefits of the project may become known to the researchers. If this occurs, you will be told about this new information. This new information may mean that you can no longer participate in this research. If this occurs, the person supervising the research will stop your participation. In all cases, you will be offered all available care to suit your needs and medical condition.

9. Results of Project

At completion of the exercise test your individual results will be explained to you as per our usual practice. A summary of the final group results will be available at the completion of the project. If you wish to receive this information, please indicate on the consent form.

10. Further Information or Any Problems

If you require further information or if you have any problems concerning this project (for example, any side effects), you can contact Lisa Hanson, Dr Helen McBurney, or Ms Deborah Ludeman. The researchers responsible for this project are

Lisa Hanson	Phone (03) 5454 8783 Mobile 0438862423
Dr Helen McBurney	Phone (03) 5454 7021
Deborah Ludeman	Phone (03) 5454 8783

11. Other Issues

If you have any complaints about any aspect of the project, the way it is being conducted or any questions about your rights as a research participant, then you may contact either:

Name: Angela Crombie

Position: Secretary, Human Research Ethics Committee, Bendigo Health

Telephone: (03) 5454 6407

Or

Name: Barbara Doherty

Position: Secretary, Faculty of Human Ethics Committee, Faculty of Health Sciences

Telephone: (03) 9479 1794

You will need to report the name of one of the researchers given in section 10 above.

12. Participation is Voluntary

Participation in any research project is voluntary. If you do not wish to take part you are not obliged to. If you decide to take part and later change your mind, you are free to withdraw from the project at any stage.

Your decision whether to take part or not to take part, or to take part and then withdraw, will not affect your routine treatment, your relationship with those treating you or your relationship with Bendigo Health.

Before you make your decision, a member of the research team will be available to answer any questions you have about the research project. You can ask for any information you want. Sign the Consent Form only after you have had a chance to ask your questions and have received satisfactory answers.

If you decide to withdraw from this project, please notify a member of the research team before you withdraw. This notice will allow that person or the research supervisor to inform you if there are any health risks or special requirements linked to withdrawing.

13. Ethical Guidelines

This project will be carried out according to the *National Statement on Ethical Conduct in Research Involving Humans* (June 1999) produced by the National Health and Medical Research Council of Australia. This statement has been developed to protect the interests of people who agree to participate in human research studies.

The ethical aspects of this research project have been approved by the Human Research Ethics Committee of Bendigo Health and LaTrobe University, Faculty of Health Science.

14. Reimbursement for your costs

You will not be paid for your participation in this project.

Consent Form

Version: 1 **Date:** 1st October 2006
Site: Cardiac Rehabilitation; Physiotherapy Department, The Bendigo Hospital

Full Project Title: The Reproducibility of the Six Minute Walk Test in Patients Referred to Cardiac Rehabilitation**.**

I have read, or have had read to me, and I understand the Participant Information version ***1*** dated ***1st October 2006***.

I freely agree to participate in this project according to the conditions in the Participant Information.

I will be given a copy of the Participant Information and Consent Form to keep

The researcher has agreed not to reveal my identity and personal details if information about this project is published or presented in any public form.

Participant's Name (printed) ……

Signature Date

Name of Witness to Participant's Signature (printed) ……

Signature Date

Declaration by researcher*: I have given a verbal explanation of the research project, its procedures and risks and I believe that the participant has understood that explanation.

Researcher's Name (printed) ……

Signature Date

* A senior member of the research team must provide the explanation and provision of information concerning the research project.

☐ I wish to receive a summary of the final results of this project. I understand that the researchers will mail the results to me, and will access my Bendigo Health Medical Record for my postal address.

Signature Date

Note: All parties signing the Consent Form must date their own signature.

Revocation of Consent Form

Version: 1 **Date:** 1st October 2006
Site: Cardiac Rehabilitation; Physiotherapy Department, The Bendigo Hospital

Full Project Title: The Reproducibility of the Six Minute Walk Test in Patients Referred to Cardiac Rehabilitation**.**

I hereby wish to WITHDRAW my consent to participate in the research proposal described above and understand that such withdrawal WILL NOT jeopardise any treatment or my relationship with Bendigo Health.

Participant's Name (printed) …….

Signature Date

Participant Information and Consent Form

Date: 8th March 2007
Site: Cardiac Rehabilitation; Physiotherapy Department, The Bendigo Hospital

Full Project Title: A Comparison of the Incremental Shuttle Walk Test and the Six Minute Walk Test in Patients Referred to Cardiac Rehabilitation

Principal Researcher: Lisa Hanson (Senior Physiotherapist, PhD Student)

Principal Researcher: Dr Helen McBurney (Associate Professor, Supervisor)

Principal Researcher: Dr Nicholas Taylor (Professor, Supervisor)

Associate Researcher: Deborah Ludeman (Cardiac Rehabilitation Co-ordinator)

This Participant Information and Consent Form is 7 pages long. Please make sure you have all the pages.

1. Your Consent

You are invited to take part in this research project.

This Participant Information contains detailed information about the research project. Its purpose is to explain to you as openly and clearly as possible all the procedures involved in this project before you decide whether or not to take part in it.

Please read this Participant Information carefully. Feel free to ask questions about any information in the document. You may also wish to discuss the project with a relative or friend or your local health worker. Feel free to do this.

Once you understand what the project is about and if you agree to take part in it, you will be asked to sign the Consent Form. By signing the Consent Form, you indicate that you understand the information and that you give your consent to participate in the research project.

You will be given a copy of the Participant Information and Consent Form to keep as a record.

2. Purpose and Background

The purpose of this project is to investigate the number of incremental shuttle walk tests (ISWT) that are needed to be performed to provide the best test result, and later to compare the ISWT with the six minute walk test (6MWT).

A total of 30 people will participate in this project.

It is not known how many ISWTs need to completed by patients with cardiac disease in order to get a best test result. Two previous research papers have reported the benefit of a practice test, however, one paper showed consistent results without a practice test. It is also not known if the ISWT provides patients with cardiac disease with a higher level of function with the opportunity to demonstrate a better performance in the ISWT than the 6MWT.

You are invited to participate in this research project because you have been referred to cardiac rehabilitation, and as part of your physiotherapy assessment you would be completing an exercise test.

This project is a part of PhD research by Lisa, and the results may be used to help Lisa to obtain a PhD degree.

3. Procedures

Participation in this project will involve completion of walk tests during your initial assessment. In this project there will be two groups of participants.

- The first group will complete three ISWTs during one afternoon. A 20 minute rest break will occur between each test. This involves more tests than the usual care (i.e., three tests instead of two). This walk test involves walking up and down a 20 metre corridor for a maximum of 12 minutes. Walking speed is set by a pre-recorded CD. Upon discharge from the program an addition two ISWTs will be completed. This is consistent with usual care.
- The second group will complete two different walk tests. Participants in this group will complete both the ISWT and the 6MWT during the initial physiotherapy assessment. A 20 minute rest break will occur between each test. The ISWT involves walking up and down a 20 metre corridor for a maximum of 12 minutes. Walking speed is set by a pre-recorded CD. The 6MWT involves walking up and down a 20 metre corridor for six minutes, with the aim to cover as much ground in that time.

It is estimated that your initial physiotherapy assessment will take 15 to 45 minutes longer than usual.

4. Possible Benefits

The exercise test is used to guide the exercise program we develop for you when you start the program. The walk test also provides us with important information to develop a home walking program. The more accurate the results of the walk test, the more appropriate the exercise program will be for you.

We cannot guarantee or promise that you will receive any benefits from this project.

5. Possible Risks

The test may provoke symptoms such as shortness of breath or tiredness in the legs. You may find that you feel puffed while you are walking, this will generally settle within two to five minutes after completing the test. If this is your first attempt at exercise, you may find that your muscles will be sore from walking. As the tests are self limiting, you are able to stop at any stage.

In the event of a side effect all the usual safety procedures will be followed.

If you find the test distressing in any way, you may suspend or end your participation in the project.

6. Alternatives to Participation

You are not under any obligation to participate. If you chose not to participate you will still receive the usual physiotherapy assessment, which includes an exercise test. You will receive the same care within the program. If you do not participate, the results of your exercise test will not be used for the research.

7. Privacy, Confidentiality and Disclosure of Information

Any information obtained in connection with this project and that can identify you will remain confidential. Any information that can identify you will be stored in your file within Bendigo Health. It will only be disclosed with your permission, except as required by law.

If you give us your permission by signing the Consent Form, we plan to photocopy this research sheet, remove any identifying features such as your name and address and use the information for further analysis. Information that will be retained includes: age, gender, cardiac diagnosis and intervention, any relevant medical history, and medications, and any walking problems. These results along with the results of your walk test will be entered into a statistics computer program and further analysed.

This de-identified information will be stored in a locked filing cabinet in the offices of La Trobe University Physiotherapy School Bendigo, specifically E5 room 505. Only myself and Dr Helen McBurney will have access to this information. The de-identified data will be archived and shredded after 7 years. In any publication, information will be provided in such a way that you cannot be identified. Only group data will be reported, no individuals will be able to be identified.

8. New Information Arising During the Project

During the research project, new information about the risks and benefits of the project may become known to the researchers. If this occurs, you will be told about this new information. This new information may mean that you can no longer participate in this research. If this occurs, the person supervising the research will stop your participation. In all cases, you will be offered all available care to suit your needs and medical condition.

9. Results of Project

At completion of the exercise test your individual results will be explained to you as per our usual practice. A summary of the final group results will be available at the completion of the project. If you wish to receive this information, please indicate on the consent form.

10. Further Information or Any Problems

If you require further information or if you have any problems concerning this project, you can contact Lisa Hanson, Dr Helen McBurney, or Ms Deborah Ludeman. The researchers responsible for this project are

Lisa Hanson Phone (03) 5454 8783

Dr Helen McBurney Phone (03) 5454 7021 Please note that Dr Helen McBurney is a Doctor of Philosophy, not a Medical Doctor.

Dr Nicholas Taylor Phone (03) 9479 5860 Please not that Dr Nicholas Taylor is a Doctor of Philosophy, not a Medical Doctor.

Deborah Ludeman Phone (03) 5454 8783

If you develop any medical problems outside of the assessment, you should attend your GP or emergency department. If you develop chest pain, please follow your usual chest pain management strategy, and call an ambulance if required.

11. Other Issues

If you have any complaints about any aspect of the project, the way it is being conducted or any questions about your rights as a research participant, then you may contact either:

Name: Angela Crombie

Position: Secretary, Human Research Ethics Committee, Bendigo Health

Telephone: (03) 5454 6407

Or

Name: Barbara Doherty

Position: Secretary, Faculty Human Ethics Committee, Faculty of Health Sciences, La Trobe University

Telephone: (03) 9479 1794 E-mail: B.Doherty@latrobe.edu.au

You will need to report the name of one of the researchers given in section 10 above.

12. Participation is Voluntary

Participation in any research project is voluntary. If you do not wish to take part you are not obliged to. If you decide to take part and later change your mind, you are free to withdraw from the project at any stage.

Your decision whether to take part or not to take part, or to take part and then withdraw, will not affect your routine treatment, your relationship with those treating you or your relationship with Bendigo Health.

Before you make your decision, a member of the research team will be available to answer any questions you have about the research project. You can ask for

any information you want. Sign the Consent Form only after you have had a chance to ask your questions and have received satisfactory answers.

If you decide to withdraw from this project, please notify a member of the research team before you withdraw. This notice will allow that person or the research supervisor to inform you if there are any health risks or special requirements linked to withdrawing.

13. Ethical Guidelines

This project will be carried out according to the *National Statement on Ethical Conduct in Research Involving Humans* (June 1999) produced by the National Health and Medical Research Council of Australia. This statement has been developed to protect the interests of people who agree to participate in human research studies.

The ethical aspects of this research project have been approved by the Human Research Ethics Committee of Bendigo Health and LaTrobe University, Faculty of Health Science..

14. Reimbursement for your costs

You will not be paid for your participation in this project.

Consent Form

Version: 1 **Date:** 8th March 2007
Site: Cardiac Rehabilitation; Physiotherapy Department, The Bendigo Hospital

Full Project Title: A Comparison of the Incremental Shuttle Walk Test and the Six Minute Walk Test in Patients Referred to Cardiac Rehabilitation**.**

I have read, or have had read to me, and I understand the Participant Information version ***1*** dated ***1st October 2006***.

I freely agree to participate in this project according to the conditions in the Participant Information.

I will be given a copy of the Participant Information and Consent Form to keep

The researcher has agreed not to reveal my identity and personal details if information about this project is published or presented in any public form.

Participant's Name (printed) ……

Signature Date

Name of Witness to Participant's Signature (printed) ……

Signature Date

Declaration by researcher*: I have given a verbal explanation of the research project, its procedures and risks and I believe that the participant has understood that explanation.

Researcher's Name (printed) ……

Signature Date

* A senior member of the research team must provide the explanation and provision of information concerning the research project.

☐ I wish to receive a summary of the final results of this project. I understand that the researchers will mail the results to me, and will access my Bendigo Health Medical Record for my postal address.

Signature Date

Note: All parties signing the Consent Form must date their own signature.

Revocation of Consent Form

Version: 1 **Date:** 8th March 2007
Site: Cardiac Rehabilitation; Physiotherapy Department, The Bendigo Hospital

Full Project Title: A Comparison of the Incremental Shuttle Walk Test and the Six Minute Walk Test in Patients Referred to Cardiac Rehabilitation**.**

I hereby wish to WITHDRAW my consent to participate in the research proposal described above and understand that such withdrawal WILL NOT jeopardise any treatment or my relationship with Bendigo Health.

Participant's Name (printed) …….

Signature Date

Participant Information and Consent Form

Date: June 28, 2010
Site: Cardiac Rehabilitation; Physiotherapy Department, St John of God Healthcare, Bendigo

Full Project Title: The Measurement Properties of the Modified 10m Incremental Shuttle Walk Test when Used in Cardiac Rehabilitation. Part A.

Principal Researcher: Lisa Hanson (Senior Clinician, PhD Student)

Principal Researcher: Dr Nicholas Taylor (Professor, Supervisor)

Principal Researcher: Dr Helen McBurney (Associate Professor, Supervisor)

This Participant Information and Consent Form is 6 pages long. Please make sure you have all the pages.

1. Your Consent

You are invited to take part in this research project.

This Participant Information contains detailed information about the research project. Its purpose is to explain to you as openly and clearly as possible all the procedures involved in this project before you decide whether or not to take part in it.

Please read this Participant Information carefully. Feel free to ask questions about any information in the document. You may also wish to discuss the project with a relative or friend or your local health worker. Feel free to do this.

Once you understand what the project is about and if you agree to take part in it, you will be asked to sign the Consent Form. By signing the Consent Form, you indicate that you understand the information and that you give your consent to participate in the research project.

You will be given a copy of the Participant Information and Consent Form to keep as a record.

2. Purpose and Background

The purpose of this project is to determine if the results of the cardiopulmonary exercise test (stress test) correlate with results of the modified 10m incremental shuttle walk test (ISWT).

A total of 15 people will participate in this project.

It is not known if the results of the modified 10m ISWT correlate with the results of the stress test ordered by your cardiologist. Previous research has indicated that there is a strong relationship, however, this research has looked at a specific group of people with cardiac disease.

You are invited to participate in this research project because you have been referred to your cardiologist for a stress test.

This project is a part of PhD research by Lisa, and the results may be used to help Lisa to obtain a PhD degree.

3. Procedures

Participation in this project will involve completion of two additional walk tests in the same week that you complete your stress test. A minimum of a 20 minute rest break will occur between each walk test. This walk test involves walking up and down a flat indoor corridor for a maximum of 12 minutes. Walking speed is set by a pre-recorded CD.

It is estimated that this additional assessment will take no longer than 60 minutes.

4. Possible Benefits

We cannot guarantee or promise that you will receive any benefits from this project.

5. Possible Risks

The test may provoke symptoms such as shortness of breath or tiredness in the legs. You may find that you feel puffed while you are walking, this will generally settle within two to five minutes after completing the test. If this is your first attempt at exercise, you may find that your muscles will be sore from walking. As the test is self limiting, you are able to stop at any stage.

In the event of a side effect all the usual safety procedures will be followed.

If you find the test distressing in any way, you may suspend or end your participation in the project.

6. Alternatives to Participation

You are not under any obligation to participate. If you chose not to participate you will still receive the usual level care by your cardiologist.

7. Privacy, Confidentiality and Disclosure of Information

Any information obtained in connection with this project and that can identify you will remain confidential. It will only be disclosed with your permission, except as required by law.

If you give us your permission by signing the Consent Form, we will record your results of the modified 10m ISWT and match these with your results of your stress test. Once results have been matched we will remove any identifying features such as your name and address and use the information for analysis. Personal information that will be retained includes: age, gender, cardiac diagnosis and intervention, any relevant medical history, and medications, and any walking problems. These results along with the results of your walk test and stress test will be entered into a statistics computer program and analysed.

This de-identified information will be stored in a locked filing cabinet in the offices of La Trobe University Physiotherapy School Bendigo, specifically Level 1 of the West Wing in room W2-01D. Only Lisa Hanson, Dr Nicholas Taylor and Dr Helen McBurney will have access to this information. The de-identified data will be archived and shredded after 7 years. In any publication, information will be provided in such a way that you cannot be identified. Group data will be reported, no individuals will be able to be identified.

8. New Information Arising During the Project

During the research project, new information about the risks and benefits of the project may become known to the researchers. If this occurs, you will be told about this new information. This new information may mean that you can no longer participate in this research. If this occurs, the person supervising the research will stop your participation. In all cases, you will be offered all available care to suit your needs and medical condition.

9. Results of Project

At completion of the exercise test your individual results will be explained to you as per our usual practice. A summary of the final group results will be available at the completion of the project. If you wish to receive this information, please indicate on the consent form.

10. Further Information or Any Problems

If you require further information or if you have any problems concerning this project, you can contact Lisa Hanson, Dr Nicholas Taylor or Dr Helen McBurney. The researchers responsible for this project are

Lisa Hanson	Phone (03) 5454 7020
Dr Nicholas Taylor	Phone (03) 9479 5860 Please not that Dr Nicholas Taylor is a Doctor of Philosophy, not a Medical Doctor.
Dr Helen McBurney	Phone (03) 5173 8196 or (03) 5173 8216 Please note that Dr Helen McBurney is a Doctor of Philosophy, not a Medical Doctor.

If you develop any medical problems outside of the assessment, you should attend your GP or emergency department. If you develop chest pain, please follow your usual chest pain management strategy, and call an ambulance if required.

11. Other Issues

St John of God Health Care Ethics Committee and the Faculty Human Ethics Committee, Faculty of Health Sciences, La Trobe University have given ethical approval for the conduct of this study. If you have any concerns of complaints regarding any aspect of the project, the way it is being conducted or any questions about your rights as a research participant, then you may contact either:

Name:	Ms Gorette De Jesus
Position:	Executive Officer, St John of God Health Care Ethics Committee
Telephone:	(08) 9382 6940
Email:	ethics@sjog.org.au
Or	
Name:	Neil McDonald
Position:	Secretary, Faculty Human Ethics Committee, Faculty of Health Sciences, La Trobe University
Telephone:	(03) 9479 2357
E-mail:	n.mcdonald@latrobe.edu.au

You will need to report the name of one of the researchers given in section 10 above. This will be managed confidentially. Your concerns will be drawn to the attention of the Committee that is monitoring the study.

12. Participation is Voluntary

Participation in any research project is voluntary. If you do not wish to take part you are not obliged to. If you decide to take part and later change your mind, you are free to withdraw from the project at any stage.

Your decision whether to take part or not to take part, or to take part and then withdraw, will not affect your routine treatment, your relationship with those treating you or your relationship with St John of God Health Care.

Before you make your decision, a member of the research team will be available to answer any questions you have about the research project. You can ask for any information you want. Sign the Consent Form only after you have had a chance to ask your questions and have received satisfactory answers.

If you decide to withdraw from this project, please notify a member of the research team before you withdraw. This notice will allow that person or the research supervisor to inform you if there are any health risks or special requirements linked to withdrawing.

13. Ethical Guidelines

This project will be carried out according to the *National Statement on Ethical Conduct in Research Involving Humans* (June 1999) produced by the National Health and Medical Research Council of Australia. This statement has been developed to protect the interests of people who agree to participate in human research studies.

The ethical aspects of this research project have been approved by the Human Research Ethics Committee of St John of God Health Care and LaTrobe University, Faculty of Health Science.

14. Reimbursement for your costs

You will not be paid for your participation in this project.

Consent Form

Version: 1 Date: June 28, 2010
Site: Cardiac Rehabilitation; Physiotherapy Department, St John of God Health Care

Full Project Title: The Measurement Properties of the Modified 10m Incremental Shuttle Walk Test when Used in Cardiac Rehabilitation. Part A.

I have read, or have had read to me, and I understand the Participant Information version ***1*** dated ***June 28, 2010***.

I freely agree to participate in this project according to the conditions in the Participant Information.

I will be given a copy of the Participant Information and Consent Form to keep

The researcher has agreed not to reveal my identity and personal details if information about this project is published or presented in any public form.

Participant's Name (printed) ……

Signature Date

Name of Witness to Participant's Signature (printed) ……

Signature Date

Declaration by researcher*: I have given a verbal explanation of the research project, its procedures and risks and I believe that the participant has understood that explanation.

Researcher's Name (printed) ……

Signature Date

* A senior member of the research team must provide the explanation and provision of information concerning the research project.

☐ I wish to receive a summary of the final results of this project. I understand that the researchers will mail the results to me, and will access my St John of God Medical Record for my postal address.

Signature Date

Note: All parties signing the Consent Form must date their own signature.

Revocation of Consent Form

Version: 1 **Date:** June 28, 2010
Site: Cardiac Rehabilitation; Physiotherapy Department, St John of God Health Care

Full Project Title: The Measurement Properties of the Modified 10m Incremental Shuttle Walk Test when Used in Cardiac Rehabilitation. Part A.

I hereby wish to WITHDRAW my consent to participate in the research proposal described above and understand that such withdrawal WILL NOT jeopardise any treatment or my relationship with St John of God Health Care.

Participant's Name (printed) ……

Signature Date

Participant Information and Consent Form

Date: June 28, 2010
Site: Cardiac Rehabilitation; Physiotherapy Department, St John of God Health Care, Bendigo

Full Project Title: The Measurement Properties of the Modified 10m Incremental Shuttle Walk Test when Used in Cardiac Rehabilitation. Part B.

Principal Researcher: Lisa Hanson (Senior Clinician, PhD Student)

Principal Researcher: Dr Nicholas Taylor (Professor, Supervisor)

Principal Researcher: Dr Helen McBurney (Associate Professor, Supervisor)

This Participant Information and Consent Form is 6 pages long. Please make sure you have all the pages.

1. Your Consent

You are invited to take part in this research project.

This Participant Information contains detailed information about the research project. Its purpose is to explain to you as openly and clearly as possible all the procedures involved in this project before you decide whether or not to take part in it.

Please read this Participant Information carefully. Feel free to ask questions about any information in the document. You may also wish to discuss the project with a relative or friend or your local health worker. Feel free to do this.

Once you understand what the project is about and if you agree to take part in it, you will be asked to sign the Consent Form. By signing the Consent Form, you indicate that you understand the information and that you give your consent to participate in the research project.

You will be given a copy of the Participant Information and Consent Form to keep as a record.

2. Purpose and Background

The purpose of this project is to investigate measurement properties of the modified 10m incremental shuttle walk test (ISWT) when used in cardiac rehabilitation, i.e., to determine if the test provides an accurate measure of the physical fitness of people who attend cardiac rehabilitation.

A total of 50 people will participate in this project.

The validity, responsiveness and thus interpretability of the 10m ISWT when used to assess physical fitness in a general cardiac rehabilitation program remains largely unknown. Previous research involved a group of people with specific cardiac disease (male patients following open heart surgery), and in this instance suggest that the test was valid and responsive to change.

You are invited to participate in this research project because you have been referred to cardiac rehabilitation.

This project is a part of PhD research by Lisa, and the results may be used to help Lisa to obtain a PhD degree.

3. Procedures

Participation in this project will involve completion of walk tests during your initial assessment. Two modified 10m ISWTs will be completed on admission to cardiac rehabilitation. A minimum of a 15 minute rest break will occur between each test. In addition you will be asked to complete questionnaires that measure quality of life, likelihood of depression and self efficacy. Throughout the cardiac rehabilitation program you will be asked to record your activity and pedometer readings in an exercise diary. Upon discharge from cardiac rehabilitation you will be asked to repeat the walk tests and the questionnaires. These assessments are more involved the usual care (i.e., usual care does not include an individual physiotherapy assessment on admission to and discharge from cardiac rehabilitation).

The walk test involves walking up and down a flat indoor corridor for a maximum of 12 minutes. Walking speed is set by a pre-recorded CD.

It is estimated that your initial physiotherapy assessment will take 60 minutes longer than usual.

4. Possible Benefits

The walk test is used to guide the exercise program we develop for you when you start the program. The walk test also provides us with important information to develop a home walking program. The more accurate the results of the walk test, the more appropriate the exercise program will be for you.

We cannot guarantee or promise that you will receive any benefits from this project.

5. Possible Risks

The test may provoke symptoms such as shortness of breath or tiredness in the legs. You may find that you feel puffed while you are walking, this will generally settle within two to five minutes after completing the test. If this is your first attempt at exercise, you may find that your muscles will be sore from walking. As the test is self limiting, you are able to stop at any stage.

In the event of a side effect all the usual safety procedures will be followed.

If you find the test distressing in any way, you may suspend or end your participation in the project.

6. Alternatives to Participation

You are not under any obligation to participate. If you chose not to participate you will receive the usual level of care within the program

7. Privacy, Confidentiality and Disclosure of Information

Any information obtained in connection with this project and that can identify you will remain confidential. Any information that can identify you will be securely stored at St John of God Health Care in Bendigo. It will only be disclosed with your permission, except as required by law.

If you give us your permission by signing the Consent Form, we plan to record information on the recording sheet. Following completion of cardiac rehabilitation we will remove any identifying features such as your name and address and use the information for further analysis. Information that will be retained includes: age, gender, cardiac diagnosis and intervention, any relevant medical history, and medications, and any walking problems. These results along with the results of your walk test, exercise diary and questionnaires will be entered into a statistics computer program and further analysed.

This de-identified information will be stored in a locked filing cabinet in the offices of La Trobe University Physiotherapy School Bendigo, specifically Level 1 of the West Wing in room W2-01D. Only Lisa Hanson, Dr Nicholas Taylor and Dr Helen McBurney will have access to this information. The de-identified data will be archived and shredded after 7 years. In any publication, information will be provided in such a way that you cannot be identified. Group data will be reported, no individuals will be able to be identified.

8. New Information Arising During the Project

During the research project, new information about the risks and benefits of the project may become known to the researchers. If this occurs, you will be told about this new information. This new information may mean that you can no longer participate in this research. If this occurs, the person supervising the research will stop your participation. In all cases, you will be offered all available care to suit your needs and medical condition.

9. Results of Project

At completion of the exercise test your individual results will be explained to you as per our usual practice. A summary of the final group results will be available at the completion of the project. If you wish to receive this information, please indicate on the consent form.

10. Further Information or Any Problems

If you require further information or if you have any problems concerning this project, you can contact Lisa Hanson, Dr Nicholas Taylor or Dr Helen McBurney. The researchers responsible for this project are

Lisa Hanson	Phone (03) 5454 7020
Dr Nicholas Taylor	Phone (03) 9479 5860 Please not that Dr Nicholas Taylor is a Doctor of Philosophy, not a Medical Doctor.
Dr Helen McBurney	Phone (03) 5173 8196 or (03) 5173 8216 Please note that Dr Helen McBurney is a Doctor of Philosophy, not a Medical Doctor.

If you develop any medical problems outside of the assessment, you should attend your GP or emergency department. If you develop chest pain, please follow your usual chest pain management strategy, and call an ambulance if required.

11. Other Issues

St John of God Health Care Ethics Committee and the Faculty Human Ethics Committee, Faculty of Health Sciences, La Trobe University have given ethical approval for the conduct of this study. If you have any concerns of complaints regarding any aspect of the project, the way it is being conducted or any questions about your rights as a research participant, then you may contact either:

Name:	Ms Gorette De Jesus
Position:	Executive Officer, St John of God Health Care Ethics Committee
Telephone:	(08) 9382 6940
Email:	ethics@sjog.org.au
Or	
Name:	Neil McDonald
Position:	Secretary, Faculty Human Ethics Committee, Faculty of Health Sciences, La Trobe University
Telephone:	(03) 9479 2357
E-mail:	n.mcdonald@latrobe.edu.au

You will need to report the name of one of the researchers given in section 10 above. This will be managed confidentially. Your concerns will be drawn to the attention of the Committee that is monitoring the study.

12. Participation is Voluntary

Participation in any research project is voluntary. If you do not wish to take part you are not obliged to. If you decide to take part and later change your mind, you are free to withdraw from the project at any stage.

Your decision whether to take part or not to take part, or to take part and then withdraw, will not affect your routine treatment, your relationship with those treating you or your relationship with St John of God Health.

Before you make your decision, a member of the research team will be available to answer any questions you have about the research project. You can ask for any information you want. Sign the Consent Form only after you have had a chance to ask your questions and have received satisfactory answers.

If you decide to withdraw from this project, please notify a member of the research team before you withdraw. This notice will allow that person or the research supervisor to inform you if there are any health risks or special requirements linked to withdrawing.

13. Ethical Guidelines

This project will be carried out according to the *National Statement on Ethical Conduct in Research Involving Humans* (June 1999) produced by the National Health and Medical Research Council of Australia. This statement has been developed to protect the interests of people who agree to participate in human research studies.

The ethical aspects of this research project have been approved by the Human Research Ethics Committee of St John of God Health Careand LaTrobe University, Faculty of Health Science.

14. Reimbursement for your costs

You will not be paid for your participation in this project.

Consent Form

Version: 1 Date: June 28, 2010
Site: Cardiac Rehabilitation; Physiotherapy Department, St John of God Health Care Bendigo

Full Project Title: The Measurement Properties of the Modified 10m Incremental Shuttle Walk Test when used in Cardiac Rehabilitation. Part B.

I have read, or have had read to me, and I understand the Participant Information version ***1*** dated ***June 28, 2010***.

I freely agree to participate in this project according to the conditions in the Participant Information.

I will be given a copy of the Participant Information and Consent Form to keep

The researcher has agreed not to reveal my identity and personal details if information about this project is published or presented in any public form.

Participant's Name (printed) ……

Signature Date

Name of Witness to Participant's Signature (printed) ……

Signature Date

Declaration by researcher*: I have given a verbal explanation of the research project, its procedures and risks and I believe that the participant has understood that explanation.

Researcher's Name (printed) ……

Signature Date

* A senior member of the research team must provide the explanation and provision of information concerning the research project.

☐ I wish to receive a summary of the final results of this project. I understand that the researchers will mail the results to me, and will access my St John of God record for my postal address.

Signature Date

Note: All parties signing the Consent Form must date their own signature.

Revocation of Consent Form

Version: 1 Date: June 28, 2010
Site: Cardiac Rehabilitation; Physiotherapy Department, St John of God Healthcare Bendigo

Full Project Title: The Measurement Properties of the Modified 10m Incremental Shuttle Walk Test when used in Cardiac Rehabilitation. Part B.

I hereby wish to WITHDRAW my consent to participate in the research proposal described above and understand that such withdrawal WILL NOT jeopardise any treatment or my relationship with St John of God Healthcare.

Participant's Name (printed) ……

Signature Date

Appendix 4: Data Recording Sheets

Data recording sheets were used for data collection for the empirical studies.

UR #……
Surname ……
Name……
DOB……
(UR Sticker if available)

Data Collection Sheet - Initial Assessment

Project: The Reproducibility of the Six Minute Walk Test in Patients Referred to Cardiac Rehabilitation.

Date/Time Assessor:

Cardiac Diagnosis/Intervention: Cardiac Signs and Symptoms: Cardiac Medications		
Relevant Medical History: Other Relevant Medications		
Mobility Status: Gait Aid:		
Age: Gender:	Weight: Height:	
Walk 1: Date/Time: Pre: BP HR RR SpO2 **6MW Distance:** Immediately Post BP HR RR SpO2 RPE 2 Minutes Post BP HR RR SpO2 RPE Further Monitoring BP HR RR SpO2 RPE Symptoms Provoked:	**Walk 2:** Date/Time: Pre: BP HR RR SpO2 **6MW Distance:** Immediately Post BP HR RR SpO2 RPE 2 Minutes Post BP HR RR SpO2 RPE Further Monitoring BP HR RR SpO2 RPE Symptoms Provoked:	**Walk 3:** Date/Time: Pre: BP HR RR SpO2 **6MW Distance:** Immediately Post BP HR RR SpO2 RPE 2 Minutes Post BP HR RR SpO2 RPE Further Monitoring BP HR RR SpO2 RPE Symptoms Provoked:

UR #……
Surname …….
Name……
DOB…….
(UR Sticker if available)

Data Collection Sheet – Discharge Assessment

Project: The Reproducibility of the Six Minute Walk Test in Patients Referred to Cardiac Rehabilitation.

Walk 4	**Walk 5**
Date/Time:	Date/Time:
Pre:	Pre:
BP HR	BP HR
RR SpO2	RR SpO2
6MW Distance:	**6MW Distance:**
Immediately Post	Immediately Post
BP HR	BP HR
RR SpO2	RR SpO2
RPE	RPE
2 Minutes Post	2 Minutes Post
BP HR	BP HR
RR SpO2	RR SpO2
RPE	RPE
Further Monitoring	Further Monitoring
BP HR	BP HR
RR SpO2	RR SpO2
RPE	RPE
Symptoms Provoked:	Symptoms Provoked:

Other Comments (i.e., report any adverse responses)

☐ Discussed individual results with participant

Signature: Date:

UR #……
Surname ……
Name……

Data Collection Sheet - Initial Assessment

Project: A Comparison of the Incremental Shuttle Walk Test and the Six Minute Walk Test in Patients Referred to Cardiac Rehabilitation
Study One: Repeated ISWTs

Date/Time Assessor:

Cardiac Diagnosis/Intervention:

Cardiac Signs and Symptoms:

Cardiac Medications

Relevant Medical History:

Other Relevant Medications

Mobility Status / Gait Aid:
Current and Previous Exercise Tolerance:

Age: Weight:
Gender: Height:

Walk 1:	**Walk 2:**	**Walk 3:**
Date/Time:	**Date/Time:**	**Date/Time:**
Pre:	Pre:	Pre:
BP HR	BP HR	BP HR
RR SpO2	RR SpO2	RR SpO2
ISWT: Time Shuttle Distance Peak HR/RPE	**ISWT:** Time Shuttle Distance Peak HR/RPE	**ISWT** Time Shuttle Distance Peak HR/RPE
Immediately Post	Immediately Post	Immediately Post
BP HR	BP HR	BP HR
RR SpO2	RR SpO2	RR SpO2
RPE	RPE	RPE
2 Minutes Post	2 Minutes Post	2 Minutes Post
BP HR	BP HR	BP HR
RR SpO2	RR SpO2	RR SpO2
RPE	RPE	RPE
Further Monitoring	Further Monitoring	Further Monitoring
BP HR	BP HR	BP HR
RR SpO2	RR SpO2	RR SpO2
RPE	RPE	RPE
Symptoms Provoked:	Symptoms Provoked:	Symptoms Provoked:

UR #……
Surname ……
Name……

Data Collection Sheet - Discharge Assessment -

Project: A Comparison of the Incremental Shuttle Walk Test and the Six Minute Walk Test in Patients Referred to Cardiac Rehabilitation
Study One: Repeated ISWTs

Walk 4: **Date/Time:** Pre: BP HR RR SpO2 **ISWT:** Time Shuttle Distance Peak HR/RPE <u>Immediately Post</u> BP HR RR SpO2 RPE <u>2 Minutes Post</u> BP HR RR SpO2 RPE <u>Further Monitoring</u> BP HR RR SpO2 RPE <u>Symptoms Provoked:</u>	**Walk 5:** **Date/Time:** Pre: BP HR RR SpO2 **ISWT:** Time Shuttle Distance Peak HR/RPE <u>Immediately Post</u> BP HR RR SpO2 RPE <u>2 Minutes Post</u> BP HR RR SpO2 RPE <u>Further Monitoring</u> BP HR RR SpO2 RPE <u>Symptoms Provoked:</u>

Signature: Date

LA TROBE UNIVERSITY

CODE:
Date of Assessment:
Time:
Assessor

Project: The Measurement Properties of the Modified 10m Incremental Shuttle Walk Test when Used in Cardiac Rehabilitation. Part A.

Cardiac History (Include: diagnosis, intervention, signs and symptoms and medications):

Other Relevant Medical History and medications:

Mobility Status / Gait Aid:

Current

Previous Exercise Tolerance:

Age:	Weight:	BMI:
Gender:	Height:	Hip/Waist Ratio

TEST 1: Date

	Pre	Peak	Post test 0 min	Post test 2 min	Post test 5 min
BP					
RR					
SpO2					
HR					
RPE					
Symptoms					

Level 1 ☐☐☐
Level 2 ☐☐☐☐
Level 3 ☐☐☐☐☐
Level 4 ☐☐☐☐☐☐
Level 5 ☐☐☐☐☐☐☐
Level 6 ☐☐☐☐☐☐☐☐
Level 7 ☐☐☐☐☐☐☐☐☐
Level 8 ☐☐☐☐☐☐☐☐☐☐
Level 9 ☐☐☐☐☐☐☐☐☐☐☐
Level 10 ☐☐☐☐☐☐☐☐☐☐☐☐
Level 11 ☐☐☐☐☐☐☐☐☐☐☐☐☐
Level 12 ☐☐☐☐☐☐☐☐☐☐☐☐☐☐

Test 2: Date

	Pre	Peak	Post test 0 min	Post test 2 min	Post test 5 min
BP					
RR					
SpO2					
HR					
RPE					
Symptoms					

Level 1 ☐☐☐
Level 2 ☐☐☐☐
Level 3 ☐☐☐☐☐
Level 4 ☐☐☐☐☐☐
Level 5 ☐☐☐☐☐☐☐
Level 6 ☐☐☐☐☐☐☐☐
Level 7 ☐☐☐☐☐☐☐☐☐
Level 8 ☐☐☐☐☐☐☐☐☐☐
Level 9 ☐☐☐☐☐☐☐☐☐☐☐
Level 10 ☐☐☐☐☐☐☐☐☐☐☐☐
Level 11 ☐☐☐☐☐☐☐☐☐☐☐☐☐
Level 12 ☐☐☐☐☐☐☐☐☐☐☐☐☐☐

LA TROBE UNIVERSITY

CODE:
Date of Assessment:
Time:
Assessor

Cardiopulmonary Exercise Test Result

LA TROBE UNIVERSITY

CODE:
Date of Assessment:
Time:
Assessor

Project: The Measurement Properties of the Modified 10m Incremental Shuttle Walk Test when Used in Cardiac Rehabilitation. Part B.

Cardiac History (Include: diagnosis, intervention, signs and symptoms and medications):

Other Relevant Medical History and medications:

Mobility Status / Gait Aid:

Current

Previous Exercise Tolerance:

Age:	Weight:	BMI:
Gender:	Height:	Hip/Waist Ratio

TEST 1: Date

	Pre	Peak	Post test 0 min	2 min	5 min
BP					
RR					
SpO2					
HR					
RPE					
Symptoms					

Level 1 ☐☐☐
Level 2 ☐☐☐☐
Level 3 ☐☐☐☐☐
Level 4 ☐☐☐☐☐☐
Level 5 ☐☐☐☐☐☐☐
Level 6 ☐☐☐☐☐☐☐☐
Level 7 ☐☐☐☐☐☐☐☐☐
Level 8 ☐☐☐☐☐☐☐☐☐☐
Level 9 ☐☐☐☐☐☐☐☐☐☐☐
Level 10 ☐☐☐☐☐☐☐☐☐☐☐☐
Level 11 ☐☐☐☐☐☐☐☐☐☐☐☐☐
Level 12 ☐☐☐☐☐☐☐☐☐☐☐☐☐☐

Test 2: Date

	Pre	Peak	Post test 0 min	2 min	5 min
BP					
RR					
SpO2					
HR					
RPE					
Symptoms					

Level 1 ☐☐☐
Level 2 ☐☐☐☐
Level 3 ☐☐☐☐☐
Level 4 ☐☐☐☐☐☐
Level 5 ☐☐☐☐☐☐☐
Level 6 ☐☐☐☐☐☐☐☐
Level 7 ☐☐☐☐☐☐☐☐☐
Level 8 ☐☐☐☐☐☐☐☐☐☐
Level 9 ☐☐☐☐☐☐☐☐☐☐☐
Level 10 ☐☐☐☐☐☐☐☐☐☐☐☐
Level 11 ☐☐☐☐☐☐☐☐☐☐☐☐☐
Level 12 ☐☐☐☐☐☐☐☐☐☐☐☐☐☐

LA TROBE UNIVERSITY

CODE:
Date of Assessment:
Time:
Assessor

Discharge Assessment:
Reported Exercise:

Weight | BMI | Hip:Waist ratio

Test 3: Date

	Pre	Peak	Post test 0 min	Post test 2 min	Post test 5 min
BP					
RR					
SpO2					
HR					
RPE					
Symptoms					

Level 1 ☐☐☐
Level 2 ☐☐☐☐
Level 3 ☐☐☐☐☐
Level 4 ☐☐☐☐☐☐
Level 5 ☐☐☐☐☐☐☐
Level 6 ☐☐☐☐☐☐☐☐
Level 7 ☐☐☐☐☐☐☐☐☐
Level 8 ☐☐☐☐☐☐☐☐☐☐
Level 9 ☐☐☐☐☐☐☐☐☐☐☐
Level 10 ☐☐☐☐☐☐☐☐☐☐☐☐
Level 11 ☐☐☐☐☐☐☐☐☐☐☐☐☐
Level 12 ☐☐☐☐☐☐☐☐☐☐☐☐☐☐

Test 4: Data

	Pre	Peak	Post test 0 min	Post test 2 min	Post test 5 min
BP					
RR					
SpO2					
HR					
RPE					
Symptoms					

Level 1 ☐☐☐
Level 2 ☐☐☐☐
Level 3 ☐☐☐☐☐
Level 4 ☐☐☐☐☐☐
Level 5 ☐☐☐☐☐☐☐
Level 6 ☐☐☐☐☐☐☐☐
Level 7 ☐☐☐☐☐☐☐☐☐
Level 8 ☐☐☐☐☐☐☐☐☐☐
Level 9 ☐☐☐☐☐☐☐☐☐☐☐
Level 10 ☐☐☐☐☐☐☐☐☐☐☐☐
Level 11 ☐☐☐☐☐☐☐☐☐☐☐☐☐
Level 12 ☐☐☐☐☐☐☐☐☐☐☐☐☐☐

Admission

SF 36: Overall Score:

PF:
RP:
BP:
GH:
Physical Health

VT:
SF:
RE:
MH:
Mental Health

Scores:
BCD:

Self Efficacy

Discharge

SF 36: Overall Score:

PF:
RP:
BP:
GH:
Physical Health

VT:
SF:
RE:
MH:
Mental Health

Scores:
BCD:

Self Efficacy

Global Rating of Change (upon discharge only):

Appendix 5: Calculation of Group and Individual Confidence Limits for the Symptom-Limited Exercise Test for a Specific 10 m ISWT Distance

The work in this Appendix relates to Chapter 4, Section 4.3.4.2, Predictive criterion validity of the 10 m ISWT. This appendix presents the calculation for the group and individual confidence limits for the symptom-limited exercise at a specific 10 m ISWT distance.

The 95% confidence limits for the symptom-limited exercise test score for the group at a specific 10 m ISWT score (*X*) were calculated (Altman & Gardner, 2000, p. 76; Howell, 2012, p. 275).

$$95\%CI\left(grY_{group}\right) = (a + bX_i) \pm \left(t_{1-\frac{\alpha}{2}}\right)\left(s_{Y \cdot X}\sqrt{\frac{1}{N} + \frac{(X_i - \bar{X})^2}{(N-1)s_X^2}}\right)$$

Where $a + bX_i$ is the regression equation for a specific value for *X* (X_i); $t_{1-\frac{\alpha}{2}}$ is a constant for *t* for a 95% confidence interval, two-tailed at the degrees of freedom and $\left(s_{Y \cdot X}\sqrt{\frac{1}{N} + \frac{(X_i - \bar{X})^2}{(N-1)s_X^2}}\right)$ is the standard error of the regression equation for X_i; $\bar{X}$ is the mean of the *X* scores in the sample; and s_X^2 is the squared standard deviation of *X* (Howell, 2012, p. 275).

The 95% individual confidence limits were calculated for the symptom-limited exercise test specific 10 m ISWT scores using the formula (Howell, 2012, p. 275).

$$95\%CI(Y_{individual}) = (a + bX_i) \pm \left(t_{1-\frac{\alpha}{2}}\right)\left(s_{Y \cdot X}\sqrt{1 + \frac{1}{N} + \frac{(X_i - \bar{X})^2}{(N-1)s_X^2}}\right)$$

Where $s_{Y \cdot X}\sqrt{1 + \frac{1}{N} + \frac{(X_i - \bar{X})^2}{(N-1)s_X^2}}$ is the standard error for a specific value for *X*. Where $s_{Y \cdot X}$ is the standard error of estimiate, N is the number in the sample, $\bar{X}$ is the mean of the *X* scores in the sample, and s_X^2 is the squared standard deviation of *X* (Howell, 2012, p. 275).

Appendix 6: Systematic Review Search Strategy

In this thesis, the systematic review was published in chapter 7. The search strategy was design to be highly sensitive for retrieving studies on measurement properties. The selection of key terms for the measurement properties was based on the guidelines by Terwee et al. (2009) and the recommendations by the COSMIN group (Mokkink et al., 2010c). Table A6.1 outlines the key terms used.

Table A6.1

Key Search Terms used for the Systematic Review

Key Word
Population
Cardiac rehabilitation OR Cardiovascular rehabilitation OR Cardi* OR Myocardi* OR Heart* OR Coronar* OR Cardiac Patient / Cardiac Patients OR Heart (MH) OR Coronary disease (MH) OR Heart Disease (MH) OR Cardiovascular diseases (MH), OR Rehabilitation, Cardiac (MH)
AND
Field Exercise test
Exercise test* (+/-MH) OR Exercise test, cardiopulmonary (MH) OR Exercise test, muscular (MH) OR Walk test OR Shuttle walk OR Six min OR 6MWT OR SMWT OR ISWT OR SWT OR Submaximal exercise test OR Functional test OR Fitness test OR Run test OR Step test OR Ergometry, cycle OR Treadmill test
AND
Psychometric Property
Psychometrics (MH) OR Measurement error (MH) OR Reliability and Validity (MH) OR Reproducibility of results (MH) OR Measurement properties OR Reliab* OR Valid* OR Consistency OR Clinometr* OR Reproducib* OR Responsive* OR Interpretability OR Interpretability OR Instrumentation[sh] OR methods[sh] OR validation studies[pt] OR comparative study[pt] OR psychometrics[MeSH] OR clinimetr*[tw] OR clinometr*[tw] OR outcome assessment (health care)[MeSH] OR outcome assessment[tiab] OR outcome measure[tw] OR observer variation[MeSH] OR

Health Status Indicators[MeSH] OR reproducibility of results[MeSH] OR reproducib*[tiab] OR discriminant analysis[MeSH] OR reliab*[tiab] OR unreliab*[tiab] OR valid*[tiab] OR coefficient[tiab] OR homogeneity[tiab] OR homogeneous[tiab] OR internal consistency[tiab] OR (chronbach*[tiab] AND (alpha[tiab] or alphas[tiab])) OR (item[tiab] AND (correlation[tiab] OR selection[tiab] OR reduction[tiab])) OR agreement[tiab] OR precision[tiab] OR imprecision[tiab] OR precise values[tiab] OR test-retest[tiab] OR (test[tiab] AND retest[tiab]) OR (reliab*[tiab] AND (test[tiab] OR retest[tiab])) OR stability[tiab] OR interrater[tiab] OR interrater[tiab] OR inter-rater[tiab] OR intrarater OR intra-rater[tiab] OR intertester[tiab] OR inter-tester[tiab] OR intratester[tiab] OR intra-tester[tiab] OR interobserver[tiab] OR inter-observer[tiab] OR intraobserver[tiab] OR intra-observer[tiab] OR intertechnician[tiab] OR inter-technician[tiab] OR intratechnician[tiab] OR intra-technician[tiab] OR interexaminer[tiab] OR inter-examiner[tiab] OR intraexaminer[tiab] OR intra-examiner[tiab] OR interassay[tiab] OR inter-assay[tiab] OR intraassay[tiab] OR intra-assay[tiab] OR interparticipant[tiab] OR inter-participant[tiab] OR intraparticipant[tiab] OR intra-participant OR kappa[tiab] OR Kappa's[tiab] OR kappas[tiab] OR repeatab*[tiab] OR ((replicab*[tiab] OR repeated[tiab]) AND (measure[tiab] OR measures[tiab] OR findings[tiab] OR result[tiab] OR test[tiab] OR tests[tiab])) OR generalisa[tiab] OR generaliza[tiab] OR concordance[tiab] OR (intraclass[tiab] AND correlation[tiab]) OR discriminative[tiab] OR known group[tiab] OR factor analysis[tiab] OR factor analyses[tiab] OR dimension[tiab] OR subscale[tiab] OR (multitrait[tiab] AND scaling[tiab] AND (analysis[tiab] OR analyses[tiab])) OR item discriminant[tiab] OR interscale correlation*[tiab] OR error[tiab] OR errors[tiab] OR individual variability[tiab] OR (variability[tiab] AND (analysis[tiab] OR values[tiab]) OR (uncertainty[tiab] AND (measurement[tiab] OR measuring[tiab])) OR "standard error of measurement"[tiab] OR sensitive*[tiab] OR responsive*[tiab] OR ((minimal[tiab] OR minimally[tiab] OR clinical[tiab]) AND (important[tiab] OR significant[tiab] OR detectable[tiab]) AND (change[tiab] OR differences[tiab])) OR (small*[tiab] AND (real[tiab] OR detectable[tiab]) AND (change[tiab] OR difference[tiab])) OR meaningful change[tiab] OR "ceiling effect"[tiab] OR "floor effect"[tiab] OR "Item response model"[tiab] OR IRT[tiab] OR Rasch[tiab] OR "differential item functioning"[tiab] OR DIF[tiab] OR "computer adaptive testing"[tiab] OR 'item bank"[tiab] OR "cross-cultural equivalence"[tiab]) ……..

Appendix 7: Systematic Review Data Extraction Forms

Table A7.1 shows the data extraction form used for the studies included in the systematic review.

Table A7.2 shows the summary sheets for scoring the COSMIN checklist with 4-point scale (Terwee et al., 2012).

Table A7.1

Data extraction for the Systematic Review

Date:	Name:
Reference	
Measurement properties: RR ME CrV CoV Re In	
Study setting and country/language:	
Population (n): Age *M* (SD) Gender M Diagnosis	
Inclusion criteria:	Exclusion criteria
Participant selection method described	
Field Exercise test:	
Operating procedures:	Outcome measures
Results: including field test scores; ceiling or floor; statistics as per COSMIN Y/N	
Description of other measures	
Other comments	
Measurement properties COSMIN quality scoring	
Other	

Note. RR = relative reliability; ME = measurement error; CrV = criterion validity; CoV = construct validity (hypothesis testing); Re = responsiveness; In = Interpretability.

Table A7.2

COSMIN Methodological Quality Scoring using the 4-Point Scale

Citation
Box B (RR)
1
2
3 SS
4
5
6
7
8
9
10
11
12
13
14
Worst Score
Box C (ME)
1
2
3 ss
4
5
6
7
8
9
10
11
Worst Score

Citation
Box F (CoV)
1
2
3 ss
4
5
6
7
8
9
10
Worst Score
Box H (CrV)
1
2
3SS
4
5
6
7
Worst Score
Box I (Re)
1
2
3 SS
4
5
6
7
8
9
10
11
12

	Citation
13	
14	
15	
16	
17	
18	
Worst score	

Note. RR = relative reliability; ME = measurement error; CrV = criterion validity; CoV = construct validity; Re = responsiveness; SS = sample size and refers to the question regarding the adequacy of the sample size. The labels Box B - Box I reflext the boxes in the COSMIN 4-point scale checklist. The grey shaded rows indicate questions that relate to statistical analysis.

Appendix 8: Systematic Review Studies Excluded after Full Text and Studies Included through Manual Search Strategies

There were 78 studies included in this systematic review. The appendix shows the studies that were excluded in the full text review of the systematic review. The systematic review was presented in Chapter 7. The references for the studies that were excluded, based on inclusion and exclusion criteria, are not included in the reference list of this thesis, unless cited throughout the thesis.

Studies excluded based on Full Text Review

Exclusion criteria: Device testing (*n* = 1).

Jehn, M., Schmidt-Trucksäess, A., Schuster, T., Hanssen, H., Weis, M., Halle, M., & Koehler, F. (2009). Accelerometer-based quantification of 6-minute walk test performance in patients with chronic heart failure: applicability in telemedicine. *Journal of Cardiac Failure, 15*(4), 334-340. doi:10.1016/j.cardfail.2008.11.011

Exclusion criterion: No psychometric properties

Golabchi, A., Basati, F., Kargarfard, M. & Sadeghi, M. (2012). Can cardiac rehabilitation programs improve functional capacity and left ventricular diastolic function in patients with mechanical reperfusion after ST elevation myocardial infarction?: A double-blind clinical trial. *ARYA Atherosclerosis* ***8***(3), 125.

Haeffener, M. P., Ferreira, G. M., Barreto, S. S. M., Arena, R. & Dall'Ago, P. (2008). Incentive spirometry with expiratory positive airway pressure reduces pulmonary complications, improves pulmonary function and 6-minute walk distance in patients undergoing coronary artery bypass graft surgery. *American Heart Journal 156*(5), 900.e1-900.e8. doi:10.1016/j.ahj.2008.08.006

Hirschhorn, A. D., Richards, D., Mungovan, S. F., Morris, N. R. & Adams, L. (2008). Supervised moderate intensity exercise improves distance walked at hospital discharge following coronary artery bypass graft surgery--a randomised controlled trial. *Heart, Lung & Circulation 17*(2), 129-138. doi:10.1016/j.hlc.2007.09.004

Jain, A., Myers, G. H., Sapin, P. M. & O'Rourke, R. A. (1993). Comparison of symptom-limited and low level exercise tolerance tests early after myocardial infarction. *Journal of the American College of Cardiology 22*(7), 1816-1820. doi:10.1016/0735-1097(93)90763-Q

Jelinek, H. F., Huang, A. Q., Khandoker, A. H., Chang, D, & Kiat, H. (2013). Cardiac rehabilitation outcomes following a 6-week program of PCI and CABG Patients. *Frontiers in Physiology Oct*(4), article number 302. *doi*:10.3389/fphys.2013.00302

Jolly, K., Lip, G. Y. H., Taylor, R. S., Raftery, J., Mant, J., Lane, D., . . . Stevens, A. (2009). The Birmingham rehabilitation uptake maximisation study (BRUM): a randomised controlled trial comparing home-based with centre-based cardiac rehabilitation. *Heart, 95*, 36-42. doi:10.1136/hrt.2007.127209

Jolly, K., Taylor, R., Lip, G. Y., Greenfield, S., Raftery, J., Mant, J., . . . & Stevens, A. (2007). The Birmingham Rehabilitation Uptake Maximisation Study (BRUM). Home-based compared with hospital-based cardiac rehabilitation in a multi-ethnic population: Cost-effectiveness and patient adherence. *Health technology assessment 11*(35), 1-118. doi:10.3310/hta11350

Juneau, M., Colles, P. Théroux, P., de Guise, P. Pelletier, G., Lam, J., & Waters, D. (1992). Symptom-limited versus low level exercise testing before hospital discharge after myocardial infarction. *Journal of the American College of Cardiology 20*(4): 927-933. doi:10.1016/0735-1097(92)90195-S

Kervio, G., Ville, N. S., Leclercq, C., Daubert, J. C., & Carré, F. (2004b). Cardiorespiratory adaptations during the six-minute walk test in chronic heart failure patients. *European Journal of Cardiovascular Prevention and Rehabilitation, 11*(2), 171-177. doi:10.1097/01.hjr.0000119964.42813.98

Kloch, B. M., Klocek, M., Czarnecka, D., Wojciechowska, W., Wiliński, J., & Kawecka, J. K. (2011). Impact of cardiac resynchronisation therapy on physical ability and quality of life in patients with chronic heart failure. *Kardiologia polska, 70*(6), 581-588.

Lee, K. W., Blann, A. D., Ingram, J., Jolly, K., & Lip, G. Y. (2005). Incremental shuttle walking is associated with activation of haemostatic and haemorheological markers in patients with coronary artery disease: the Birmingham rehabilitation uptake maximisation study (BRUM). *Heart, 91*(11), 1413-1417. doi:10.1136/hrt.2004.050005

Maniar, S., Sanderson, B. K. & Bittner, V. (2009). Comparison of baseline characteristics and outcomes in younger and older patients completing cardiac rehabilitation. *Journal of Cardiopulmonary Rehabilitation & Prevention 29*(4): 220-229. doi:10.1097/HCR.0b013e3181ac7870

Oerkild, B., Frederiksen, M., Hansen, J. F. & Prescott, E. (2012). Home-based cardiac rehabilitation is an attractive alternative to no cardiac rehabilitation for elderly patients with coronary heart disease: Results from a randomised clinical trial. *BMJ Open 2*(6), e001820. doi:10.1136/bmjopen-2012-001820

Oliveira, B. G., Velasquez-Melendez, G., Rincon, L. G., Ciconelli, R. M., Sousa, L, A., Ribeiro, A. L. (2008). Health-related quality of life in Brazilian pacemaker patients. *Pacing and Clinical Electrophysiology 31*(9), 1178-1183. doi:10.1111/j.1540-8159.2008.01159.x

Scalvini, S., Zanelli, E., Comini, L., Dalla Tomba, M., Troise, G., Febo, O., & Giordano, A. (2013). Home-based versus in-hospital cardiac rehabilitation after cardiac surgery: A non-randomized controlled study. *Physical Therapy 93*(8): 1073-1083.

Solak, Ö., Yaman, F., Ulaşlı, A. M., Eroğlu, S., Akçi, Ö., Özkeçeci, G., . . . & Dündar, Ü. (2015). Improvement in quality of life, functional capacity, and depression level after cardiac rehabilitation. *Turkish Journal of Physical Medicine and Rehabilitation, 61*(2), 130-136. doi:10.5152/tftrd.2015.76093

Stewart, R. A., Szakewsja, D. She, L., Lee, K, L., Drazner, M. H., Lubiszewska, B., . . . & Choudhary, S. K. (2014). Exercise capacity and mortality in patients with ischemic left ventricular dysfunction randomized to coronary artery bypass graft surgery or medical therapy: an analysis from the STICH trial (Surgical Treatment for Ischemic Heart Failure). *JACC Heart Failure 2*(4): 335-343. doi:10.1016/j.jchf.2014.02.009

Witham, M. D., Fulton, R. L., Greig, C. A., Johnston, D. W., Lang, C. C. & Boyers, D. (2012). Efficacy and cost of an exercise program for functionally impaired older patients with heart failure a randomized controlled trial. *Circulation: Heart Failure 5*(2) 209-216. doi:10.1161/CIRCHEARTFAILURE.111.963132

Exclusion criteria: Exercise test not a field test.

Currie, P. J., Kelly, M. J., & Pitt, A.(1983). Comparison of supine and erect bicycle exercise electrocardiography in coronary heart disease: accentuation of exercise-induced ischemic ST depression by supine posture. *American Journal of Cardiology 52*(10): 1167-1173. doi:10.1016/0002-9149(83)90568-4

Dugmore, L. D., Tipson, R. J., Phillips, M. H., Flint, E. J., Stentiford, N. H., Bone, M. F. & Littler, W. A. (1999). Changes in cardiorespiratory fitness, psychological wellbeing, quality of life, and vocational status following a 12 month cardiac exercise rehabilitation programme. *Heart 81*(4): 359-366. doi:10.1136/hrt.81.4.359

Gardener, A. W. and Ades, P. A. (1993). Reliability of cardiopulmonary exercise measures and their response to endurance training in older patients with coronary disease. *Journal of Cardiopulmonary Rehabilitation 13*(3): 188-193.

Johnson, N. P., Wu, E., Bonow, R. O., & Holly, T. A. (2008). Relation of exercise capacity and body mass index to mortality in patients with intermediate to high risk of coronary artery disease. *American Journal of Cardiology 102*(8): 1028-1033. doi:10.1016/j.amjcard.2008.06.017

Kakos, L. S., Szabo, A. J., Gunstad, J., Stanek, K. M., Waechter, D., Hughes, J., ... & Rosneck, J. (2010). Reduced executive functioning is associated with poorer outcome in cardiac rehabilitation. *Preventive cardiology, 13*(3), 100-103. doi:10.1111/j.1751-7141.2009.00065.x

Kulcu, D. G., Kurtais, Y., Tur, B. S., Gulec, S., & Seckin, B. (2007). The effect of cardiac rehabilitation on quality of life, anxiety and depression in patients with congestive heart failure. A randomized controlled trial, short-term results. *Europa Medicophysica, 43*(4), 489-497.

Marburger, C. T., Brubaker, P., Pollock, W., Morgan, T., & Kitzman, D. (1998). Reproducibility of cardiopulmonary exercise testing in elderly patients with congestive heart failure. *American Journal of Cardiology 82*(7), 905-909.

Exclusion criteria: focus of study is prognosis, morbidity or mortality

Cacciatore, F., Abete, P., Mazzella, F., Furgi, G., Nicolino, A., Longobardi, G., ... & Ferrara, N. (2012). Six-minute walking test but not ejection fraction predicts mortality in elderly patients undergoing cardiac rehabilitation following coronary artery bypass grafting. *European Journal of Preventive Cardiology, 19*(6), 1401-1409. doi:10.1177/1741826711422991

Curtis, J. P., Rathore, S. S., Wang, Y., & Krumholz, H. M. (2004). The association of 6-minute walk performance and outcomes in stable outpatients with heart failure. *Journal of Cardiac Failure, 10*(1), 9-14. doi:10.1016/j.cardfail.2003.08.010

Fishbein, D. P., Hellkamp, A. S., Mark, D. B., Walsh, M. N., Poole, J. E., Anderson, J., ... & SCD-HeFT Investigators. (2014). Use of the 6-min walk distance to identify variations in treatment benefits from implantable cardioverter-defibrillator and amiodarone: Results from the SCD-HeFT (Sudden Cardiac Death in Heart Failure Trial). *Journal of the American College of Cardiology 63,* 2560-2568. doi:10.1016/j.jacc.2014.02.602

Freyssin, C., Verkindt, C., Prieur, F., Benaich, P., Maunier, S., & Blanc, P. (2012). Cardiac rehabilitation in chronic heart failure: effect of an 8-week, high-intensity interval training versus continuous training. *Archives of Physical Medicine and Rehabilitation 93*, 1359-1364. doi:10.1016/j.apmr.2012.03.007

Hülsmann, M., Berger, R., Sturm, B., Bojic, A., Woloszczuk, W., Bergler-Klein, J., & Pacher, R. (2002). Prediction of outcome by neurohumoral activation, the six-minute walk test and the Minnesota Living with Heart Failure Questionnaire in

an outpatient cohort with congestive heart failure. *European Heart Journal, 23*(11), 886-891. doi: 10.1053/euhj.2001.3115

Passantino, A., Lagioia, R., Mastropasqua, F., & Scrutinio, D. (2006). Short-term change in distance walked in 6 min is an indicator of outcome in patients with chronic heart failure in clinical practice. *Journal of the American College of Cardiology, 48*(1), 99-105. doi:10.1016/j.jacc.2006.02.061

Exclusion criteria: Testing surgical/medical intervention

Chan, K. L., Tang, A. S., Achilli, A., Sassara, M., Bocchiardo, M., Gaita, F., . . . & Gras, D. (2003). Functional and echocardiographic improvement following multisite biventricular pacing for congestive heart failure. *The Canadian Journal of Cardiology, 19*(4), 387-390.

Gotzmann, M., Hehen, T., Germing, A., Lindstaedt, M., Yazar, A., Laczkovics, A., . . . & Bojara, W. (2010). Short-term effects of transcatheter aortic valve implantation on neurohormonal activation, quality of life and 6-minute walk test in severe and symptomatic aortic stenosis. *Heart, 96*(14), 1102-1106. doi:10.1136/hrt.2009.180661

Green, P., Cohen, D. J., Généreux, P., McAndrew, T., Arnold, S. V., Alu, M., . . . & Dvir, D. (2013). Relation between six-minute walk test performance and outcomes after transcatheter aortic valve implantation (from the PARTNER trial). *The American Journal of Cardiology, 112*(5), 700-706. doi:10.1016/j.amjcard.2013.04.046

Studies found on Manual Search Strategies

Cahalin, L. P., Mathier, M. A., Semigran, M. J., Dec, G. W., & DiSalve, T. G. (1996). The six-minute walk test predicts peak oxygen uptake and survival in patients with advanced heart failure. *Chest, 110*(2), 325-332. doi:10.1378/chest.110.2.325

Cervie, P., Olper, L., De Santi, F., & Pierini, S. M., C. (2012). Reliability and minimal important difference of the treadmill six minute walking test in people following cardiac surgery. *Italian Journal of Physiotherapy, 2*(4), 142-150.

Chiaranda, G., Myers, J., Mazzoni, G., Terranova, F., Bernardi, E., Grossi, G., . . . Grazzi, G. (2012). Peak oxygen uptake prediction from a moderate, perceptually regulated, 1-km treadmill walk in male cardiac patients. *Journal of Cardiopulmonary Rehabilitation and Prevention, 32*(5), 262-269. doi:10.1097/HCR.0b013e3182663507

Delahaye, N., Cohen-Solal, A., Faraggi, M., Czitrom, D., Foult, J. M., Daou, D., . . . Le Guludec, D. (1997). Comparison of left ventricular responses to the six minute walk test, stair climbing, and maximal upright bicycle exercise in patients with congestive heart failure due to idiopathic dilated cardiomyopathy. *The American Journal of Cardiology, 80*(1), 65-79. doi: 10.1016/S0002-9149(97)00285-3

Gayda, M., Temfemo, A., Choquet, D., & Ahmaidi, S. (2004). Cardiorespiratory requirements and reproducibility of the six-minute walk test in elderly patients with coronary artery disease. *Archives of Physical Medicine and Rehabilitation, 85*(9), 1538-1543. doi:10.1016/j.apmr.2003.11.037

Guazzi, M., Dickstein, K., Vicenzi, M., & Arena, R. (2009). Six-minute walk test and cardiopulmonary exercise testing in chronic heart failure: A comparative analysis on clinical and prognostic insights. *Circulation, 2*(6), 549-555. doi:10.1161/CIRCHEARTFAILURE.109.881326

Hanson, L. C., Taylor, N. F., & McBurney, H. (2016). The 10 m incremental shuttle walk test is a highly reliable field exercise test for patients referred to cardiac rehabilitation: a retest reliability study. *Physiotherapy, 102*(3), 243-248. doi:10.1016/j.physio.2015.08.004

Houghton, A. R., Harrison, M., Cowley, A. J., & al, e. (2002). Assessing exercise capacity, quality of life and haemodynamics in heart failure: Do the tests tell us the same thing. *European Journal of Heart Failure, 4(3)*, 289-295. doi:10.1016/S1388-9842(01)00236-7

Ingle, L., Rigby, A. S., Nabb, S., Jones, P. K., Clark, A. L., & Cleland, J. G. (2006). Clinical determinants of poor six-minute walk test performance in patients with left ventricular systolic dysfunction and no major structural heart disease. *European Journal of Heart Failure, 8*, 321-325. doi:10.1016/j.ejheart.2005.08.006

Lewis, M. E., Newal, C., Townend, J. N., Hill, S. L., & Bonser, R. S. (2001). Incremental shuttle walk test in the assessment of patients for heart transplantation. *Heart, 86*, 183-187. doi:10.1136/heart.86.2.183

Peeters, & Mets. (1996). The 6-minute walk as an appropriate exercise testing in elderly patients with chronic heart failure. *Journal of Gerontology: Medical Sciences, 51*, M147-151. doi:10.1093/gerona/51A.4.M147

Riley, M., McParland, J., Stanford, C. F., & Nicholls, D. P. (1992). Oxygen consumption during corridor walk testing in chronic cardiac failure. *European Heart Journal, 13*(6), 789-793. doi:10.1136/thx.54.3.213

Verrill, D., Barton, C., Beasley, W., Lippard, M., & King, C. (2003). Six-minute walk performance and quality of life comparisons in North Carolina cardiac rehabilitation programs. *Heart Lung, 32*, 41-51. doi: 10.1067/mhl.2003.7

Wright, D. J., Khan, K. M., Gossage, E. M., & Saltissi, S. (2001). Assessment of a low-intensity cardiac rehabilitation programme using the six-minute walk test. *Clinical Rehabilitation, 15*, 119-124. doi:10.1191/026921501671364449

Appendix 9: Results of the Quality Review

This appendix reports the results of the methodological quality assessment of each of the studies included in the systematic review presented in Chapter 7 of the thesis. A methodological quality review was performed on each measurement property addressed in each study retrieved for the systematic review.

Two examiners completed the methodological quality assessment independently and then discussed the results. The results of the methodological assessment after discussion are presented for relative reliability (Table A9.1), measurement error (Table A9.2), criterion validity (Table A9.3), construct validity (Table A9.4) and responsiveness (Table A9.5).

Table A9.1

Summary of COSMIN Quality Scoring for Box B. Reliability: Relative Measures

First author (date)	Field test	Questions from Box B. Reliability: Relative Measures														Worst score
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	
Bellet 2011	6MWT	+++	++	+	+++	++	+++	+	+	+++	+++	++	NA	NA	NA	+
Cahalin 1996	6MWT	++	++	0	+++	++	+++	++	+++	++	+++	++	NA	NA	NA	++*
Carvalho 2011	6MWT	++	+	0	+++	++	+++	++	+++	+++	+++	+	NA	NA	NA	+*
Cervie 2012	Other	+++	+++	+	+++	++	+++	+	+++	++	+++	+++	NA	NA	NA	+
Corvea-Tindel 2009	6MWT	+++	++	0	+++	+	+	+	+	+	+	+	NA	NA	NA	+*
De Greef	Other	++	+	+	+++	+	+++	++	+++	+++	+++	+++	NA	NA	NA	+
Demers 2001	6MWT	++	+	+++	+++	++	+++	++	+++	+	+++	+++	NA	NA	NA	+
Fowler 2005	ISWT	+++	++	+	+++	++	+++	++	+++	++	+++	++	NA	NA	NA	++*
Green 2001	ISWT	++	+	0	+++	++	+++	++	+++	++	+	++	NA	NA	NA	+*
Gremeaux 2012	Other	++	+	+	+++	++	+	++	+	+	+	++	NA	NA	NA	+
Guimarães 2008	Other	++	+	0	+++	++	+++	++	+++	++	+++	++	NA	NA	NA	+*
Hamilton 2000	6MWT	+++	++	++	+++	++	+	++	+	++	+++	+++	NA	NA	NA	+
Hanson 2012	6MWT	+++	++	+^	+++	++	+++	++	+++	+++	+++	+++	NA	NA	NA	++*
Hanson 2016	ISWT	+++	+++	++^	+++	++	+++	++	+++	+++	+++	+++	NA	NA	NA	++
Ingle 2005	6MWT	+++	++	++	+++	+++	+++	+++	+	+++	+++	++	NA	NA	NA	+
Jolly 2008	ISWT	+++	++	+++	+++	++	+++	++	+++	++	+++	++	NA	NA	NA	++
Kervio 2004 A	6MWT	+++	++	0	+++	++	+++	++	+++	+	+++	++	NA	NA	NA	+*

First author (date)	Field test	Questions from Box B. Reliability: Relative Measures														Worst score
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	
Lewis 2001	ISWT	+++	++	0	+++	++	++	++	+++	+++	+++	+	NA	NA	NA	+*
Mandic Walker 2013	6MWT	+++	++	++	+++	++	+++	+	+	++	+	+	NA	NA	NA	+
Morales 1999	6MWT	++	++	0	+++	++	+	+	++	+	+++	+	NA	NA	NA	+*
	ISWT	++	++	0	+++	++	+	+	++	+	+++	+	NA	NA	NA	+*
Nogueira 2006	6MWT	+++	+++	0	+++	++	+++	++	+++	++	+++	++	NA	NA	NA	++*
O'Keefe 1998	6MWT	+++	+++	0	+++	++	+++	+	+	++	+	+++	NA	NA	NA	+*
Olper 2011	6MWT	+++	++	0^	+++	++	+++	++	+++	++	+++	+++	NA	NA	NA	++*
	Other	+++	++	0^	+++	++	+++	++	+++	++	+++	+++	NA	NA	NA	++*
Pinna 2000	6MWT	+++	+++	+++	+++	++	+++	++	+++	++	+	++	NA	NA	NA	+
Zugck 2000	6MWT	++	++	0	+++	++	+++	++	+++	++	+++	++	NA	NA	NA	++*

Note. +++ excellent, ++ good + fair 0 poor * Second worst score, denotes a change of COSMIN score after the removal of sample size item from the ratings. ^ adequate justification of sample size

Table A9.2

Summary of COSMIN Quality Scoring for Box C. Measurement Error

First author (date)	Field test	Questions from Box C. Measurement error: absolute measures											Worst score
		1	2	3	4	5	6	7	8	9	10	11	
Bellet 2011	6MWT	+++	++	+	+++	++	+++	++	+	+++	+++	+++	+
Cervie 2012	Other	+++	+++	+	+++	++	+++	+	+++	++	+++	+++	+
De Greef 2005	Other	++	+	+	+++	+	+++	++	+++	+++	0	+++	0
Fowler 2005	ISWT	+++	++	+	+++	++	+++	++	+++	++	+++	++	++*
Guimarães 2008	Other	++	+	0	+++	+++	+++	++	+++	++	+++	++	+*
Hanson 2012	6MWT	+++	+++	+	+++	++	+++	++	+++	+++	+++	+++	++*
Hanson 2016	ISWT	+++	+++	++^	+++	++	+++	++	+++	+++	+++	+++	++
Ingle 2005	6MWT	+++	++	++	+++	+++	+++	+++	+	+++	+++	+++	+
Jolly 2008	ISWT	+++	+++	+++	+++	++	+++	++	+++	++	+++	+++	++
Kervio 2004a	6MWT	+++	++	0	+++	++	+++	++	+++	+	+++	0	0
Nogueira 2006	6MWT	+++	+++	0	+++	++	+++	++	+++	++	+++	+++	++*
Olper 2011	6MWT	+++	++	0^	+++	++	+++	++	+++	++	+++	+++	++*
	Other	+++	++	0^	+++	++	+++	++	+++	++	+++	+++	++*
Opasich 1998	6MWT	++	++	+++	+++	++	+++	++	+++	++	+	+++	+
Pinna 2000	6MWT	+++	+++	+++	+++	++	+++	++	+++	++	+	+++	+
Pulz 2008	6MWT	++	++	++	+++	++	+++	++	+++	++	+++	+++	++
	ISWT	++	++	++	+++	++	+++	++	+++	++	+++	+++	++
Riley 1992	6MWT	++	++	0	+++	+++	+++	++	+++	++	+++	0	0

Note. +++ excellent, ++ good + fair 0 poor * Second worst score, denotes a change of COSMIN score after the removal of sample size item from the ratings. ^ adequate justification of sample size

Table A9.3

Summary of COSMIN Quality Scoring for Box H. Criterion Validity

First author Date	Field Test	Questions from Box H. Criterion validity							Worst score
		1	2	3	4	5	6	7	
Beatty 2012	6MWT	+++	+++	+++	+	+++	+++	NA	+
Cahalin 1996	6MWT	++	++	+	+++	+++	+++	NA	++*
Carvalho 2011	6MWT	++	+	0	+++	+++	+++	NA	+*
Casillas 2015	6MWT	+++	++	+++	+++	+++	+++	NA	++
	Other	+++	++	+++	+++	+++	+++	NA	++
Cheetham 2005	6MWT	++	+	0	+++	+++	+++	NA	+*
Chiaranda 2012	Other	+++	+++	+	+++	+	+++	NA	+
De Greef 2005	Other	++	+	0	+++	+++	+++	NA	+*
De Sousa 2008	6MWT	++	+	0	+++	+++	+++	NA	+*
Delahaye 1997	6MWT	++	++	0	+++	+++	+++	NA	++*
	Other	++	++	0	+++	+++	+++	NA	++*
Doutreleau 2009	6MWT	++	+	0	+++	+++	+++	NA	+*
Forman 2012	6MWT	+++	+++	+++	+++	+++	+++	NA	+++
Fowler 2005	ISWT	+++	+++	+	+++	+++	+++	NA	+++*
Gayda 2004	6MWT	++	+	0	+++	+	+++	NA	+*
	Other	++	+	0	+++	+	+++	NA	+*
Green 2001	6MWT	++	+	0	+++	+	+++	NA	+*
Green 2001	ISWT	++	+	0	+++	+	+++	NA	+*
Gremeaux 2012	Other	++	+	+	+++	+	+++	NA	+

First author Date	Field Test	Questions from Box H. Criterion validity							Worst score
		1	2	3	4	5	6	7	
Guazzi 2009	6MWT	+++	+++	+++	+++	+	+++	NA	+
Guyatt 1985	6MWT	+++	++	0	+	+++	+++	NA	+*
Hamilton 2000	6MWT	+++	++	++	++	+	+++	NA	+
Houghton 2002	Other	++	++	0	+	+++	+++	NA	+*
Jehn 2009	6MWT	+++	++	+	+++	+++	+++	NA	++*
Kervio 2004a	6MWT	+++	++	0	+++	+++	+++	NA	++*
Kristjansdottir 2004	6MWT	++	++	0	++	+++	+++	NA	++*
Langenfeld 1990	6MWT	+++	++	+	+++	+	+++	NA	+
Lewis 2001	ISWT	+++	++	0	+++	+++	+++	NA	++*
Lipkin 1986	6MWT	++	++	0	+++	+++	+++	NA	++*
Lucas 1999	6MWT	++	++	+++	+++	+	+++	NA	+
Maldonado-Martin 2006	6MWT	++	++	++	++	+++	+++	NA	++
Mandic Walker 2013	6MWT	++	++	++	+++	+++	+++	NA	++
	ISWT	++	++	++	+++	+++	+++	NA	++
Meyer 2003	Other	++	++	++	++	+	+++	NA	+
Morales 1999	6MWT	++	++	+	+++	+++	+++	NA	++*
Morales 2000	6MWT	++	++	+	+++	+++	+++	NA	++*
Opasich 2001	6MWT	++	++	+++	+++	+	+	NA	+
Peeters 1996	6MWT	+++	++	0	+	+	+++	NA	+
Pulz 2008	6MWT	++	++	++	+++	+++	+++	NA	++
	ISWT	++	++	++	+++	+++	+++	NA	++
Riley 1992	6MWT	++	++	0	+++	+++	+++	NA	++

First author Date	Field Test	Questions from Box H. Criterion validity							Worst score
		1	2	3	4	5	6	7	
Zugck 2000	6MWT	++	++	+++	+++	+++	+++	NA	++

Note. +++ excellent, ++ good + fair 0 poor * Second worst score, denotes a change of COSMIN score after the removal of sample size item from the ratings. ^ adequate justification of sample size

Table A9.4

Summary of COSMIN Quality Scoring for Box F. Hypothesis Testing (Construct Validity)

First author (date)	Field test	Questions from Box F. Hypothesis testing (construct validity)										Worst score
		1	2	3	4	5	6	7	8	9	10	
Ades 2003	6MWT	+++	+	+	+	++	++	+++	0	+	+++	0
Allison 2004	6MWT	+++	+++	++	++	+++	++	+++	+	+	+++	+
Alosco 2012	Other	++	+	+++	+	+++	++	++	+	+	+++	+
Araya-Ramírez 2010	6MWT	++	+	+++	0	++	++	+	0	+	+++	0
Bajraktari 2011	6MWT	++	+	++	+	++	++	++	NA	+++	+++	+
Baldasseroni 2014	6MWT	++	+	+++	+	++	++	+	NA	+++	+++	+
Baptista 2012	6MWT	+++	++	++	0	++	++	+++	0	+++	+++	0
Beatty 2012	6MWT	+++	++	+++	+	++	++	NA	NA	+	+++	+
Bittner 1993	6MWT	++	++	+++	0	++	++	++	NA	+++	+++	0
Bittner 2000	6MWT	+++	++	+++	0	++	++	+	0	+	+++	0
Chien 2011	6MWT	++	+	+	+	++	++	++	+	0	+++	0
Cipriano 2010	6MWT	+++	+++	0	0	++	++	+	+	+++	++	0
Corvea-Tindel 2009	6MWT	++	+	++	+	++	++	++	+	+++	++	+
Delahaye 1997	6MWT	++	++	0	0	++	++	+	0	+++	+++	0
	Other	++	++	0	0	++	++	+	0	+++	+++	0
Demers 2001	6MWT	++	+	+++	+	++	++	+	+	+++	+++	+
Flynn 2009	6MWT	++	+	+++	++	++	++	++	+	+	++	+
Forman 2012	6MWT	+++	+++	+++	0	++	++	0	0	+++	+++	0
Gremeaux 2012	6MWT	++	+	+	+	++	++	+	0	+	+++	0

	Other	++	+	+	+	++	++	+	0	+	+++	0
Guazzi 2009	6MWT	+++	+++	+++	+	++	++	+	0	+	+++	0
Guyatt 1985	6MWT	+++	++	0	0	++	++	+	+	0	+++	0
Hamilton 2000	6MWT	+++	++	++	+	++	++	++	+	+++	+++	+
Houghton 2002	Other	++	++	0	0	++	++	++	+	+++	+++	0
Ingle 2006	6MWT	++	++	+++	0	++	++	0	0	+++	+++	0
Juenger 2002	6MWT	++	++	+++	+	++	++	+++	++	+++	+++	+
Karapolat 2008	6MWT	+++	+++	+++	+/0	++	++	++	0	+	+++	0
Mandic Hodge 2013	ISWT	+++	+++	+	0	++	++	+++	0	+	+++	0
Nogueira 2010	Other	++	++	+	0	++	++	+++	+	+	+++	0
O'Keefe 1998	6MWT	+++	+++	0	+	++	++	+++	0	+++	+++	0
Olper 2011	6MWT	+++	++	0	+	++	++	+++	+++	+	+++	+*
	Other	+++	++	0	+	++	++	+++	+++	+	+++	+*
Opasich 2004	6MWT	+++	++	+++	0	++	++	0	0	+++	+++	0
Pepera 2013	ISWT	++	++	0	0	++	++	+	+	+	+++	0
Pulz 2008	6MWT	++	++	++	0	++	++	+++	+++	+++	+++	0
	ISWT	++	++	++	0	++	++	+++	+++	+++	+++	0
Verrill 2003	6MWT	++	++	+++	0	++	++	+++	++	+++	+++	0
Westlake 2005	6MWT	+++	+++	+++	+	++	++	+++	+++	+	+++	+

Note. +++ excellent, ++ good + fair 0 poor * Second worst score, denotes a change of COSMIN score after the removal of sample size item from the ratings. ^ adequate justification of sample size

Table A9.5

Summary of COSMIN Quality Scoring for Box I. Responsiveness

First author (date)	Field Test	Questions from Box I. Responsiveness																		Score
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	
Araya-	6MW	++	+	++	++	0	++	++	+	++	++	+	0	+	++	NA	NA	NA	NA	0
Ramírez	T			+	+		+	+							+					
2010																				
Bellet 2011	6MW	++	++	+	++	++	++	+	0	++	++	NA	NA	++	+	NA	NA	NA	NA	0
	T	+			+	+	+							+						
Bittner 2000	6MW	++	++	++	++	++	++	++	0	++	++	+	0	+	+	NA	NA	NA	NA	0
	T	+		+	+	+	+	+												
Cervie 2012	Other	++	++	+	++	++	++	++	+	++	++	+	0	0	+	NA	NA	NA	NA	0
		+	+		+	+	+	+												
Cheetham	6MW	++	+	0	++	++	++	+	NA	NA	NA	NA	NA	NA	NA	++	++	++	NA	+*
2005	T				+	+	+									+	+	+		
Demers	6MW	++	+	++	++	++	+	++	0	++	++	++	++	+	+	NA	NA	NA	MA	0
2001	T			+	+	+		+				+	+							
Fowler	ISW	++	++	0	++	++	++	++	0	++	++	+	+	+	+	NA	NA	NA	NA	0
2005	T	+	+		+	+														
Gary 2004	6MW	++	+	+	++	++	++	+	++	++	++	+	+	+	+	NA	NA	NA	NA	+
	T				+	+	+		+	+										
Gremeaux	6MW	++	+	+	++	++	++	++	+	++	++	++	0	+	+	NA	NA	NA	NA	0
2009	T				+	+	+	+				+								

First author (date)	Field Test	Questions from Box I. Responsiveness																		Score
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	
	Other	++	+	+	++	++	++	++	+	++	++	++	0	+	+	NA	NA	NA	NA	0
					+	+	+	+				+								
Gremeaux	6MW	++	+	0	++	++	++	++	+	++	++	++	+	+	+	NA	NA	NA	NA	+*
2011	T				+	+	+					+								
Gremeaux	6MW	++	+	+	++	++	++	++	+	++	++	+	++	+	+	NA	NA	NA	NA	+
2012	T				+	+	+	+					+							
	Other	++	+	+	++	++	++	++	+	++	++	+	++	+	+	NA	NA	NA	NA	+
					+	+	+	+					+							
Houchen-	ISW	++	+	++	++	++	++	++	+	++	++	++	+	++	++	NA	NA	NA	NA	+
Wolloff	T			+	+	+	+	+				+		+	+					
2015																				
Ingle 2005	6MW	++	++	++	++	++	++	++	+	++	++	++	+	++	++	NA	NA	NA	NA	+
	T	+		+	+	+		+				+		+	+					
Kavanagh	6MW	++	++	0	++	++	++	++	0	++	++	++	+	++	++	NA	NA	NA	NA	0
1996	T	+	+		+	+	+	+				+		+	+					
Mandic	ISW	++	++	+	++	++	++	++	0	++	++	++	0	+	++	NA	NA	NA	NA	0
Hodge 2013	T	+	+			+	+	+				+			+					
Meyer 2003	Other	++	++	+	++	++	++	++	+	++	++	++	0	+	++	NA	NA	NA	NA	0
		+	+		+	+	+	+				+			+					
O'Keefe	6MW	++	++	+	++	++	++	++	0	++	++	++	0	+	++	NA	NA	NA	NA	0
1998	T	+			+	+	+	+							+					
Olper 2011	6MW	++	++	+	++	++	+	++	+	++	++	++	++	0	0	NA	NA	NA	NA	+
	T	+			+	+														

First author (date)	Field Test	Questions from Box I. Responsiveness																		Score
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	
	Other	++	++	+	++	++	+	++	+	++	++	++	++	0	0	NA	NA	NA	NA	0
		+			+	+														
Verrill 2003	6MW	++	++	++	++	+	++	++	0	++	++	+	+	++	++	NA	NA	NA	NA	0
	T			+	+			+						+	+					
Wright	6MW	++	++	++	++	++	++	++	0	++	++	++	NA	+	++	NA	NA	NA	NA	0
2001	T	+	+	+	+	+	+	+				+			+					

Note. +++ excellent, ++ good + fair 0 poor * Second worst score, denotes a change of COSMIN score after the removal of sample size item from the ratings. ^ adequate justification of sample size

Appendix 10: Study Results of Measurement Properties

For the systematic literature review, two reviewers extracted data relating to results of the measurement properties independently and then discussed each study for accuracy. This appendix summarises the results of the measurement properties presented in each study for the 6MWT (Table A10.1), the 10 m ISWT (Table A10.2) and the other field exercise tests (Table A10.3).

Table A10.1

Results of the Measurement Properties of the 6MWT

First author (date)	Method. Quality Review	Results	Comments
Relative Reliability			
Bellet 2011	Fair	Baseline *ICC* .93, 95% CI [.90, .95]; 6/52 Follow-up *ICC* .95, 95% CI [.91, .98]; 6/12 Follow-up *ICC* .91, 95% CI [.82, .95]	2 6MWTs, Same day
Cahalin 1996	Good*	*ICC* .96	2 6MWTs, Same day
Carvalho 2011	Fair*	*r* (Spearman's rho) = .93	2 6MWTs, Same day
Corvea-Tindel 2009	Fair*	*r* = .92	No detail
Demers 2001	Fair	Baseline *ICC* .9; 18/52-Follow-up *ICC* .88; 43/52-Follow-up *ICC* .91; All walks *ICC* .81 (random effects model) and .80 (mixed effects model)	2 6MWTs each session
Hamilton 2009	Fair	*ICC* .97	3 6MWTs, non-consecutive days
Hanson 2012	Good*	Walk-1-2-3 $ICC_{2,1}$ .94	3 6MWTs, 1 week
Ingle 2005	Fair	*ICC* .80, 95% CI [.69, .87]	2 6MWTs, 1 year
Kervio 2004	Fair*	CHF_D *ICC* .99; CHF_P *ICC* .98	2 6MWTs, same day
Mandic Walker 2013	Fair	*r* = .95	2 6MWTs, same day
Morales 1999	Fair*	Walk-2-3 *r* = .98	3 6MWTs, 2 weeks
Nogueira 2006	Good*	*ICC* .88, 95% CI [.79, .94]	3 6MWTs, same day
O'Keefe 1998	Fair*	*ICC* .91	2 6MWTs 3-8 weeks

First author (date)	Method. Quality Review	Results	Comments
Olper 2011	Good*	Baseline *ICC* .96, 95% CI [.90, .98]; Follow-up *ICC* .96, 95% CI [.91, .98]	3 6MWTs, same day
Pinna 2000	Fair	*ICC* .96	2 6MWTs, same day Eliminated outliers
Zugck 2000	Good*	*ICC* .96	3 6MWTs, consecutive days
Measurement Error			
Bellet 2011	Fair	Walk-1-2 95%LOA -.40 to 71 m	2 6MWTs, Same day
Hanson 2012	Good*	Walk-1-2 95%LOA -2 to 106 m; Walk-2-3 95%LOA -29 to 95 m	3 6MWTs, same day -1 week
Ingle 2005	Fair	95%LOA -162 to 145	
Kervio 2004	Poor	CHFD CV 2.0%; CHFP CV 1.9%	2 6MWTs, same day
Nogueira 2006	Good	Walk-1-2 95%LOA -48 to 84 m; Walk-2-3 95%LOA -33 to 49 m	3 6MWTs, same day
Olper 2011	Good*	Walk-2-3 Pre-rehabilitation SEM 19 m (SDC 53 m), Walk-2-3 Post-rehabilitation SEM 14 m (39 m)	2 6MWTs, 1 year
Opasich 1998	Fair	Walk-1-2 SEM 15 m, real change 10% mean walk test performed in a single session (with 99% confidence).	2 6MWTs, Same day – 1 day
Pinna 200	Fair	SEM 16 m	2 6MWTs, Same day; Eliminated outliers
Pulz 2008	Good	7 ± 40 m (i.e., -33 to 47 m)	2 6MWTs, same day (30 min rest)
Riley 1992	Poor	Walk-1-2 CV 9%; Walk-2-3 CV 7%	3 tests, each 1 week apart No SEM/LOA
Criterion Validity (concurrent)			
		Association of 6MWT distance unless otherwise specified.	

First author (date)	Method. Quality Review	Results	Comments
Beatty 2012	Fair	METS *r* = .66	Criterion: Treadmill
Cahalin 1996	Good*	R$\dot{V}O_2$ *r* = .64	Criterion: Bicycle
Carvalho 2011	Fair*	Walk-2: A$\dot{V}$O2 r = .70	Criterion: Bicycle
Casillas 2015	Good	HR_{max} .*r* = .23	Criterion: Treadmill
Cheetham 2005	Fair*	R$\dot{V}O_2$ *r* = .81 A$\dot{V}O_2$ *r* = .70	Criterion: Treadmill
De Sousa2008	Fair*	Est VO_{2Max} *r* = .71	Criterion: Treadmill
Delahaye 1997	Good*	R$\dot{V}O_2$.*r* = .61 Watts *r* = .60	Criterion: Bicycle
Doutreleau 2009	Fair*	A$\dot{V}O_2$ *r* = .46 W_{peak} *r* = .42 $VO_{2@VT}$ *r* = .53; 6MWTD x weight (kg) and A$\dot{V}O_2$.r = .74.	Criterion: Bicycle
Forman 2012	Excellent	R$\dot{V}O_2$.*r* = .54 VE/VCO_2 slope *r* = -.26	Criterion: Bicycle
Gayda 2004	Fair*	R$\dot{V}O_2$.*r* = .56	Criterion: Bicycle
Green 2001	Fair*	R$\dot{V}O_2$.*r* = .67	Criterion: Treadmill
Guazzi 2009	Fair	A$\dot{V}O_2$.*r* = .68 $\dot{V}O_{2.@AT}$ *r* = .63 VE/VCO_2 slope *r* = -.38	Criterion: Bicycle
Guyatt 1985	Fair*	Results of cycle test *r* = .42	Criterion: Bicycle
Hamilton 2000	Fair	METs *r* = .69	Criterion: Treadmill
Jehn 2009	Good*	R$\dot{V}O_2$.*r* = .77, 6MWT R$\dot{V}O_2$. and CPET R$\dot{V}O_2$.*r* = .72	Criterion: Bicycle
Kervio 2004a	Good*	CHFD R$\dot{V}O_2$.*r* = .88, VT *r* = .81, CHFP R$\dot{V}O_2$.*r* = .62 and VT *r* = .77	Criterion: Treadmill

First author (date)	Method. Quality Review	Results	Comments
Kristjansdottir 2004	Good*	W_{max} r = .93	Criterion: Bicycle
Langenfeld 1990	Fair	6MWT (Watts) and CPET (Watts) r = .74	Criterion: Bicycle
Lipkin 1986	Good*	Estimated from data r = .69 for whole sample (n = 26)	Criterion: Treadmill
Lucas 1999	Fair	R$\dot{V}O_2$.r = .52, VO_2@AT r = .39, VO_2 10-20ml/kg/min: R$\dot{V}O_2$.r = .28	Criterion: Bicycle
Maldonado-Martin 2006	Good	R$\dot{V}O_2$.r = .54, SHF r = .62, DHF r = .45, VT r = .23; R$\dot{V}O_2$ = 0.01426 x 6MWT(m) + 7.222	Criterion: Bicycle
Mandic Walker 2013	Good	R$\dot{V}O_2$.r = .72	Criterion: Bicycle
Morales 1999	Good*	R$\dot{V}O_2$.r = .69	Criterion: Bicycle
Morales 2000	Good*	R$\dot{V}O_2$.r = .69	Criterion: Bicycle
Opasich 2001	Fair	R$\dot{V}O_2$.r = .59, NYHA II r = .46and III r = .38. Distance ↓in NYHA III than II (p < .001)	Criterion: Bicycle
Peeters 1996	Fair	Treadmill distance Kendall's Tau = .688 (includes data from 9 elderly controls)	Criterion: Treadmill
Pulz 2008	Good	R$\dot{V}O_2$.r = .76	Criterion: Treadmill
Riley 1992	Good	R$\dot{V}O_2$ r = .88	Criterion: Treadmill
Zugck 2000	Good	Walk 1 R$\dot{V}O_2$.r = .68 (♂ r = .69, ♀ r = .59), Walk 2 R$\dot{V}O_2$.r = .71, Walk 3 R$\dot{V}O_2$.r = .74	Criterion: Bicycle
Criterion Validity (Predictive)			
Cahalin	Good*	$R\dot{V}O_2 = (0.03 \times \text{ISWT(m)}) + 3.98$	Criterion: Bicycle

First author (date)	Method. Quality Review	Results	Comments
Cheetham	Fair*	$R\dot{V}O_2 = (0.038 \times 6MWT(m)) + 0.9906$	Criterion: Treadmill
Guazzi 2009	Fair	$A\dot{V}O_2 = \left(3.3216 \times 6MWT(m)\right) - 20.052$	Criterion: Bicycle
Maldonado-Martin 2006	Good	$R\dot{V}O_2 = (0.01426 \times 6MWT(m)) + 7.222$	Criterion: Bicycle
Mandic Walker 2013	Good	$R\dot{V}O_2 = (0.029 \times 6MWT(m)) + 2.109$ $R\dot{V}O_2 = (0.015 \times 6MWT(m)) + (0.239 \times$ Chair stands in 30 sec $(n)) - \left(0.218 \times$ Body fat $(\%)\right) + 12.258$	Criterion: Bicycle
Morales 1999	Good*	6MWT(m) not independent predictor of $R\dot{V}O_2$	Criterion: Bicycle
Pulz 2008	Good	Area under ROC = .89. Cut-off to predict $R\dot{V}O_2 < 14$ = 490 m (sens 83%, spec 83%)	Criterion: Treadmill
Construct Validity (hypothesis testing)			
Ades 2003	Poor	MOS SF 36: "Walk a mile" *r* = .62; "walk several blocks" *r* = .54, "walk 1 block" *r* = .38	
Allison 2004	Fair	Self-efficacy *r* = .44, PASE *r* = .30; 6/52: Self-efficacy *r* = .84; PASE r = .49; 12/52: Self-efficacy *r* = .84, PASE *r* = .61	
Araya-Ramírez 2010	Poor	↑Weight in highest tertile group (6MWT > 436 m) compared with middle and lower tertile ($p = .014$); ↑MOS SF-36 PCS in highest tertile group ($p < .001$); ↑SF-36 MCS in highest tertile group ($p = .005$)	

First author (date)	Method. Quality Review	Results	Comments
Bajraktari 2011	Fair	Age *r* = -.355 BMI *r* = -.108, waist hip ratio r = .107, NYHA class *p* < .002	
Baldesseroni 2014	Fair	Age *r* = -.47 ♀ < ♂ (*p* = .001)	
Baptista 2012	Poor	SF36PF *r* = .29; SF36P *r* = .22; SF36V *r* = .27	
Beatty 2012	Fair	Lower quartile more likely to be ♀, ↑age, ↑BMI (*p* < .05)	
Bittner 1993	Poor	Age *r* = -.34 ♀ < ♂ (*p* < .001)	
Bittner 2000	Poor	Age older < younger (*p* = .002); BMI ns; Self-reported activity (*p* = .015); MOS SF36 Physical function (*p* < .0001), Role Physical ns	
Chien 2011	Poor	HADS-A *r* = -.31; HADS-D *r* = -.40	
Cipriano 2015	Poor	LVEF *r* = .70; NYHA *r* = .73; MLHFQ no correlation	
Corvea-Tindel 2009	Fair	HFFSI *r* = .36; Age *r* = -.22; LVEF *r* = -.34; C-QLI .04; HFSS .13; PSS.07; SESS=.14; MAAC-D *r* = -.02; MAAC-A *r* = -.13; MAAC-H *r* = -.09	
Delahaye 1997	Poor	Stair climbing time = -.82	
Demers 2001	Fair	Disease specific QOL *r* = -.26; NYHA *r* = -.43	
Flynn 2009	Fair	euroQol VAS *r* = .11, 95% CI [.07, .15]; KCCQ overall score *r* = .27, 95% CI [.23, .31]; KCCQ clinical summary *r* = .322, 95% CI [.29, .36]; KCCQ PL *r* = .35, 95% CI [.31, .38]; KCCQ TS *r* = .24, 95% CI [.21, .28]; KCCQ SL *r* = .21, 95% CI [.17, .25]; KCCQ SE *r* = .05, 95% CI [.01, .09];	

First author (date)	Method. Quality Review	Results	Comments
		KCCQ QOL *r* = .17, 95% CI [.13, .21]	
Forman 2012	Poor	Age *r* = -.23; BMI *r* = -.13; LVEF *r* = .06; BDI II *r* = -.12	
Gremeaux 2012	Poor	200mFWT time Pre-rehabilitation: *r* = -.57; Post-rehabilitation *r* = -.93	
Guazzi 2009	Poor	LVEF% *r* = .06	
Guyatt 1985	Poor	NYHA *r* = -.45; SAS *r* = -.37	
Hamilton 2000	Fair	Age older < younger; ♂>♀ ($p < .01$); NYHA *r* = -.60; DASI *r* = .50; MOS SF36 *r* = .47; MOS SF36PF *r* = .62; MOS SF36RF *r* = .36; MOS SF36BP *r* = .33; MOS SF36GH *r* = .21; MOS SF36V *r* = .35; MOS SF36RE *r* = .23	
Ingle 2006	Poor	Independent predictors of ↓ 6MWT (m): Age ≥75 years (OR 4.0, 95% CI [2.4, 6.4]; BMI <20 (OR 3.4, 95% CI [1.6, 7.3]; ♀ (OR=2.0, 95% CI [1.3, 3.0]; self-perceived depression and anxiety; worsening feelings of health status	
Juenger 2001	Fair	MOS SF36PF *r* = -.56 MOS SF36RP *r* = .28' MOS SF36BP *r* = .26, MOS SF36GH *r* = .34; MOS SF36V *r* = .41; MOS SF36SF *r* = .29; MOS SF36RE *r* = .24, MOS SF36MH *r* = .29	
Karapolat 2008	Poor	MOS SF36PF *r* = .48, MOS SF36RP *r* = .28; MOS SF36BP *r* = .35, MOS SF36GH *r* = .17; MOS SF36Vt *r* = .29; MOS SF36SF *r* = .30; MOS SF36RE *r* = .21, SF36MH *r* = .07. MLHQF PS *r* = -.33, MLHFQ MS *r* = -.04; MLHFQ TS *r* = -.31	

First author (date)	Method. Quality Review	Results	Comments
Nogueira 2010	Poor	MOS SF36: CF *r* = .1; FA *r* = .01; Pain *r* = .1; GHS *r* = .1; V *r* = .2; SA *r* = .2; EA *r* = .4; MH *r* = .2. MLHFQ *r* = -.5	2 6MWTs, 1 hr rest, results of Test-2 used in analysis
O'keefe 1998	Poor	Total CHQ *r* = -.79; CHQ dyspnoea dimension *r* = -.58	*A-priori* hypothesis
Olper 2011	Fair*	TM 6MWT pre-rehabilitation *r* = .72 and post-rehabilitation *r* = .67	
Opasich 2004	Poor	Gender *r* = -.35; ♂>♀ ($p < .0001$); Comorbidity Y<N ($p < .0001$); LVEF in ♂ *r* = .11; LVEF<50% Y<N ($p = .016$), in ♀ LVEF < 50% ($p = .85$)	?Insufficient females with hF
Pulz 2008	Poor	6MWT(m) > ISWT(m) ($p < .0001$), 21 ± 39% more in 6MWT. ↑ impairment → ↑ difference between 6MWT and ISWT	
Riley 1992	Poor	NYHA II > III ($p < .05$) II > IV ($p < .05$) no difference III and IV (p=_)	
Verrill 2003	Poor	F-P QOLI scores for pre-rehabilitation ($.12 \geq r \leq .19$) and post-rehabilitation ($.20 \geq r \leq .24$)	
Westlake 2005	Fair	Depression *r* = -.33	
Zugck 2000	Fair	NYHA *r* =-.58,	
Responsiveness			
Araya-Ramírez 2010	Poor	ΔBody weight *r* = -.125; ΔSF-36 PCS *r* = .224; Δ6MWT distance and initial 6MWT distance *r* = -.465; ΔSF-36 MCS not related; attend ≥25 sessions improved 6MWT 4.1% more than those attending ≤ 24 sessions ($p = .012$); improvement was greatest	

First author (date)	Method. Quality Review	Results	Comments
		in lower tertiles vs middle and highest ($p < .001$)	
Bellet 2011	Poor	Test 1 ES to 6/52 .30, & to 6/12 .50; Test 2 ES to 6/52 .45, & to 6/12 .53	
Bittner 2000	Poor	Walk distance pre- and post- CR $r = .83$; Age (β = -.253), ♂ (β = .213), percent change in training METS (β = .369), change in self-reported physical activity (β = .345)	
Cheetham 2005	Fair*	Criterion hypothesis testing: $\Delta R\dot{V}O_2$ $r = .08$; $\Delta A\dot{V}O_2$ $r = .14$	
Demers 2001	Poor	ES Small -.02 to .16	
Gary 2004	Fair	Group differences between control group and Intervention group for distance walked were significant	
Gremeaux 2009	Poor	SRM 1.1 (strong)	
Gremeaux 2011	Fair*	Criterion hypothesis testing: ΔMETs $r = .59$	
Grenmeaux 2012	Fair	SRM 1.11	
Ingle 2005	Fair	ΔSymptom severity $r = -.75$	
Kavanaugh 1996	Poor	Δ6MWT distance positively correlated with fatigue, dyspnoea, emotional function, mastery	
O'Keefe 1998	Poor	Responsiveness coefficient 1.73, ES Improvement .85 and deterioration 2.13 GRC $r = .78$, ΔCHQ total $r = .70$, dyspnoea CHQ dimension $r = .60$, fatigue CHF dimension $r = .58$, emotion CHQ dimension $r = .47$,	
Olper 2011	Fair	2/52 Endurance treadmill training: ES .60	

First author (date)	Method. Quality Review	Results	Comments
Verrill 2003	Poor	♂ ES .65; ♀ ES .63. Improvements in QOLI Health and Function scores did not associate with improvements in 6MWT performance	
Wright 2001	Poor	Significant improvement in 6MWT in 6/52 cardiac rehabilitation program, not significant change in 6MWT those not attending 6/52	

Note. 6MWT = six minute walk test; ICC = intraclass correlation coefficient; CI = confidence intervals; r = Pearson's product moment correlation coefficient; LOA = limits of agreement; SEM = standard error of measurement; CV = coefficient of variation; METs = metabolic equivalent; $R\dot{V}O_2$ = relative peak oxygen uptake; $A\dot{V}O_2$ = absolute peak oxygen uptake; HR = heart rate; est = estimated; VO2max = maximum oxygen uptake; W = work; VE/VCO2 = minute ventilation per carbon dioxide slope; CPET = cardiopulmonary exercise test; CHF-D = chronic heart failure with optimised pharmacological management; CHF-P = chronic heart failure with pacing; AT = anaerobic threshold; ROC = receiver operating curve; MOS SF 36 = medical outcomes study 36-item short form; PCS = physical component scale; MCS = mental component scale; BMI = body mass index; HADS-A = hospital anxiety depression scale calculated for anxiety; HADS-D = hospital anxiety depression scale calculated for depression; LVEF = left ventricular ejection fraction; NYHA = New York Heart Association; MLHFQ = Minnesota living with heart failure questionnaire; HFFSI = heart failure functional status inventory; C-QLI = cardiac quality of life index; HFSS = health functioning subscale; PSS = psychological and spiritual subscale; SESS = socioeconomic subscale; MAAC-D = multiple affect adjective checklist for depression; MAAC-A = multiple affect adjective checklist for Anxiety; MAAC-H = multiple affect adjective checklist for hostility; QOL = quality of life; Euro-QOL VAS = European quality of life visual analogue scale; KCCQ = Kansas City Cardiomyopathy Questionnaire; PL = Physical Limitation, TS = total score, SL = social limitation; SE = self-efficacy; BDIII = Beck depression inventory; DASI = duke activity status index; MOS SF 36 scales PF = physical function; RF = role physical; BP = bodily pain; GH = general health; V = vitality; RE = role emotional; MLHQF PS = physical score; MS = mental score; TS = total score; ♂ = men; ♀ = women; CHQ = chronic heart failure questionnaire; F-P QOLI = Ferrans and Power's quality of life index; ES = effect size; SRM = standardise response mean; GRC = global rating of change.

Table A10.2

Results of the Measurement Properties of the 10 m ISWT

First author (date)	quality review	Results	Comments
Relative Reliability			
Fowler 2005	Good*	ISWT-1-2 *ICC* .94, 95% CI [.89, 97]; ISWT-2-3 *ICC* .99, 95% CI [.99, .99]	3 tests, 1 week
Green 2001	Fair*	ISWT-1-2 r = .98	2 tests, 1 week
Hanson 2016	Good	n = 10 $ICC_{agreement}$ ISWT-1-2 .990, 95% CI [.961, .998] ISWT-2-3 .997, 95% CI [.990, .999]; n = 62 $ICC_{agreement}$ ISWT-1-2 .990, 95% CI [.928, .997]	3 tests, same day
Jolly 2008	Good	ISWT-1-2 *ICC* .94	2 tests, same day
Lewis 2001	Fair*	ISWT-2-3 r = .90	3 tests, same day
Morales 1999	Fair*	ISWT-2-3 r = .99	3 tests within 2 weeks
Measurement Error			
Fowler 2005	Good*	ISWT-1-2 95% LOA [24, 56]; Repeatability Coefficient 122 m Test 2-3 95% LOA [-2, 5]; Repeatability Coefficient 21 m	3 tests, 1 week
Hanson 2016	Good	Pilot study (n = 10) ISWT-1-2 95% LOA [-6, 36], $SEM_{agreement}$ 23 m; ISWT-2-3 95% LOA [-7, 17], $SEM_{agreement}$ 12 m. Main study (n = 62) ISWT-1-2 95% LOA [12, 22], $SEM_{agreement}$ 17 m (4% of grand mean)	3 tests, same day
Jolly 2008	Good	95% LOA 23, 36	2 tests, same day
Pulz 2008	Good	M_{diff} (SD_{diff}) 8±45 m no significant difference between Walk 1 and 2	2 tests, 30 mins rest
Criterion Validity (concurrent)			

First author (date)	quality review	Results	Comments
		Association with 10 m ISWT distance unless otherwise specified	
Fowler 2005	Excellent*	R$\dot{V}O_2$.*r* = .79 to .87	Criterion: Treadmill
Green 2001	Fair*	R$\dot{V}O_2$.*r* = .83	Criterion: Treadmill
Lewis 2001	Good*	R$\dot{V}O_2$.*r* = .73	Criterion: Treadmill
Mandic Walker 2013	Good	R$\dot{V}O_2$.*r* = .72	Criterion: Bicycle
Morales 1999	Good*	R$\dot{V}O_2$.*r* = .83	Criterion: Bicycle
Morales 2000	Good*	R$\dot{V}O_2$.*r* = .83	Criterion: Bicycle
Pulz 2008	Good	R$\dot{V}O_2$.*r* = .79.	Criterion; Treadmill
Criterion Validity (concurrent)			
Fowler 2005	Excellent*	$\text{R}\dot{\text{V}}\text{O}_2 = (0.03 \times \text{ISWT(m)}) + 7.81$	Criterion: Treadmill
Green 2001	Good*	$\text{R}\dot{\text{V}}\text{O}_2 = \big(0.0283 \times \text{ISWT(m)}\big) + 4.2355$	Criterion: Treadmill
Mandic Walker 2013	Good	$\text{R}\dot{\text{V}}\text{O}_2 = (3.021 \times \text{ISWT speed (km/hr)}) - 0.007$ $\text{R}\dot{\text{V}}\text{O}_2 = (1.564 \times \text{ISWT speed (km/hr)}) - \big(0.219 \times \text{body fat (\%)}\big) + (0.296 \times \text{Chair stands 30 sec } (n)) + 11.399$	Criterion: Bicycle
Morales 1999	Good*	$\text{R}\dot{\text{V}}\text{O}_2 = (0.023 \times \text{ISWT(m)}) + 5.9$ Cut off values to predict R$\dot{V}O_2$ < 14 = 450m (Sens 100%, Spec = 89%	Criterion: Bicycle
Pulz 2008	Good	ISWT area under ROC = .91, best cut off to predict R$\dot{V}O_2$.< 14 was 380 m (Sens 90% Spec 87%).	Criterion; Treadmill

First author (date)	quality review	Results	Comments
Construct Validity			
Mandic Hodge 2013	Poor	Self-reported physical activity at follow-up r = .52	
Pepera 2013	Poor	Predictors of performance: Step length at 66% maximum walking speed (r^2 = .68), leg length (r^2 = .58), height (r^2 = .57). ISWT (m) = (10.7 x height(cm)) – 1316 (SEE = 90 m). Using the equation, the actual distance walk and predicted distance (r = .69)	
Pulz 2008	Poor	Statistically significant differences in distance walked between the 6MWT and the ISWT P<.0001. The more impaired the patient, the higher the difference in 6MWT and ISWT difference. Patients walk on average, 21 ± 39% more in 6MWT than ISWT.	
Responsiveness			
Fowler 2005	Poor	6/52 Cardiac rehabilitation program: ES: .55 (moderate); M_{Diff} 82 m, 95% CI [53, 110]	
Houchen-Wolloff 2015	Fair	6/52 Cardiac rehabilitation program: ES: .38 (small); Mean Diff 65 m, 95% CI [55, 80]; Δ6MWT: significant difference between groups, *post hoc* analysis better and about the same (mean diff 56 m p < .001), between slightly better and about the same (mean diff 41 m p < .05), non-significant difference between better and slightly better (mean diff 15 m)	

Note. ISWT – incremental shuttle walk test; ICC = intraclass correlation coefficient; CI = confidence intervals; r = Pearson's product moment correlation; SEM = standard error of measurement; LOA = limits of agreement; Mdiff = mean difference; SDdiff = standard deviation of the difference; R$\dot{V}O_2$. Relative oxygen uptake; ROC = receiver operating curve; sens = sensitivitiy; spec = specificity; ES = effect size.

Table A10.3

Results of the Measurement Properties of the Alternate Field Exercise Tests

First author (date)	Field Test	quality review	Results	Comments
Relative Reliability				
Cervie 2012	TM 6MWT	Fair	Test-2-3 $ICC_{2,1}$ .927, 95% CI [.869, .960]	3 TM 6MWTs, same day (30 min rest)
De Greef 2005.	Modified GFE	Fair	Test-1-2 Including maximum score (n = 46): *ICC* .98, 95% CI [.97, .99] Test-1-2 Excluding maximum score: (n = 29) *ICC* .92, 95% CI [.91, .93]	2 Modified GFE, 1 week
Gremeaux 2012	200 m FWT	Fair	Test-1-2 *ICC* = .97	2 200 m FWT, same day (30 min rest)
Guimarães 2008	TM 6MWT	Fair*	Test-1-2 *ICC* = .88	2 TM 6MWTs, 2 days
Olper 2011	TM 6MWT	Good*	Test-2-3 Pre-cardiac rehabilitation $ICC_{3,1}$ = .97, 95% CI [.93, .98], Test-2-3 Post-cardiac rehabilitation $ICC_{3,1}$ = .94, 95% CI [.87, .98]	3 TM 6MWT, 2 days
Measurement Error				
Cervie 2012	TM 6MWT	Fair	Test-2-3 SEM 27.87 (SDC = 77 m); MIC 51 - 55	MIC 54 m; 3 TM 6MWTs, same day (30 min rest)
De Greef 2005	Modified GFE	Poor	Bland Altman plots 95% LOA Walk 1-2 -13 to 13	2 Modified GFE, 1 week

First author (date)	Field Test	quality review	Results	Comments
Guimarães 2008	TM 6MWT	Fair*	Bland Altman plots 95% LOA Walk 1-2 -72 to 63 m (estimated)	2 TM 6MWTs, 2 days
Olper 2011	TM 6MWT	Good*	Test-2-3 Pre-rehabilitation SEM = 23 m (SDC = 64 m), Test-2-3 Post-rehabilitation SEM = 18 m (SDC = 50 m).	3 TM 6MWT, 2 days
Concurrent Criterion Validity				
Casillas et al 2015	20 m FWT	Good	Test-1 HR_{max} *r* = -.26	Criterion: Treadmill
De Greef 2005	Modified GFE protocol	Fair*	R$\dot{V}O_2$ *r* = .77	Criterion: Bicycle
Delahaye 1997	Stair climbing	Good*	R$\dot{V}O_2$.*r* = -.66; Watts *r* = -.56	Criterion: Bicycle
Gayda 2003	20 m shuttle walk test	Fair*	R$\dot{V}O_2$ *r* = .91; Maximal ventilation *r* = .61; Maximal heart rate *r* = .80; Maximal speed *r* = .89	Criterion: Treadmill
Gremeaux 2012	200 m FWT	Fair	Pre-CR: W_{peak} *r* = -.42; Post-CR: W_{peak} *r* = -.46	Criterion: Bicycle
Houghton 2002	100 m FWT	Fair*	Treadmill SLET Time *r* = -.64	Criterion: Treadmill
Meyer 2003	TM 6MWT	Fair	$\dot{V}O_2$ expressed as percent predicted VO_{2max} *r* = .58 .	Criterion: Bicycle
Predictive Criterion Validity				
Casillas 2015	200 m FWT	Good	$HR_{max} = 128.35 - (0.59 \times age) + (0.31 \times HR\ 200mFWT)$	Criterion: Treadmill
Chiaranda 2012	1km treadmill walk test	Fair	Beta blocked group (*n* = 66): $R\dot{V}O_2 = (2.79 \times 1km\ TM\ walk\ speed) - (0.49 \times BMI) - (0.14 \times age) + 33.42$. Testing of	Criterion: Treadmill

First author (date)	Field Test	quality review	Results	Comments
			equations in different sample, r = .71 (SEE 2.3ml/kg/min) Not beta blocked group (n = 44): $R\dot{V}O_2 = (4.41 \times 1kmTM\ walk\ speed) - (0.40 \times BMI) - (0.30 \times age) - (0.11 \times 1km\ TM\ HR_{max}) + 46.11$ $R\dot{V}O_2$ = (2.79 x 1 km TM walk speed) – (0.49 x BMI) – (0.14 x age) + 33.42. Testing of equations in different sample, r = .64 (SEE = 3.8 ml/kg/min).	
Construct Validity				
Alosco 2012	2 min step test	Fair	Age r = -.20; BDI-II r =-.18; 3MS r = .28; Executive function r = .22-.29; Language r = .19 to .29	
Delahaye 1997	Stair climbing	Good*	6MWD r = -.82	
Gremeaux 2012	200 m FWT	Fair	6MWT: Pre-CR r = -.57 – Post-CR r = -.93 SF36 Physical component summary Pre: r = -.77 Mental component summary r = -.1 Significant difference in two groups (achieved <90W or achieved >100W in CPET)	
Houghton 2002	100 m FWT	Fair*	Pedometer r = -.10; QOL r = -.44	
Olper 2011	TM 6MWT	Poor	Pre-CR 6MWT (m) r = .72, Post-CR 6MWT (m) r = .67.	
Responsiveness				

First author (date)	Field Test	quality review	Results	Comments
Cervie 2012	TM 6MWT	Poor	2/52 endurance treadmill training: Non-significant difference between distance and global rating of change (Mean diff: “did not improve” 125 (49); “a little better” 100 (17); “somewhat better” 105 (37); “much better” 116 (74).)	3 tests, used the mean score of test-2 and test-3
Gremeaux 2009	200 m FWT	Poor	SRM 1.11 (strong)	
Gremeaux 2012	200 m FWT	Fair	SRM 1.11 (strong)	
Meyer 2003	Tm 6MWT	Poor	ΔTM 6MWT distance (m) and ΔO_2 uptake at VT (ml/kg/min) r = .53	
Olper 2011	TM 6MWT	Poor	2/52 endurance treadmill training: ES: .9, M_{diff} 109 m (91)	

Note. TM = treadmill; ICC = intraclass correlation coefficient; GFE = Groningen Fitness for the Elderly; FWT = fast walk test; SEM = standard error of measurement; LOA = limits of agreement; HR = heart rate; $R\dot{V}O_2$.= relative peak oxygen uptake; W = work; $\dot{V}O_2$ = maximum oxygen uptake; BMI = body mass index; SEE = standard error of estimate; BDI-II = Beck depression inventory; 3MS = modified mini-mental state examination; CR = cardiac rehabilitation; SRM = standardised response mean; O_2 = oxygen; VT = ventilatory threshold; ES = effect size.

Appendix 11: Summary of the Operating Procedures used for the 6MWT

Table A11.1

Operating procedures for the 6MWT

First author (date)	Standard administration	Corridor length (m)	Instructions	Encouragement Y/N Time interval Phrase standardised Y/N	Number of walks, Rest between testing, test used in data analysis (if specified)	Assessment, Pre-, Post- and monitoring during the test
Ades 2003	Cress 1997	?	?	?	Assume 1	?
Adsett 2011	ATS 2002	25	Greatest distance	Encourage Y Time 1 min Phrases Y	2 Rest: vary	Pre: BP Post: BP, time to recovery Monitoring: HR, SpO_2, RPE
Allison 2004	No	?	Limit to Borg RPE 10-11	?	Assume 1	?
Araya-Ramírez	Steele 1996	30	Greatest distance e	?	Assume 1	Pre: Resting ECG, HR, BP Post: BP, RPE Monitoring: HR
Bajraktari 2011	Guyatt 1985	15	Greatest distance	?	1	?
Baldasseroni 2014	Guyatt 1985	30	?	?	1	?
Baptista 2012	ATS 2002	?	?	?	1	?
Beatty 2012	ATS 2002	44	Greatest distance	Encourage: Y Time: 1 min	1	?

First author (date)	Standard administration	Corridor length (m)	Instructions	Encouragement Y/N Time interval Phrase standardised Y/N	Number of walks, Rest between testing, test used in data analysis (if specified)	Assessment, Pre-, Post- and monitoring during the test
Bellet 2011	ATS 2002	30	?	Phrases: Y Encourage Y Time ? Phrases Y	2 Rest: ≥10 mins	Pre: HR, BP, RPE Post: HR, BP, RPE Monitoring: HR, BP, RPE, Rest breaks
Bittner 1993	Guyatt 1985	30.5	Greatest distance	Encourage Y Time 30 sec Phrases Y	1	?
Bittner 2000	Limited detail	? track	Greatest distance	Encourage Y Time 30 sec Phrases Y	Assume 1	Pre: HR, BP, SpO_2 Post: HR, BP, SpO_2 Exertion, angina, & dyspnoea scales Monitoring ?
Cahalin 1996	Guyatt 1985	50	Walk as far as possible	No	Same day practice	Pre: HR, ECG, BP, RPE Post: HR, ECG, BP, RPE Monitoring: HR, ECG, BP, RPE
Carvalho 2011	Limited detail	30	Maximum speed	Yes	2 Rest: 30-45 mins	Pre: HR, BP Post: HR, BP Monitoring: HR
Casillas 2015	Limited detail	50	Greatest distance	No	1	Pre: HR, BP Post: HR, BP Monitoring: HR

First author (date)	Standard administration	Corridor length (m)	Instructions	Encouragement Y/N Time interval Phrase standardised Y/N	Number of walks, Rest between testing, test used in data analysis (if specified)	Assessment, Pre-, Post- and monitoring during the test
Cipriano 2010	AACVPR	30 (circular)	?	?	?	?
Cheetham 2005	Guyatt 1985	?	Greatest distance	No	Familiarisation 1/52 before	?
Chien 2011	ATS 2002	30	Walk at own pace	?	1	?
Corvea-Tindel 2009	Limited detail	?	Greatest distance	No	1	?
Delahaye 1997	Guyatt 1985	?	Walk as quickly as possible	Encourage Y Time ? Phrases ?	1	?
Demers 2001	Limited detail	≥ 20 m	Greatest distance	Encourage Y Time 30 sec Phrases Y	Baseline: 1 Others: 2	?
de Sousa 2008;	ATS 2002	34	Maximum speed	Encourage Y Time 2 & 4 mins Phrases Y	2 and familiarisation 3 tests if difference is ?10% Rest: 20 min Outcome: best distance	Pre: ECG, HR, BP Post: ECG, HR, BP Monitoring: ECG, HR
Doutreleau 2009	ATS 2002	13m	Greatest distance	Provided: Yes Time: 1 min Phrases: standardised	1	Pre: HR SpO2 Post HR SpO2 Monitoring: HR SpO2
Flynn 2009	Limited detail	?	?	?	1	?

First author (date)	Standard administration	Corridor length (m)	Instructions	Encouragement Y/N Time interval Phrase standardised Y/N	Number of walks, Rest between testing, test used in data analysis (if specified)	Assessment, Pre-, Post- and monitoring during the test
Forman 2012	Limited detail	20-25 m	Greatest distance	Provided: Yes Time: Yes Phrases: ?	1	?
Gary 2004	Guyatt 1985	?	Greatest distance	?	1	?
Gayda 2004	Limited detail	20	Greatest distance	Provided: Yes Time: 30 sec Phrases: ?	1	Pre: HR, VO_2, $\dot{V}_E$ Post: HR, VO_2, $\dot{V}_E$ Monitoring: HR, VO_2, $\dot{V}_E$
Green 2001	Guyatt 1985	120	Greatest distance	?	Familiarisation	Pre: ? Post ? Monitoring HR, VO_2
Gremeaux 2009	ATS 2002	50	Greatest distance	Provided: Yes Time: 30 sec Phrases: standard	2 (1st test Familiarisation), Outcome: 2nd test	Pre: BP Post: BP, RPE Monitoring: HR
Gremeaux 2012	ATS 2002	50	Greatest distance	Provided: yes Time: 30 sec Phrases: Y	2 (1st test Familiarisation), Outcome: 2nd test	Pre: ? Post: ? Monitoring: HR
Gremeaux 2011	ATS 2002	50	Greatest distance	Provided Y Time 30 sec Phrases Y	Familiarisation	Pre: BP Post: BP Monitoring: ?
Guazzi 2009	No detail	?	Greatest distance	?	2 (separated days) First was familiarisation Outcome: 2nd test	?

First author (date)	Standard administration	Corridor length (m)	Instructions	Encouragement Y/N Time interval Phrase standardised Y/N	Number of walks, Rest between testing, test used in data analysis (if specified)	Assessment, Pre-, Post- and monitoring during the test
Guyatt 1985	Guyatt 1985	33	Greatest distance	Encouraged group Provided Y Time 30 sec Phrases Y Unencouraged group N	6 (2 week interval) Outcome: mean score	?
Hamilton 2000	No detail	98.62 Rectangle	Greatest distance	No	3 Outcome: best test score	?
Hanson 2012	Guyatt 1985	20	Greatest distance	Provided: Yes Time: Yes Phrases: Yes	3 Outcome: comparisons with 1st, 2nd and 3rd walk	Pre: BP Post: BP Monitoring: HR, SpO2
Ingle 2006	Roul 1998	15	Greatest distance	?	1	?
Ingle 2005	Guyatt 1985	20-25	Greatest distance	Provided: Yes Time: 2, 4 mins Phrases: yes	1	?
Jehn 2009	?	40	Greatest distance; but pace complete full 6 mins	?	1 (but previous familiarisation)	Pre: VO_2 Post: VO_2 Monitoring: VO_2
Jelinek 2013	?	18	Walk up and down track as many times	No	2 Outcome: best score	?
Juenger 2002	No detail	?	?	?	1	?

First author (date)	Standard administration	Corridor length (m)	Instructions	Encouragement Y/N Time interval Phrase standardised Y/N	Number of walks, Rest between testing, test used in data analysis (if specified)	Assessment, Pre-, Post- and monitoring during the test
Karapolat	Guyatt 1985	20 m	At own pace, rest permitted			
Kervio 2004a	Guyatt 1985	18	Greatest distance	Yes/30sec/Yes	Unclear which score was used	Pre: HR VO_2 Post: HR VO_2 Monitoring: HR, VO_2
Kervio 2004b	Guyatt 1985	18	Regular pace as far as they could	Yes/30sec/Yes	1	Pre ? Post ? Monitoring: HR, VO_2, VCO_2, V_E, V_T, RR
Kristjánsdóttir 2004	Guyatt 1985	35	Greatest distance, to exhaustion	No	4 Outcome: best test score	Pre: HR, BP Post: HR, BP, RPE
Langenfeld 1990	Guyatt 1985	20	Greatest distance	No	1	Pre: ? Post: ? Monitoring: HR
Lipkin 1986;	McGavin 1976	20	Greatest distance, to exhaustion	Yes, as needed	3 Outcome: 3rd test	
Lucas 1999	Lipkin 1986	?	Greatest distance	?	1	?
Maldonado-Martin 2006	Guyatt	? track	Greatest distance, own pace	No	1	?

First author (date)	Standard administration	Corridor length (m)	Instructions	Encouragement Y/N Time interval Phrase standardised Y/N	Number of walks, Rest between testing, test used in data analysis (if specified)	Assessment, Pre-, Post- and monitoring during the test
Mandic Walker 2013	?	50 Rectangle	Greatest distance Allowed to jog	Yes Time: ? Standardised: Yes	2 Outcome: best test score	Pre: ? Post: ? Monitoring: HR, RPE
Morales 1999	Lipkin 1986	20	Greatest distance own pace, no further	Yes, as necessary with standard phrases	2 Outcome: ?	?
Morales 2000	?	?	Greatest distance, own pace	?	2 Outcome: 2nd test	?
Nogueira 2006	AACVPR 2013	33	Greatest distance, own pace	30 sec standardised	3 Outcome: 3rd test	Pre: ? Post: ? Monitoring: HR ECG
O'Keefe 1998	Guyatt 1985	25	Greatest distance, own pace	Test, 30 sec standardised	1	?
Olper 2011	ATS 2002	18	Greatest distance, own pace	Yes, 1 min, standardised	3, Outcome: Used best test	Pre: HR, BP Post: HR BP RPE Monitoring: ECG
Opasich 2004	ATS 2002	35	Max pace	No	1	?
Opasich 1998	Limited detail	34	Greatest distance, to exhaustion	No	Familiarisation, then 2 tests Outcome: mean score	?

First author (date)	Standard administration	Corridor length (m)	Instructions	Encouragement Y/N Time interval Phrase standardised Y/N	Number of walks, Rest between testing, test used in data analysis (if specified)	Assessment, Pre-, Post- and monitoring during the test
Opasich 2001	Limited detail	34	Greatest distance, until Exhaustion,	No	Familiarisation, then 2 tests. Outcome: Mean score	?
Peeters 1996	Limited detail	60 m Circle	Walk as quickly as possible to cover longest distance Accompanied by observer	Yes, 30 seconds	1	Pre: BP, HR Post: BP, HR
Pinna 2000	Limited detail	34	Greatest distance, until Exhaustion	No	2 (> 30mins rest break) and additional familiarisation Outcome: mean score	?
Pulz 2008	ATS 2002	30	Greatest distance Own pace	Yes, 1 min, standardised	2 Outcome: ?	Pre: ? Post: BP, RPE Monitoring: HR, ECG
Rejeski 2002	No detail	?	Greatest distance	No	1	?
Riley 1992	Limited detail	25	Greatest distance	?	3 Outcome: 3rd test	Pre ? Post: ? Monitoring VO2
Verrill 2003	Limited detail	?	Greatest distance	Provided: Yes Time: Yes	Assumable 1	Pre: HR BP Post: HR BP RPE

First author (date)	Standard administration	Corridor length (m)	Instructions	Encouragement Y/N Time interval Phrase standardised Y/N	Number of walks, Rest between testing, test used in data analysis (if specified)	Assessment, Pre-, Post- and monitoring during the test
				Phrases: Yes		Monitoring: ?
Westlake 2005	Limited detail	?	Research assistant walked behind	?	1	?
Wright 2001	Guyatt 1985	20	Greatest distance	No	Familiarisation (no detail)	Pre: HR, BP Post HR, BP, RPE
Zugck 2000	Guyatt 1985	132	Greatest distance	No	1 (but 3 for *n* = 10)	?

Note. ATS = American Thoracic Society; BP = blood pressure; HR = heart rate; SpO2 = oxygen saturation measured via pulse oximetry; RPE = rate of perceived exertion; ECG = electrocardiogram; AACPR = American Association of Cardiovascular and Pulmonary Rehabilitation; VO_2, oxygen uptake; $\dot{V}_E$ = ventilatory threshold; VCO_2, = volume of carbon dioxide expired; V_E, = pulmonary ventilation; V_T = tidal ventilation; RR = respiratory rate.

References

Acute Coronary Syndrome Guidelines Working Group. (2006). Guidelines for the management of acute coronary syndromes 2006. *Medical Journal of Australia, 184*(8), S9-S29.

Ades, P., Savage, M., Cress, M., Brochu, N., & Poehlman, E. (2003). Resistance training on physical performance in disabled older female cardiac patients. *Medicine and Science in Sports and Exercise, 35*, 1265-1270. doi:10.1249/01.MSS.0000079044.21828.0E ·

Ades, P. A., Maloney, A., Savage, P. D., & Carhart, R. L. J. (1999). Determinants of physical functioning in coronary patients. *Archives of Internal Medicine, 159*(19), 2357-2360. doi:10.1001/archinte.159.19.2357

Ades, P. A., Savage, P. D., Brawner, C. A., Lyon, C. E., Ehrman, J. K., Bunn, J. Y., & Keteyian, S. J. (2006). Aerobic capacity in patients entering cardiac rehabilitation. *Circulation, 113*, 2706-2712. doi:10.1161/CIRCULATIONAHA.105.606624

Ades, P. A., Savage, P. D., Tischler, M. D., Poehlman, E. T., Dee, J., & Niggel, J. (2002). Determinants of disability in older coronary patients. *American Heart Journal, 143*, 151-156.

Ainsworth, B. E., Haskell, W. L., Whitt, M. C., Irwin, M. L., Swartz, K. H., Emplaincourt, P. O., . . . Leon, A. S. (2000). Compendium of physical activities: An update of activity codes and MET intensities. *Medicine and Science in Sports and Exercise, 32*(9; SUPP/1) S498-S504.

Allison, M. J., & Keller, C. (2004). Self-efficacy intervention effect on physical activity in older adults. *Western Journal of Nursing Research, 26*(1), 31-46. doi:10.1177/0193945903259350

Alosco, M. L., Spitznagel, M. B., Raz, N., Cohen, R., Sweet, L. H., Colbert, L. H., . . . Gunstad, J. (2012). The 2-minute step test is independently associated with cognitive function in older adults with heart failure. *Aging Clinical and Experimental Research, 24*(5), 468-474. doi:10.3275/8186

Alston, L., Allender, S., Peterson, K., Jacob, J., & Nichols, M. (2017). Rural inequalities in the Australian burden of ischaemic heart disease: A systematic review. *Heart, Lung and Circulation, 26*(2), 122-133. doi:10.1016/j.hlc.2016.06.1213

Altman, D. G., & Bland, J. M. (1999). How to randomise. *British Medical Journal, 319*(7211), 703-704.

Altman, D. G., & Gardner, M. J. (2000). Regression and correlation. In D. Altman, D. Machin, T. Bryant, & M. Gardner (Eds.), *Statistics with confidence* (2nd ed., pp. 73-92). London: British Medical Journal Books.

American Association of Cardiovascular and Pulmonary Rehabilitation. (2013). *Guidelines for cardiac rehabilitation and secondary prevention programs* (5th ed.). Champaign, IL: Human Kinetics.

American Association of Cardiovascular and Pulmonary Rehabilitation. (2016). Performance measure for improvement in functional capacity at completion of cardiac rehabilitation. Retrieved from AACVPR Program Certification website: http://www.aacvpr.org/Portals/0/Program%20Certification/Performance%20M

easures/Measures/CR%20Functional%20Capacity%20Performance%20Measure_10.24.16.pdf

American College of Sports Medicine. (2010a). *ACSM's guidelines for exercise testing and prescription* (8th ed.). Baltimore, MD: Williams & Wilkins.

American College of Sports Medicine. (2010b). *ACSM's health-related physical fitness assessment manual* (3rd ed.). Baltimore, MD: Lippincott Williams & Wilkins.

American Thoracic Society. (2002). ATS Statement: Guidelines for the six- minute walk test. *American Journal of Respiratory and Critical Care Medicine, 166*(1), 111-117. doi:10.1164/ajrccm.166.1.at1102

American Thoracic Society, & American College of Chest Physicians. (2003). ATS/ACCP Statement on cardiopulmonary exercise testing. *American Journal of Respiratory and Critical Care Medicine, 167*(2), 211-277. doi:10.1164/rccm.167.2.211

Amireault, S., & Godin, G. (2014). Validation of using fitness center attendance electronic records to assess the frequency of moderate/vigorous leisure-time physical activity among adults. *Measurement in Physical Education and Exercise Science, 18*(1), 13-30. doi:10.1080/1091367X.2013.831765

Anderson, K. N., Anderson, L. E., & Glanze, W. D. (1997). *Mosby's medical, nursing, & allied health dictionary*. Philadelphia: Sauders.

Anderson, L., Olridge, N., Thompson, D. R., Zwisler, A.-D., Rees, K., Martin, N., & Taylor, R. S. (2016). Exercise-based cardiac rehabilitation for coronary heart disease: Cochrane systematic review and meta-analysis. *Journal of American College of Cardiology, 67*(1), 1-12. doi:10.1016/j.jacc.2015.10.044

Anokye, N. K., Trueman, P., Green, C., Pavey, T. G., & Taylor, R. S. (2012). Physical activity and health related quality of life. *BMC Public Health, 12*(1), 324. doi:10.1186/1471-2458-12-624

Araya-Ramírez , F., Briggs, K. K., Bishop, S. R., Miller, C. E., Moncada-Jiménez, J., & Grandjean, P. W. (2010). Who is likely to benefit from phase II cardiac rehabilitation? *Journal of Cardiopulmonary Rehabilitation, 30*, 93-100. doi:10.1097/HCR.0b013e3181d0c17f

Arnold, H. J., Sewell, L., & Singh, S. (2007). A comparison of once- versus twice-weekly supervised phase III cardiac rehabilitation. *British Journal of Cardiology, 14*(1), 45-48.

Arnott, A. S. (1997). Assessment of functional capacity in cardiac rehabilitation: Reproducibility of a 10-m shuttle walk test. *Coronary Health Care, 1*(1), 30-36. doi:10.1016/S1362-3265(97)80037-5

Audelin, M. C., Savage, P. D., & Ades, P. A. (2008). Changing clinical profile of patients entering cardiac rehabilitation/secondary prevention programs: 1996 to 2006. *Journal of Cardiopulmonary Rehabilitation and Prevention, 28*(5), 299-306. doi:10.1097/01.HCR.0000336139.48698.26

Australian Bureau of Statistics. (1995). *National Health Survey: SF36 population norms, Australia*. (Australian Bureau of Statistics Catalogue No. 4399.0). Canberra, Australia. Retrieved from Australian Bureau of Statistics website: http://www.abs.gov.au/AUSSTATS/abs@.nsf/mf/4399.0

Australian Institute for Health and Welfare. (2017). How many Australians have cardiovascular disease? Retrieved from http://aihw.gov.au/cardiovascular-disease/prevalence/

Babu, A. D., Lopez-Jimenez, F., Thomas, R. J., Isaranuwatchai, W., Herdy, A. H., Hoch, J. S., & Grace, S. L. (2016). Advocacy for outpatient cardiac rehabilitation globally. *BMC Health Services Research, 16*(1), 471. doi:10.1186/s12913-016-1658-1

Bajraktari, G., Elezi, S., Berisha, V., Lindqvist, P., Rexhepaj, N., & Henein, M. Y. (2011). Left ventricular asynchrony and raised filling pressure predict limited exercise performance assessed by 6 minute walk test. *International Journal of Cardiology, 146*(3), 385-389. doi:10.1016/j.ijcard.2009.07.018

Balady, G. J., Ades, P. A., Bittner, V. A., Franklin, B. A., Gordon, N. F., Thomas, R. J., . . . Yancy, C. W. (2011). Referral, enrolment, and delivery of cardiac rehabilitation/secondary prevention programs at clinical centers and beyond: a presidential advisory from the American Heart Association. *Circulation, 124*(25), 2951-2960. doi:10.1161/CIR.0b013e31823b21e2

Balady, G. J., Arena, R., Sietsema, K., Myers, J., Coke, L., Fletcher, G. F., . . . Milani, R. V. (2010). Clinician's guide to cardiopulmonary exercise testing in adults: A scientific statement from the American Heart Association. *Circulation, 122*(2), 191-225. doi:10.1161/CIR.0b013e3181e52e69

Balady, G. J., Williams, M. A., Ades, P. A., Bittner, V., Comoss, P., Foody, J. M., . . . Southard, D. (2007). Core components of cardiac rehabilitation/secondary prevention programs: 2007 update A scientific statement from the American Heart Association Exercise, Cardiac Rehabilitation, and Prevention Committee, the Council on Clinical Cardiology; the Councils on Cardiovascular Nursing, Epidemiology and Prevention, and Nutrition, Physical Activity, and Metabolism; and the American Association of

Cardiovascular and Pulmonary Rehabilitation. *Circulation, 115*(20), 2675-2682. doi:10.1161/CIRCULATIONAHA.106.180945

Baldasseroni, S., Pratesi, A., Orso, F., Foschini, A., Marella, A. G., Bartoli, N., . . . Tarantini, F. (2014). Age-related impact of depressive symptoms on functional capacity measured with 6-minute walking test in coronary artery disease. *European Journal of Preventive Cardiology, 21*(5), 647-654. doi:10.1177/2047487313494581

Bandura, A. (2006). Guide for constructing self-efficacy scales. In F. Pajares & T. Urdan (Eds.), *Self-efficacy beliefs of adolescents* (pp. 307-337). Greenwich, CT: Information Age Publishing.

Bansal, V., Hill, K., Dolmage, T. E., Brooks, D., Woon, L. J., & Goldstein, R. S. (2008). Modifying track layout from straight to circular has a modest effect on the 6-min walk distance. *Chest, 133*(5), 1155-1160. doi:10.1378/chest.07-2823

Baptista, V. C., Palhares, L. C., de Oliveira, P. P., Silveira Filho, L. M., Vilarinho, K. A., Severino, E. S., . . . Petrucci, O. (2012). Six-minute walk test as a tool for assessing the quality of life in patients undergoing coronary artery bypass grafting surgery. *Brazilian Journal of Cardiovascular Surgery, 27*(2), 231-239. doi:10.5935/1678-9741.20120039

Bartels, B., De Groot, J. F., & Terwee, C. B. (2013). The six-minute walk test in chronic pediatric conditions: A systematic review of measurement properties. *Physical Therapy, 93*(4), 529-541.

Beatty, A., Schiller, N., & Whooley, M. (2012). Six-minute walk test as a prognostic tool in stable coronary heart disease. *Archives of Internal Medicine, 172*(14), 1096-1102. doi: :10.1001/archinternmed.2012.2198.

Beekman, E., Mesters, I., Hendriks, E. J. M., Klassen, P. M., Gosselink, R., van Schayck, O. C. P., & de Bie, R. A. (2013). Course length of 30 metres versus 10 metres has a significant influence on six-minute walk distance in patients with COPD: an experimental crossover study. *Journal of Physiotherapy, 59*(3), 169-176. Doi: 10.1016/S1836-9553(13)70181-4

Bellet, R. N., Francis, R. L., Jacob, J. S., Healy, K. M., Bartlett, H. J., Adams, L. S., & Morris, N. R. (2011). Repeated six-minute walk tests for outcome measurement and exercise prescription in outpatient cardiac rehabilitation: A longitudinal study. *Archives of Physical Medicine and Rehabilitation, 92*(9), 1388-1394. doi:10.1016/j.apmr.2011.04.014

Bendigo Health. (2013a). About Bendigo Health. Retrieved from http://www.bendigohealth.org.au/About_Us.asp?PageID=10

Bendigo Health. (2013b). Our Services (A to Z). Retrieved from http://www.bendigohealth.org.au/Patients_Families_Services.asp?index=0&nav_letter=C

Benhorin, J., Pinsker, G., Moriel, M., Gavish, A., Tzivoni, D., & Stern, S. (1993). Ischemic threshold during two exercise testing protocols and during ambulatory electrocardiographic monitoring. *Journal of the American College of Cardiology, 22*(3), 671-677. doi:10.1016/0735-1097(93)90175-Z

Beniamini, Y., Rubensein, J. J., Zaichkowsky, L. D., & Crim, M. C. (1997). Effects of high intensity strength training on quality-of-life parameters in cardiac rehabilitation patients. *The American Journal of Cardiology, 80*(7), 841-846. doi:10.1016/S0002-9149(97)00533-X

Bergner, M., & Rothman, M. L. (1987). Health status measures: An overview and guide for selection. *Annual Reviews of Public Health, 8*, 191-210.

Beyler, N., Nusser, S., Fuller, W., & Gregory, W. (2008). *Relating self-report and accelerometer physical activity with application to NHANES 2003–2004.* Paper presented at the Proceedings of the Survey Research Methods Section.

Bittner, V., Sanderson, B., Breland, J., Adams, C., & Schumann, C. (2000). Assessing functional capacity as an outcome in cardiac rehabilitation: Role of the 6-minute walk test. *Clinical Exercise Physiology, 2*(1), 19-26.

Bittner, V., Weiner, D. H., Yusuf, S., Rogers, W. J., McIntyre, K. M., Bangdiwala, S. I., . . . Bourassa, M. G. (1993). Prediction of mortality and morbidity with a 6-minute walk test in patients with left ventricular dysfunction. *Journal of the American Medical Association, 270*(14), 1702-1707.

Bize, R., Johnson, J. A., & Plotnikoff, R. C. (2007). Physical activity level and health-related quality of life in the general adult population: A systematic review. *Preventive Medicine, 45*(6), 401-415. doi:10.1016/j.ypmed.2007.07.017

Bland, J. M., & Altman, D. G. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *The Lancet, 327*(8476), 307-310. doi:10.1016/S0140-6736(86)90837-8

Borg, G. A. V. (1990). Psychophysical scaling with applications in physical work and the perception of exertion. *Scandinavian Journal of Work, Environment and Health, 16*(suppl 1), 55-58. doi:10.5271/sjweh.1815

Borg, G. A. V., & Noble, B. J. (1973). Perceived exertion. *Exercise and Sports Sciences Reviews, 2*(1), 131-154. doi:10.1249/00003677-19740020-00006

Bowen, D. J., Kreuter, M., Spring, B., Cofta-Woerpel, L., Linnan, L., Weiner, D., . . . Fernandez, M. (2009). How we design feasibility studies. *American Journal of Preventative Medicine, 36*(5), 452-457. doi:10.1016/j.amepre.2009.02.002

Briffa, T., Allan, R., Maiorana, A., Oldenburg, B., Sammel, N., Stubbs, A., . . . On behalf of the Executive Working Group and National Forum Participants. (2006). *National Heart Foundation of Australia physical activity recommendations for people with cardiovascular disease*. Sydney, Australia: National Heart Foundation of Australia. Retrieved from http://www.heartfoundation.org.au/SiteCollectionDocuments/PAR4CVD.pdf.

British Association for Cardiovascular Prevention and Rehabilitation. (2012). *The BACPR standards and core components for cardiovascular disease prevention and rehabilitation* (2nd ed.). London: British Cardiovascular Society.

Brodie, D., Bethell, B., & Breen, S. (2006). Cardiac rehabilitation in England: a detailed national survey. *European Journal of Preventative Cardiology, 13*(1), 122-128. doi:10.1097/01.hjr.0000192741.04103.d3

Brown, K. (2003). A review to examine the use of SF-36 in cardiac rehabilitation. *British Journal of Nursing, 12*(15), 904-909. doi:10.12968/bjon.2003.12.15.11422

Brown, T. M., Hernandez, A. F., Bittner, V., Cannon, C. P., Ellrodt, G., Liang, L., & Peterson, E. D. (2009). Predictors of cardiac rehabilitation referral in coronary artery disease patients: Findings from the American Heart Association's Get with the Guidelines Program. *Journal of American College of Cardiology, 54*(6), 515-521. doi:10.1016/j.jacc.2009.02.080

Brown, W. J., Bauman, A. E., Bull, F. C., & Burton, N. W. (2012). *Development of evidence-based physical activity recommendations for adults (18-64 years)*. Retrieved from http://www.health.gov.au/internet/main/publishing.nsf/Content/health-pubhlth-strateg-phys-act-guidelines/$File/DEB-PAR-Adults-18-64years.pdf

Brubaker, C., Witta, E. L., & Angelopoulos, T. J. (2003). Comparison of changes in exercise tolerance and quality of life between congestive heart failure and coronary artery bypass graft patients following a hospital-based cardiac rehabilitation program. *Journal of Exercise Physiology, 6*(1), 18-23.

Bruce, R. A., Blackmon, J. R., Jones, J. W., & Strait, G. (1963). Exercise testing in adult normal subjects and cardiac patients. *Pediatrics, 32*, 742-756.

Bruce, R. A., & Hornsten, T. R. (1969). Exercise stress testing in evaluation of patients with ischemic heart disease. *Progress in Cardiovascular Diseases, 11*(5), 371-390.

Bruce, R. A., Kusumi, F., & Hosmer, D. (1973). Maximal oxygen intake and nomographic assessment of functional aerobic impairment in cardiovascular disease. *American Heart Journal, 85*(4), 546-562.

Buckley, J. P., Furze, G., Doherty, P., Speck, L., Connolly, S., Hinton, S., & Jones, J. L. (2013). BACPR scientific statement: British standards and core components for cardiovascular disease prevention and rehabilitation. *Heart, 99*(15), 1069-1071. doi:10.1136/heartjnl-2012-303460

Butland, R. J. A., Pang, J., Gross, E. R., Woodcock, A. A., & Geddes, D. M. (1982). Two-, six-, and 12-minute walking tests in respiratory disease. *British Medical Journal (Clinical Research Education), 284*(6329), 1607-1608. doi:10.1136/bmj.284.6329.1607

Cahalin, L. P., Mathier, M. A., Semigran, M. J., Dec, G. W., & DiSalve, T. G. (1996). The six-minute walk test predicts peak oxygen uptake and survival in patients with advanced heart failure. *Chest, 110*(2), 325-332. doi:10.1378/chest.110.2.325

Candido, E., Richards, J. A., Oh, P., Suskin, N., Arthur, H. M., Fair, T., & Alter, D. A. (2011). The relationship between need and capacity for multidisciplinary cardiovascular risk-reduction programs in Ontario. *Canadian Journal of Cardiology, 27*(2), 200-207. doi:10.1016/j.cjca.2011.01.008

Carvalho, E. E. V. D., Costa, D. C., Crescencio, J. C., de Santi, G. L., Papa, V., Marques, F., . . . Junior, L. G. (2011). Heart failure: Comparison between six-minute walk test and cardiopulmonary test. *Arquivos Brasileiros de Cardiologia, 97*(1), 59-64. doi:10.1590/S0066-782X2011005000056

Casanova, C., Celli, B. R., Casas, B. A., Cote, C., de Torres, J. P., Jardim, J., . . . Aguirre-Jaime, A. (2011). The 6-min walk distance in healthy subjects: reference standards from seven countries. *European Respiratory Journal, 37*(1), 150-156. doi:10.1183/09031936.00194909

Casillas, J. M., Joussain, C., Gremeaux, V., Hannequin, A., Rapin, A., Laurent, Y., & Benaim, C. (2015). A study of the 200-metre fast walk test as a possible new assessment tool to predict maximal heart rate and define target heart rate for exercise training of coronary heart disease patients. *Clinical Rehabilitation, 29*(2), 175-183. doi:10.1177/0269215514540922

Caspersen, C. T., Powell, K. E., & Christenson, G. M. (1985). Physical activity, exercise, and physical fitness: Definitions and distinctions for health-related research. *Public Health Reports, 100*(2), 126-131.

Cervie, P., Olper, L., De Santi, F., & Pierini, S. M., C. (2012). Reliability and minimal important difference of the treadmill six minute walking test in people following cardiac surgery. *Italian Journal of Physiotherapy, 2*(4), 142-150.

Chatterjee, S., Sengupta, S., Nag, M., Kumar, P. G., S., & Rudra, A. (2013). Cardiopulmonary exercise testing: A review of techniques and applications.

Anesthesia and Clinical Research, 4(7), 340-346. doi:10.4172/2155-6148.1000340

Cheetham, C., Taylor, R., Burke, V., O'Driscoll, G., & Green, D. J. (2005). The 6-minute walk test does not reliably detect changes in functional capacity of patients awaiting cardiac transplantation. *Journal of Heart Lung Transplant, 24*(7), 848-853. doi:10.1016/j.healun.2004.05.012

Cheng, T. Y. L., & Boey, K. W. (2002). The effectiveness of a cardiac rehabilitation on self-efficacy and exercise tolerance. *Clinical Nursing Research, 11*(1), 10-21. doi:10.1177/105477380201100102

Chew, D. P. B., Allan, R. M., Aroney, C. N., & Sheerin, N. J. (2005). National data elements for the clinical management of acute coronary syndromes. *Medical Journal of Australia, 182*(9), 1-16.

Chiaranda, G., Myers, J., Mazzoni, G., Terranova, F., Bernardi, E., Grossi, G., . . . Grazzi, G. (2012). Peak oxygen uptake prediction from a moderate, perceptually regulated, 1-km treadmill walk in male cardiac patients. *Journal of Cardiopulmonary Rehabilitation and Prevention, 32*(5), 262-269. doi:10.1097/HCR.0b013e3182663507

Chien, C. L., Lee, C. M., Wu, Y. W., & Wu, Y. T. (2011). Home-based exercise improves the quality of life and physical function but not the psychological status of people with chronic heart failure: a randomised trial. *Journal of Physiotherapy, 57*(3), 157-163. doi:10.1016/S1836-9553(11)70036-4

Cipriano G. Jr., Bernardelli, G. F., Arena, R., Oliveira, L. V., Valdez, F., & Branco, J. N. (2010). Cardio-respiratory responses of the 6-minute walk test in patients with refractory heart failure during the preoperative period for heart transplant surgery. *Monaldi Archives Chest Disease, 74*(2), 64-69.

Clark, A. M., King-Shier, K. M., Duncan, A., Spaling, M., Stone, J. A., Jaglal, S., & Angus, J. (2012). Factors influencing referral to cardiac rehabilitation and secondary prevention programs: A systematic review. *European Journal of Preventative Cardiology, 20*(4), 692-700. doi:10.1177/2047487312447846

Clark, D. M., McKenzie, D. P., Marshall, R. J., & Smith, G. C. (1994). The construction of a brief case-finding instrument for depression in the physically ill. *Integrative Psychiatry, 10*(3), 117-123.

Clark, L. A., & Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological Assessment, 7*(3), 309-319.

Cohen, J. (1977). *Statistical power analysis for behavioural sciences*. New York: Academic Press.

Cohen, R. A., Moser, D. J., Clark, M. M., Aloia, M. S., Cargill, B. R., Stefanik, S., . . . Forman, D. E. (1999). Neurocognitive functioning and improvement in quality of life following participation in cardiac rehabilitation. *American Journal of Cardiology, 83*, 1374-1378. doi:10.1016/S0002-9149(99)00103-4

Corvera-Tindel, T., Doering, L. V., Roper, J., & Dracup, K. (2009). Emotional functioning drives quality of life in men with heart failure. *Progress in Cardiovascular Nursing, 24*(1), 2-11. doi:10.1111/j.1751-7117.2009.00026.x

Costa-Santos, C., Bernardes, J., Ayres-de-Campos, D., Costa, A., & Costa, C. (2011). The limits of agreement and the intraclass correlation coefficient may be inconsistent in the interpretation of agreement. *Journal of Clinical Epidemiology, 64*, 264-269. doi: 10.1016/j.jclinepi.2009.11.010

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin, 52*(4), 281-302.

CVC Victoria. (2014a). About Us. Retrieved from http://www.cvcvictoria.com.au/Home/AboutUs.aspx

CVC Victoria. (2014b). Stress Echoes. Retrieved from http://www.cvcvictoria.com.au/DiagnosticTests/StressEcho.aspx

Dafoe, W. A., Huston, P., Wong, C., & Stokes, H. (2003). Primary and secondary prevention of cardiac disease: bridging the two solitudes. *The Canadian Journal of Cardiology, 19*(6), 709-713.

Dallal, G. E. (2008). Randomization.com. Retrieved from http://www.randomization.com/

de Greef, M. H., Sprenger, S. R., Elzenga, C. T., Popkema, D. Y., Bennekers, J. H., Niemeijer, M. G., . . . Mook, G. A. (2005). Reliability and validity of a twelve-minute walking test for coronary heart disease patients. *Perceptual and Motor Skills, 100*(2), 567-575. doi:10.2466/pms.100.2.567-575

de Sousa, L. A. P., Britto, R. R., Ribeiro, A. L., Baracho, S. M., Barros, V. C. V., Carvalho, V. T., & Parreira, V. F. (2008). Six-minute walk test in patients with permanent cardiac pacemakers. *Journal of Cardiopulmonary Rehabilitation and Prevention, 28*(4), 253-257. doi:10.1097/01.HCR.0000327183.51992.76

de Vet, H. C. W., Ostelo, R. W. J. G., Terwee, C. B., van der Roer, N., Knol, D. L., Beckerman, H., . . . Bouter, L. M. (2007). Minimally important change determined by a visual method integrating an anchor-based and a distribution-based approach. *Quality Life Research, 16*(1), 131-142. doi:10.1007/s11136-006-9109-9

de Vet, H. C. W., Terwee, C. B., Knol, D. L., & Bouter, L. M. (2006). When to use agreement versus reliability measures. *Journal of Clinical Epidemiology, 59*(10), 1033-1039. doi:10.1016/j.jclinepi.2005.10.015

de Vet, H. C. W., Terwee, C. B., Mokkink, L. B., & Knol, D. L. (2011). *Measurement in Medicine*. Cambridge: Cambridge University Press.

Delahaye, N., Cohen-Solal, A., Faraggi, M., Czitrom, D., Foult, J. M., Daou, D., . . . Le Guludec, D. (1997). Comparison of left ventricular responses to the six minute walk test, stair climbing, and maximal upright bicycle exercise in patients with congestive heart failure due to idiopathic dilated cardiomyopathy. *The American Journal of Cardiology, 80*(1), 65-79. doi: 10.1016/S0002-9149(97)00285-3

Demers, C., McKelvie, R. S., Negassa, A., & Yusuf, S. (2001). Reliability, validity, and responsiveness of the six-minute walk test in patients with heart failure. *American Heart Journal, 142*(4), 698-703. doi:10.1067/mhj.2001.118468

Dempster, M., & Donnelly, M. (2000). Measuring the health related quality of life of people with ischaemic heart disease. *Heart, 83*, 641-644. doi:10.1136/heart.83.6.641

Department of Health. (2013). Greater Bendigo (C). Retrieved from http://docs.health.vic.gov.au/docs/doc/4884D467E3A9603FCA25786600172E2A/$FILE/Greater_Bendigo_2012-v02.pdf

Department of Health and Human Services. (2014). Loddon Mallee region. Retrieved from http://www.health.vic.gov.au/regions/loddonmallee/

Dobson, F., Hinman, R. S., Hall, M., Terwee, C. B., Roos, M., & Bennell, K. L. (2012). Measurement properties of performance-based measures to assess physical function in high and knee osteoarthritis: a systematic review. *Osteoarthritis and Cartilages, 20*(12), 1548-1562. doi:10.1016/j.joca.2012.08.015

Doutreleau, S., Di Marco, P., Talha, S., Charloux, A., Piquard, F., & Geny, B. (2009). Can the six-minute walk test predict peak oxygen uptake in men with heart transplant? *Archives of Physical Medicine and Rehabilitation, 90*(1), 51-57. doi:10.1016/j.apmr.2008.07.010

Dugmore, L. D., Tipson, R. J., Phillips, M. H., Flint, E. J., Stentiford, N. H., Bone, M. F., & Littler, W. A. (1999). Changes in cardiorespiratory fitness, psychological wellbeing, quality of life, and vocational status following a 12 month cardiac rehabilitation programme. *Heart, 81*, 359-366. doi:10.1136/hrt.81.4.359

Dybkaer, R. (2011). Definitions of 'measurement'. *Accreditation and Quality Assurance, 16*(10), 479. doi:10.1007/s00769-011-0808-8

Everett, B., Salamonson, Y., & Davidson, P. M. (2009). Bandura's exercise self-efficacy scale: Validation in an Australian cardiac rehabilitation setting. *International Journal of Nursing Studies, 46*, 824-829. doi:10.1016/j.ijnurstu.2009.01.016

Fitzpatrick, R., Davey, C., Buxton, M. J., & Jones, D. R. (1998). Evaluating patient-based outcome measures for use in clinical trials. *Health Technology Assessment, 21*(4). Retrieved from https://static1.squarespace.com/static/58a757f81b631b594055bed4/t/58cc3cf3c534a5ffdbcf9f45/1489779955874/Evaluating+patient-based+outcome+measures+for+use+in+clinical+trials.pdf

Fletcher G. F., Balady G., Froelicher V. F., Hartley L. H., Haskell W. L., Pollock M. L., & M., W. I. (1995). Exercise standards: A statement for healthcare professionals from the American Heart Association. *Circulation, 91*(2), 580-615. doi10.1161/01.CIR.91.2.580

Fletcher, G. F., Balady, G. J., Amsterdam, E. A., Chaitman, B., Eckel, R., Fleg, J., . . . Bazzarre, T. (2001). AHA scientific statement: Exercise Standards for testing and training. A statement for healthcare professionals from the American Heart Association. *Circulation, 105*, 1694-1740. doi:10.1161/hc3901.095960

Flynn, K. E., Lin, L., Ellis, S. J., Russell, S. D., Spertus, J. A., Whellan, D. J., . . . Weinfurt, K. P. (2009). Outcomes, health policy, and managed care: relationships between patient-reported outcome measures and clinical measures in outpatients with heart failure. *American Heart Journal, 158*(4 Suppl), S64-71. doi:10.1016/j.ahj.2009.07.010

Flynn, K. E., Pina, I. L., Wheelan, D. J., Lin, L., Blumenthal, J. A., Ellis, S. J., . . . Weinfurt, K. P. (2009). Effects of exercise training on health status in patients with chronic heart failure. *Journal of the American Medical Association, 301*(4), 1451-1459. Doi:10.1001/jama.2009.457

Fontana, A. F., Kerns, R. D., Rosenberg, R. L., Marcus, J. L., & Colonese, K. L. (1986). Exercise training for cardiac patients: Adherence, fitness, and benefits. *Journal of Cardiopulmonary Rehabilitation, 6*(1), 4-15.

Forman, D. E., Fleg, J. L., Kitzman, D. W., Brawner, C. A., Swank, A. M., McKelvie, R. S., . . . Bittner, V. (2012). 6-min walk test provides prognostic utility comparable to cardiopulmonary exercise testing in ambulatory outpatients with systolic heart failure. *Journal of the American College of Cardiology, 60*(25), 2653-2661. doi:10.1016/j.jacc.2012.08.1010

Fowler, S. J., Singh, S. J., & Revill, S. (2005). Reproducibility and validity of the incremental shuttle walking test in patients following coronary artery bypass surgery. *Physiotherapy, 91*, 22-27. doi:10.1016/j.physio.2004.08.009

Frizelle, D. J., Lewin, R. J. P., Kaye, G., Hargreaves, C., Hasney, K., Beaumont, N., & Moniz-Cook, E. (2004). Cognitive-behavioural rehabilitation programme for patients with an implanted cardioverter defibrillator: A pilot study. *British Journal of Health Psychology, 9*, 381-392. doi:10.1348/1359107041557039

Frost, A. E., Langleben, D., Oudiz, R., Hill, N., Horn, E., McLaughlin, V., . . . Barst, R. J. (2005). The 6-min walk test (6MW) as an efficacy endpoint in pulmonary arterial hypertension clinical trials: Demonstration of a ceiling effect. *Vascular Pharmacology, 43*(1), 36-39. doi:10.1016/j.vph.2005.03.003

Gary, R. A., Sueta, C. A., Dougherty, M., Rosenberg, B., Cheek, D., Preisser, J., . . . McMurray, R. (2004). Home-based exercise improves functional performance and quality of life in women with diastolic heart failure. *Heart and Lung, 33*(4), 210-218. doi:10.1016/j.hrtlng.2004.01.004

Gayda, M., Choquet, D., Temfemo, A., & Ahmadi, S. (2003). Cardiorespiratory fitness and functional capacity assessed by the 20-meter shuttle walking test in patients with coronary artery disease. *Archives of Physical Medicine and Rehabilitation, 84*(7), 1012-1016. doi:10.1016/S0003-9993(03)00036-4

Gayda, M., Temfemo, A., Choquet, D., & Ahmaidi, S. (2004). Cardiorespiratory requirements and reproducibility of the six-minute walk test in elderly patients with coronary artery disease. *Archives of Physical Medicine and Rehabilitation, 85*(9), 1538-1543. doi:10.1016/j.apmr.2003.11.037

Gibbons, W. J., Fruchter, N., Sloan, S., & Robert, L. (2001). Reference values for a multiple repetition 6-minute walk test in health adults older than 20 years. *Journal of Cardiopulmonary Rehabilitation, 21*(2), 87-93.

Glass, G. V., & Hopkins, K. D. (1996). *Statistical methods in education and psychology* (3rd ed.). Boston: Allyn & Bacon.

Glazer, K. M., Emery, C. F., Frid, D. J., & Banyasz, R. E. (2002). Psychological predictors of adherence and outcomes among patients in cardiac rehabilitation. *Journal of Cardiopulmonary Rehabilitation, 22*, 40-46.

Gobel, F. L., Nordstrom, L. A., Nelson, R. R., Jorgensen, C. R., & Wang, Y. (1978). The rate-pressure product as an index of myocardial oxygen consumption during exercise in patients with angina pectoris. *Circulation, 57*(3), 549-556. doi:10.1161/01.CIR.57.3.549

Goble, A. J., & Worcester, M. U. C. (1999). *Best practice guidelines for cardiac rehabilitation and secondary prevention*. Melbourne, Australia: Heart Research Centre, on behalf of Department of Human Services Victoria.

Goel, K., Pack, Q. R., Lahr, B., Greason, K. L., Lopez-Jimenez, F., Squires, R. W., . . . Thomas, R. J. (2015). Cardiac rehabilitation is associated with reduced long-term mortality in patients undergoing combined heart valve and CABG surgery. *European Journal of Preventive Cardiology, 22*(2), 159-168. doi:10.1177/2047487313512219

Grace, S. L., Russell, K. L., Reid, R. D., Oh, P., Anand, S., Rush, J., . . . Stewart, D. E. (2011). Effects of cardiac rehabilitation referral strategies on utilization rates: A prospective, controlled study. *Archives of Internal Medicine, 171*(3), 235-241. doi:10.1001/archinternmed.2010.501

Green, D. J., Watts, K., Rankin, S., Wong, P., & O'Driscoll, J. G. (2001). A comparison of the shuttle and 6 minute walking tests with measured peak oxygen consumption in patients with heart failure. *Journal of Science and Medicine in Sport, 4*(3), 292-300. doi:10.1016/S1440-2440(01)80038-4

Green, P., Cohen, D. J., Genereux, P., McAndrew, T., Arnold, S. V., Alu, M., . . . Kirtane, A. J. (2013). Relation between six-minute walk test performance and

outcomes after transcatheter aortic valve implantation (from the PARTNER trial). *American Journal of Cardiology, 112*(5), 700-706. doi:10.1016/j.amjcard.2013.04.046

Gremeaux, V., Deley, G., Duclay, J., Antoine, D., Hannequin, A., & Casillas, J. M. (2009). The 200-m fast-walk test compared with the 6-min walk test and the maximal cardiopulmonary test: a pilot study. *American Journal of Physical Medicine and Rehabilitation, 88*(7), 571-578. doi:10.1097/PHM.0b013e3181aa416b

Gremeaux, V., Hannequin, A., Laroche, D., Deley, G., Duclay, J., & Casillas, J. M. (2012). Reproducibility, validity, and responsiveness of the 200-meter fast walk test in patients undergoing cardiac rehabilitation. *Clinical Rehabilitation, 26*(8), 733-740. doi:10.1177/0269215511427750

Gremeaux, V., Troisgros, O., Benaïm, S., Hannequinn, A., Laurent, Y., Casillas, J.-M., & Benaïm, C. (2011). Determining the minimal clinically important difference for the six-minute walk test and the 200-meter fast-walk test during cardiac rehabilitation program in coronary artery disease patients after acute coronary syndrome. *Archives of Physical Medicine and Rehabilitation, 92*(4), 611-619. doi:10.1016/j.apmr.2010.11.023

Guazzi, M., Dickstein, K., Vicenzi, M., & Arena, R. (2009). Six-minute walk test and cardiopulmonary exercise testing in chronic heart failure: A comparative analysis on clinical and prognostic insights. *Circulation, 2*(6), 549-555. doi:10.1161/CIRCHEARTFAILURE.109.881326

Guilford, J. P. (1954). *Psychometric methods* (2nd ed.). New York: McGraw-Hill Book Company.

Guimarães, G. V., Carvalho, V. O., & Bocchi, E. A. (2008). Reproducibility of the self-controlled six-minute walking test in heart failure patients. *Clinics (Sao Paulo), 63*(2), 201-206. doi:10.1590/S1807-59322008000200008

Guyatt, G. H., Kirshner, B., & Jaeschke, R. (1992a). Measuring health status: What are the necessary measurement properties? *Journal of Clinical Epidemiology, 45*(12), 1341-1345. doi:10.1016/0895-4356(92)90194-R

Guyatt, G. H., Kirshner, B., & Jaeschke, R. (1992b). Response A methodological framework for health status measures: Clarity or oversimplification? *Journal of Clinical Epidemiology, 45*(12), 1353-1355. doi: 10.1016/0895-4356(92)90196-T

Guyatt, G. H., Pugsley, S. O., Sullivan, M. J., Thompson, P. J., Berman, L. B., Jones, N. L., . . . Taylor, D. W. (1984). Effect of encouragement on walking test performance. *Thorax, 39*, 818-822. doi:10.1136/thx.39.11.818

Guyatt, G. H., Sullivan, M. J., Thompson, P. J., Fallen, E. L., Pugsley, S. O., Taylor, D. W., & Berman, L. B. (1985). The 6-minute walk: A new measure of exercise capacity in patient with chronic heart failure. *Canadian Medical Association Journal, 132*(8), 919-923.

Guyatt, G. H., Walter, S., & Norman, G. R. (1987). Measuring change over time: Assessing the usefulness of evaluative instruments. *Journal of Chronic Disease, 40*(2), 171-178. doi:10.1016/0021-9681(87)90069-5

Hamilton, D. M., & Haennel, R. G. (2000). Validity and reliability of the 6-minute walk test in a cardiac rehabilitation population. *Journal of Cardiopulmonary Rehabilitation, 20*(3), 156-165. doi:10.1097/00008483-200005000-00003

Hamm, L. F., Kavanagh, T., Campbell, R. B., Mertens, D. J., Beyene, J., Kennedy, J., & Shephard, R. J. (2004). Timeline for peak improvements during 52 weeks of

outpatient cardiac rehabilitation. *Journal of Cardiopulmonary Rehabilitation, 24*, 374-382.

Hammill, B. G., Curtis, L. H., Schulman, K. A., & Whellan, D. J. (2010). Relationship between cardiac rehabilitation and long-term risks of death and myocardial infarction among elderly Medicare beneficiaries. *Circulation, 121*, 63-70. doi:10.1161/CIRCULATIONAHA.109.876383

Hanson, L. C., McBurney, H., & Taylor, N. F. (2012). The retest reliability of the six minute walk test in patients referred to a cardiac rehabilitation programme. *Physiotherapy Research International, 17*, 55-61. doi:10.1002/pri.513

Hanson, L. C., Taylor, N. F., & McBurney, H. (2016). The 10 m incremental shuttle walk test is a highly reliable field exercise test for patients referred to cardiac rehabilitation: a retest reliability study. *Physiotherapy, 102*(3), 243-248. doi:10.1016/j.physio.2015.08.004

Harada, N., Chiu, V., & Steward, A. (1999). Mobility-related function in older adults: Assessment with a 6-minute walk test. *Archives of Physical Medicine and Rehabilitation, 80*(7), 837-841. doi:10.1016/S0003-9993(99)90236-8

Harvill, L. M. (1991). Standard error of measurement. *Educational Measurement: Issues and Practice, 10*, 33-41. doi:10.1111/j.1745-3992.1991.tb00195.x

Heran, B. S., Chen, J. M., Ebrahim, S., Moxham, T., Oldridge, N., Rees, K., . . . Taylor, R. S. (2011). Exercise-based cardiac rehabilitation for coronary heart disease. *Cochrane Database of Systematic Reviews, 7* Art. No. CD001800. doi:10.1002/14651858.CD001800.pub2

Holland, A. E., Spruit, M. A., Troosters, T., Puhan, M. A., Pepin, V., Saey, D., . . . Singh, S. J. (2014). An official European Respiratory Society/American Thoracic Society technical standard: field walking tests in chronic respiratory

disease. *European Respiratory Journal, 44*(6), 1428-1446. doi:10.1183/09031936.00150314

Horwich, T. B., Fonarow, G. C., Hamilton, M. A., MacLellan, W. R., Woo, M. A., & Tillisch, J. H. (2001). The relationship between obesity and mortality in patients with heart failure. *Journal of the American College of Cardiology, 38*(3), 789-795. doi:10.1016/S0735-1097(01)01448-6

Houchen-Wolloff, L., Boyce, S., & Singh, S. (2015). The minimum clinically important improvement in the incremental shuttle walk test following cardiac rehabilitation. *European Journal of Preventative Cardiology*. doi:10.1177/2047487314540840

Houghton, A. R., Harrison, M., Cowley, A. J., & al, e. (2002). Assessing exercise capacity, quality of life and haemodynamics in heart failure: Do the tests tell us the same thing. *European Journal of Heart Failure, 4(3)*, 289-295. doi:10.1016/S1388-9842(01)00236-7

Howell, D. C. (2012). *Statistical methods for psychology* (8th ed.). Australia: Wadsworth, Cengage Learning.

Husted, J. A., Cook, R. J., Farewell, V. T., & Gladman, D. D. (2000). Methods for assessing responsiveness: a critical review and recommendations. *Journal of Clinical Epidemiology, 53*(5), 459-468. doi:10.1016/S0895-4356(99)00206-1

Hutchinson, P., Meyer, A., & Marshall, B. (2015). Factors influencing outpatient cardiac rehabilitation attendance. *Rehabilitation Nursing, 40*(6), 360-367. doi:10.1002/rnj.202

Ingle, L., Rigby, A. S., Nabb, S., Jones, P. K., Clark, A. L., & Cleland, J. G. (2006). Clinical determinants of poor six-minute walk test performance in patients with left ventricular systolic dysfunction and no major structural heart disease.

European Journal of Heart Failure, 8, 321-325. doi:10.1016/j.ejheart.2005.08.006

Ingle, L., Shelton, R. J., Rigby, A. S., Nabb, S., Clark, A. L., & Cleland, J. G. (2005). The reproducibility and sensitivity of the 6-min walk test in elderly patients with chronic heart failure. *European Heart Journal, 26*(17), 1742 - 1751. doi:10.1093/eurheartj/ehi259

Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology, 59*, 45-52. doi:10.1037/0022-006X.59.1.12

Japanese Circulation Society Joint Working Group. (2014). Guidelines for rehabilitation in patients with cardiovascular disease (JCS 2012). *Circulation Journal, 78*, 2022-2093. doi:10.1253/circj.CJ-66-0094

Jefford, M., Mileshkin, L., Richards, K., Thomson, J., Matthews, J. P., Zalcberg, J., . . . Clarke, D. M. (2004). Rapid screening for depression - validation of the Brief Case-Find for Depression (BCD) in medical oncology and palliative care patients. *British Journal of Cancer, 91*, 900-906. doi:10.1038/sj.bjc.6602057

Jehn, M., Halle, M., Schuster, T., Hanssen, H., Weis, M., Koehler, F., & Schmidt-Trucksass, A. (2009). The 6-min walk test in heart failure: is it a max or sub-maximum exercise test? *European Journal of Applied Physiology, 107*(3), 317-323. doi:10.1007/s00421-009-1128-0

Jelinek, M. V., & Lown, B. (1974). Exercise stress testing for exposure of cardiac arrhythmias. *Progress in Cardiovascular Disease, XVI*(5 (March/April)), 497-522.

Jette, D. U., & Downing, J. (1994). Health status of individuals entering a cardiac rehabilitation program as measured by the Medical Outcomes Study 36-item

short form survey (SF-36). *Physical Therapy, 74*(6), 521-527. doi:10.1093/ptj/74.6.521

Jette, D. U., & Downing, J. (1996). The relationship of cardiovascular and psychological impairments to the health status of patients enrolled in cardiac rehabilitation programs. *Physical Therapy, 76*(2), 130-139. doi:10.1093/ptj/76.2.130

Jolliffe, J., Rees, K., Taylor, R., Thompson, D., Oldridge, N., & Ebrahim, S. (2001). Exercise-based rehabilitation for coronary heart disease. *Cochranne Database of Systematic Reviews 2001 1* Art No: CD001800. doi: 10.1002/14651858

Jolly, K., Lip, G. Y. H., Taylor, R. S., Raftery, J., Mant, J., Lane, D., . . . Stevens, A. (2009). The Birmingham rehabilitation uptake maximisation study (BRUM): a randomised controlled trial comparing home-based with centre-based cardiac rehabilitation. *Heart, 95*, 36-42. doi:10.1136/hrt.2007.127209

Jolly, K., Taylor, R. S., Lip, G. Y. H., & Singh, S. (2008). Reproducibility and safety of the incremental shuttle walking test for cardiac rehabilitation. *International Journal Cardiology, 125*, 144-145. doi:10.1016/j.ijcard.2007.01.037

Juenger, J., Schellberg, D., Kraemer, S., Haunstetter, A., Zugck, C., Herzog, W., & Haass, M. (2002). Health related quality of life in patients with congestive heart failure: comparison with other chronic diseases and relation to functional variables. *Heart, 87*, 235-241. doi:10.1136/heart.87.3.235

Karapolat, H., Eyigor, S., Zoghi, M., Nalbantgil, S., Yagdi, T., Durmaz, B., & Ozbaran, M. (2008). Health related quality of life in patients awaiting heart transplantation. *Tohoku J Exp Med, 214*(1), 17-25. doi:10.1620/tjem.214.17

Kavanagh, T., Myers, M. G., Baigrie, R. S., Mertens, D. J., Sawyer, P., & Shephard, R. J. (1996). Quality of life and cardiorespiratory function in chronic heart

failure: effects of 12 months' aerobic training. *Heart, 76*, 42-49. doi:10.1136/hrt.76.1.42

Keating, J. L., & Matyas, T. A. (1998). Unreliable inferences from reliable measurements. *Australian Journal of Physiotherapy, 44*(1), 5-10.

Keell, S. D., Chambers, J. S., Francis, D. P., Edwards, D. F., & Stables, R. H. (1998). Shuttle-walk test to assess chronic heart failure. *Lancet, 352*(9129), 705.

Kervio, G., Ville, N. S., Leclercq, C., Daubert, J., & Carré, F. (2004a). Intensity and daily reliability of the six-minute walk test in moderate chronic heart failure patients. *Archives of Physical Medicine and Rehabilitation, 85*(9), 1513-1518. doi:10.1016/j.apmr.2003.09.035

Kervio, G., Ville, N. S., Leclercq, C., Daubert, J. C., & Carré, F. (2004b). Cardiorespiratory adaptations during the six-minute walk test in chronic heart failure patients. *European Journal of Cardiovascular Prevention and Rehabilitation, 11*(2), 171-177. doi:10.1097/01.hjr.0000119964.42813.98

Kirshner, B., & Guyatt, G. (1985). A methodological framework for assessing health indicies. *Journal of Chronic Disease, 38*(1), 27-36. doi: 10.1016/0021-9681(85)90005-0

Knox, A. J., Morrison, J. F. J., & Muers, M. F. (1988). Reproducibility of walking test results in chronic obstructive airways disease. *American Journal of Respiratory and Critical Care Medicine, 160*, 388-392. doi:10.1136/thx.43.5.388

Kristjánsdóttir, A., Ragnarsdóttir, M., Einarsson, M. B., & Torfason, B. (2004). A comparison of the 6-minute walk test and symptom limited graded exercise test for Phase II cardiac rehabilitation of older adults. *Journal of Geriatric Physical Therapy, 27*(2), 65-68.

Kroman, S. L., Roos, E. M., Bennell, K. L., Hinman, R. S., & Dobson, F. (2014). Measurement properties of performance-based outcome measures to assess physical function in young and middle-aged people known to be at high risk of hip and/or knee osteoarthritis: a systematic review. *Osteoarthritis and Cartilages, 22*(1), 26-39. doi:10.1016/j.joca.2013.10.021

Kuczmarski, R. J., & Flegal, K. M. (2000). Criteria for definition of overweight in transition: background and recommendations for the United States. *American Journal of Clinical Nutrition, 72*(5), 1074-1081.

Langenfeld, H., Schneider, B., Grimm, W., Beer, M., Knoche, M., Riegger, G., & Kochsiek, K. (1990). The six-minute walk--an adequate exercise test for pacemaker patients? *Pacing and Clinical Electrophysiology, 13*(12 Pt 2), 1761-1765. doi:10.1111/j.1540-8159.1990.tb06886.x

Lawlor, P. R., Filion, K. B., & Eisenberg, M. J. (2011). Efficacy of exercise-based cardiac rehabilitation post-myocardial infarction: A systematic review and meta-analysis of randomized controlled trials. *American Heart Journal, 162*(4), 571-584. doi:10.1016/j.ahj.2011.07.017

Lee, K. M., Lee, K., Chung, C. Y., Ahn, S., Sung, K. H., Kim, T. W., . . . Park, M. S. (2012). Pitfalls and important issues in testing reliability using intraclass correlation coefficients in orthopaedic research. *Clinics in Orthopedic Surgery, 4*(2), 149-155. doi:10.4055/cios.2012.4.2.149

Léger, L. A., & Lambert, J. (1982). A maximal multistage 20-m shuttle run test to predict VO_2 max. *European Journal of Applied Physiology and Occupational Physiology, 49*(1), 1-12. doi:10.1007/BF00428956

Lehmann, G., & Kolling, K. (1996). Reproducibility of cardiopulmonary exercise parameters in patients with valvular heart disease. *Chest, 110*(3), 685-692. doi:10.1378/chest.110.3.685

Lewinter, C., Doherty, P., Gale, C. P., Crouch, S., Stirk, L., Lewin, R. J., . . . Bland, J. M. (2014). Exercise-based cardiac rehabilitation in patients with heart failure: a meta-analysis of randomised controlled trials between 1999 and 2013. *European Journal of Preventative Cardiology, 22*(12), 1504-1512. doi:10.1177/2047487314559853

Lewis, M. E., Newal, C., Townend, J. N., Hill, S. L., & Bonser, R. S. (2001). Incremental shuttle walk test in the assessment of patients for heart transplantation. *Heart, 86*, 183-187. doi:10.1136/heart.86.2.183

Lipkin, D. P., Scriven, A. J., Crake, T., & Poole-Wilson, A. (1986). Six minute walking test for assessing exercise capacity in chronic heart failure. *British Medical Journal (Clinical Research Education), 292*(6521), 653-655. doi:10.1136/bmj.292.6521.653

Lord, F. M., & Novick, M. R. (1968). *Statistical theories of mental test scores*. MA: Addison-Wesley Publishing Company.

Lucas, C., Stevenson, L. W., Johnson, W., Hartley, H., Hamilton, M. A., Waldon, J., . . . Eagen-Bengsten, E. (1999). The 6-min walk and peak oxygen consumption in advanced heart failure: Aerobic capacity and survival. *American Heart Journal, 138*(4), 618-624. doi:10.1016/S0002-8703(99)70174-2

Luszczynska, A., & Sutton, S. (2006). Physical activity after cardiac rehabilitation: Evidence that different types of self-efficacy are important in maintainers and relapsers. *Rehabilitation Psychology, 51*(4), 314-321. doi:/10.1037/0090-5550.51.4.314

Maldonado-Martín, S., Brubaker, P. H., Kaminsky, L. A., Moore, J. B., Stewart, K. P., & Kitzman, D. W. (2006). The relationship of 6-min walk to VO_2peak and VT in older heart failure patients. *Medicine and Science in Sports and Exercise, 38*(6), 1047-1053. doi:10.1249/01.mss.0000222830.41735.14

Mandic, S., Hodge, C., Stevens, E., Walker, R., Nye, E. R., Body, D., . . . Williams, M. J. A. (2013). Effects of community-based cardiac rehabilitation on body composition and physical function in individuals with stable coronary artery disease: 1.6-Year followup. *BioMed Research International, 2013*. doi:10.1155/2013/903604

Mandic, S., Walker, R., Stevens, E., Nye, E. R., Body, D., Barclay, L., & Williams, M. J. (2013). Estimating exercise capacity from walking tests in elderly individuals with stable coronary artery disease. *Disability & Rehabilitation, 35*(22), 1853-1858. doi: 10.3109/09638288.2012.759629

May, G. A., & Nagle, F. J. (1984). Changes in rate-pressure product with physical training of individuals with coronary artery disease. *Physical Therapy, 64*(9), 1361-1366. doi:10.1093/ptj/64.9.1361

McCreery, C., Cradock, K., Fallon, N., Duffy, R., O'Doherty, V., & Kingston, C. (2013). Cardiac rehabilitation guidelines 2013. Retrieved from the Irish Association of Cardiac Rehabilitation, http://www.iacr.info/wp-content/uploads/2015/03/IACR-Guidelines2013.pdf

McGavin, C. R., Gupta, S. P., & McHardy, G. J. (1976). Twelve-minute walking test for assessing disability in chronic bronchitis. *British Medical Journal, 1*(60`3), 822-823. doi:10.1136/bmj.1.6013.822

McGraw, K. O., & Wong, S. P. (1996). Forming inferences about some intraclass correlation coefficients. *Psychological Methods, 1*(1), 30-46. doi:10.1037/1082-989X.1.1.30

McHorney, C. A., & Tarlov, A. R. (1995). Individual-patient monitoring in clinical practice: are available health status surveys adequate? *Quality of Life Research, 4*, 293-307.

McKee, G. (2009). Are there meaingful longitudinal changes in health related quality of life- SF36, in cardiac rehabilitation patients? *European Journal of Cardiovascular Nursing, 8*(1), 40-47. doi:10.1016/j.ejcnurse.2008.04.004

Meland, E., Maeland, J. G., & Laerum, E. (1999). The importance of self-efficacy in cardiovascular risk factor change. *Scandinavian Journal of Public Health, 1*, 11-17. Doi:10.107/BF01593882

Mendis, S., Puska, P., & Norrving, B. (2011). *Global atlas on cardiovascular disease prevention and control*: World Health Organisation, Geneva. Retrieved from the World Health Organisation at http://www.who.int/cardiovascular_diseases/publications/atlas_cvd/en/

Meyer, K., & Laederach-Hofmann, K. (2003). Effects of a comprehensive rehabilitation program on quality of life in patients with chronic heart failure. *Progress in Cardiovascular Nursing, 18*(4), 169-176. doi:10.1111/j.0889-7204.2003.03013

Milani, R. V., & Lavie, C. J. (1998). Prevalence and effects of cardiac rehabilitation on depression in the elderly with coronary heart disease. *The American Journal of Cardiology, 81*, 1233-1236. doi: 10.1016/S0002-9149(98)00121-0

Milani, R. V., Lavie, C. J., & Cassidy, M. M. (1996). Effects of cardiac rehabilitation and exercise training programs on depression in patients after major coronary

events. *American Heart Journal, 132*, 726-732. doi:10.1016/S0002-8703(96)90304-X

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & The PRISMA Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Med, 6*(7), e1000097. doi:10.1371/journal.pmed1000097

Moholdt, T., Madssen, E., Rognmo, O., & Aamot, I. L. (2014). The higher the better? Interval training in coronary heart disease. *Journal of Science and Medicine in Sport, 17*, 506-510. doi:10.1016/j.jsams.2013.07.007

Mokkink, L. B., Terwee, C. B., Gibbons, E., Stratford, P. W., Alonso, J., Patrick, D. L., . . . de Vet, H. C. W. (2010a). Inter-rater agreement and reliability of the COSMIN (COnsensus-based Standards for the selection of health status Measurement INstruments) checklist. *BMC Medical Research Methodology, 10*(82). doi:10.1186/1471-2288-10-82

Mokkink, L. B., Terwee, C. B., Patrick, D. L., Alonso, J., Stratford, P. W., Knol, D. L., . . . De Vet, H. C. W. (2010b). The COSMIN checklist for assessing the methodological qualities of studies on measurement properties of health status measurement instruments: an international Delphi study. *Quality of Life Research, 19*(4), 539-549. doi:10.1007/s11136-010-9606-8

Mokkink, L. B., Terwee, C. B., Patrick, D. L., Alonso, J., Stratford, P. W., Knol, D. L., . . . De Vet, H. C. W. (2010c). The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *Journal of Clinical Epidemiology, 63*, 737-745. doi:10.1016/j.jclinepi.2010.02.006

Morales, F. J., Martinez, A., Mendez, M., Agarrado, A., Ortega, F., Fernandez-Guerra, J., . . . Burgos, J. (1999). A shuttle walk test for assessment of functional capacity in chronic heart failure. *American Heart Journal, 138*, 291-298. doi:10.1016/S0002-8703(99)70114-6

Morales, F. J., Montemayor, T., & Martinez, A. (2000). Shuttle versus six-minute walk test in the prediction of outcome in chronic heart failure. *International Journal of Cardiology, 76*(2-3), 101-105. doi:10.1016/S0167-5273(00)00393-4

Morgan, A. D., Peck, D. F., Buchanan, D. R., & McHardy, G. J. R. (1983). Effects of attitudes and beliefs on exercise tolerance in chronic bronchitis. *British Medical Journal, 286*, 171-173. doi:10.1136-bmj.286.6360.171

Morrin, L., Black, S., & Reid, R. (2000). Impact of duration in a cardiac rehabilitation program on coronary risk profile and health-related quality of life outcomes. *Journal of Cardiopulmonary Rehabilitation, 20*(2), 115-121.

Murray, C. J. L., Vos, T., Lozano, R., Naghavi, M., Flaxman, A. D., Michaud, C., . . . Lopez, A. D. (2012). Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. *The Lancet, 380*(9859), 2197-2223. doi:10.1016/S0140-6736(12)61689-4

Myers, J., Forman, D. E., Balady, G. J., Franklin, B. A., Nelson-Worel, J., Martin, B.-J., . . . Arena, R. (2014). Supervision of exercise testing by nonphysicians: A scientific statement from the American Heart Association. *Circulation, 130*(12), 1014-1027. doi:10.1161/CIR.0000000000000101

Myers, J., Gullestad, L., Vagelos, R., Do, D., Bellin, D., Ross, H., & Fowler, M. B. (1998). Clinical, hemodynamic, and cardiopulmonary exercise test

determinants of survival in patients referred for evaluation of heart failure. *Annals of Internal Medicine, 129*(4), 286-293. doi:10.7326/0003-4819-129-4-199808150-00004

National Guideline Clearinghouse. (2013). *MI - secondary prevention. Secondary prevention in primary and secondary care for patients following a myocardial infarction.* Retrieved from US Department of Health and Human Services, https://www.guideline.gov/summaries/summary/47700

National Heart Foundation. (2012). Victorian Heart Maps. Retrieved from the National Heart Foundation of Australia, http://www.heartfoundation.org.au/information-for-professionals/data-and-statistics/Pages/interactive-map-victoria.aspx

National Heart Foundation of Australia, & Australian Cardiac Rehabilitation Association. (2004). *Recommended framework for cardiac rehabilitation*. Retrieved from Australia: https://www.heartfoundation.org.au/images/uploads/publications/Recommended-framework.pdf

National Institute for Health and Care Excellence. (2013). *Cardiac rehabilitation services: Commissioning guide*. Retrieved from London: https://www.nice.org.uk/guidance/cg172/resources/cardiac-rehabilitation-services-commissioning-guide-394126887877

Neill, W. A., Branch, L. G., De Jong, G., Smith, N. E., Hogan, C. A., Corcoran, P. J., . . . Osberg, S. (1985). Cardiac disability: the impact of coronary disease on patients' daily activities. *Archives of Internal Medicine, 145*, 1642-1647. doi:10.1001/achinte.1985.00360090110017

New Zealand Guidelines Group, & Heart Foundation. (2002). *Evidence-based best practice guideline: Cadiac rehabilitation*. Retrieved from Wellington: New Zealand: http://www.moh.govt.nz/NoteBook/nbbooks.nsf/0/20B3EF98EBDD6D98CC257A5200113C96/$file/cardiac_rehabilitation.pdf

Nogueira, I. D., Servantes, D. M., Nogueira, P. A., Pelcerman, A., Salvetti, X. M., Salles, F., . . . Filho, J. A. (2010). Correlation between quality of life and functional capacity in cardiac failure. *Arquivos Brasileiros de Cardiologia, 95*(2), 238-243. doi:10.1590/S0066-782X2010005000096

Nogueira, P. A., Leal, A. C., Pulz, C., Nogueira, I. D., & Filho, J. A. (2006). Clinical reliability of the 6 minute corridor walk test performed within a week of a myocardial infarction. *International Heart Journal, 47*(4), 533-540. doi:10.1536/ihj.47.533

Noonan, V., & Dean, E. (2000). Submaximal exercise testing: clinical application and interpretation. *Physical Therapy, 80*, 782-807. doi:10.1093/ptj/80.8.782

Norman, G. R., & Streiner, D. L. (2008). *Biostatistics: the bare essentials* (3rd ed.). Hamilton, Ontario: BC Decker.

Norton, K., Norton, L., & Sadgrove, D. (2010). Position statement on physical activity and exercise intensity terminology. *Journal of Science and Medicine in Sport, 12*(2010), 496-502. doi:10.1016/j.jsams.2009.09.008

NSW Department of Health. (2006). *NSW chronic care program: Implementing rehabilitation for chronic diseases - Volume 2*. Retrieved from Sydney: http://www0.health.nsw.gov.au/policies/gl/2006/pdf/GL2006_022.pdf

O'Keefe, S. T., Lye, M., Donnellan, C., & Carmichael, D. N. (1998). Reproducibility and responsiveness of quality of life assessment and six minute walk test in elderly heart failure patients. *Heart, 80*(4), 377-382. doi:10.1136/hrt.80.4.377

Oldridge, N. (2012). Exercise-based cardiac rehabilitation in patients with coronary heart disease: Meta-analysis outcomes revisited. *Future Cardiology, 8*(5), 729-751. doi:10.2217/fca.12.34)

Oldridge, N. B., Guyatt, G. H., Fischer, M. E., & Rim, A. A. (1988). Cardiac rehabilitation after myocardial infarction: Combined experience of randomized clinical trials. *Journal of the American Medical Association, 260*, 945-950. doi:10.1001/jama.1988.03410070073031

Olper, L., Cervi, P., Santi, F., Meloni, C., & Gatti, R. (2011). Validation of the treadmill Six-Minute Walk Test in people following cardiac surgery. *Physical Therapy, 91*(4), 566-576. doi:10.2522/ptj.20100156

Opasich, C., De Feo, S., Pinna, G. D., & al, e. (2004). Distance walked in the 6-minute test soon after cardiac surgery: Toward an efficient use in the individual patient. *Chest, 126*, 1796-1801. doi: 10.1378/chest.126.6.1796

Opasich, C., Pinna, G. D., Mazza, A., Febo, O., Riccardi, P. G., Capomolla, S., & al, e. (1998). Reproducibility of the six-minute walking test in patients with chronic heart failure: practical implications. *American Journal of Cardiology, 81*, 1497-1500. doi:10.1016/S0002-9149(98)00218-5

Opasich, C., Pinna, G. D., Mazza, A., Febo, O., Riccardi, R., Riccardi, P. G., . . . Tavazzi, L. (2001). Six-minute walking performance in patients with moderate-to-severe heart failure; is it a useful indicator in clinical practice? *Eur Heart J, 22*(6), 488-496. doi:10.1053/euhj.2000.2310

Pack, Q. R., Mansour, M., Barboza, J. S., Hibner, B. A., Mahan, M. G., Ehrman, J. K., . . . Keteyian, S. J. (2013). An early appointment to outpatient cardiac rehabilitation at hospital discharge improves attendance at orientation. *Circulation, 127*(3), 349-355. doi:10.1161/CIRCULATIONAHA.112.121996

Palange, P., Ward, S. A., Carlsen, K.-H., Casaburi, R., Gallagher, C. G., Gosselink, R., . . . Whipp, B. J. (2007). Recommendations on the use of exercise testing in clinical practice. *European Respiratory Journal, 29*, 185-209. doi:10.1183/09031936.00046906

Pate, R. R., Pratt, M., Blair, S. N., Haskell, W. L., Macera, C. A., Bouchard, C., . . . Wilmore, J. H. (1995). Physical activity and public health: A recommendation from the centers for disease control and prevention and the American College of Sports Medicine. *The Journal of the American Medical Association, 273*(5), 402-407. doi:10.1001/jama.1995.03520290054029.

Patrick, J. M. (2005). *Effects of long versus short bouts of exercise in heart failure patients.* (Masters Degree Thesis), La Trobe University, Bundoora.

Patrick, J. M. (2008, August 2008). *When is one six minute walk test enough?* Paper presented at the Australian Cardiac Rehabilitation Association Conference, Alice Springs, Australia.

Pavy, B., Iliou, M.-C., Vergès-Patois, B., Brion, R., Monpère, C., Carré, F., . . . Marcadet, D. (2012). French Society of Cardiology guidelines for cardiac rehabilitation in adults. *Archives of Cardiovascular Diseases, 105*(5), 309-328. doi:http://dx.doi.org/10.1016/j.acvd.2012.01.010

Pavy, B., Iliou, M. C., Meurin, P., Tabet, J.-Y., & Corone, S. (2006). Safety of exercise training for cardiac patients. Results of the French REgistry of

Complications during cardiac rehabilitation. *Archives of Internal Medicine, 166*(21), 2329-2334. doi:10.1001/archinte.166.21.2329

Payne, G. E., & Skehan, J. D. (1996). Shuttle walking test: A new approach for evaluating patients with pacemakers. *Heart, 75*, 414-418. doi:10.1136/hrt.75.4.414

Peeters, & Mets. (1996). The 6-minute walk as an appropriate exercise testing in elderly patients with chronic heart failure. *Journal of Gerontology: Medical Sciences, 51*, M147-151. doi:10.1093/gerona/51A.4.M147

Pepera, G., Cardoso, F., Taylor, M. J. D., Peristeropoulos, A., & Sandercock, G. R. H. (2013). Predictors of shuttle walking test performance in patients with cardiovascular disease. *Physiotherapy, 99*, 317-322. doi:10.1016/j.physio.2013.01.003

Pepera, G., McAllister, J., & Sandercock, G. (2010). Long-term reliability of the incremental shuttle walking test in clinically stable cardiovascular disease patients. *Physiotherapy, 96*, 222-227. doi:10.1016/j.physio.2009.11.010

Perk, J., De Backer, G., Gohlke, H., Graham, I., Reiner, Ž., Verschuren, M., . . . Wolpert, C. (2012). European Guidelines on cardiovascular disease prevention in clinical practice (version 2012). The Fifth Joint Task Force of the European Society of Cardiology and Other Societies on Cardiovascular Disease Prevention in Clinical Practice (constituted by representatives of nine societies and by invited experts) Developed with the special contribution of the European Association for Cardiovascular Prevention & Rehabilitation (EACPR). *European Heart Journal, 33*(13), 1635-1701. doi:10.1093/eurheartj/ehs092

Pichurko, B. M. (2012). Exercising your patient: Which test(s) and when? *Respiratory Care, 57*(1), 100-113. doi:10.4187/respcare.01428

Piepoli, M. F., Corrà, U., Adamopoulos, S., Benzer, W., Bjarnason-Wehrens, B., Cupples, M., . . . Giannuzzi, P. (2014). Secondary prevention in the clinical management of patients with cardiovascular diseases. Core components, standards and outcome measures for referral and delivery. *European Journal of Preventative Cardiology, 21*(6), 664-681. doi:10.1177/2047487312449597

Piepoli, M. F., Corra, U., Benzer, W., Bjarnason-Wehrens, B., Dendale, P., Gaita, D., . . . Schmid, J. P. (2010). Secondary prevention through cardiac rehabilitation: from knowledge to implementation. A position paper from the Cardiac Rehabilitation Section of the European Association of Cardiovascular Prevention and Rehabilitation. *European Journal of Cardiovascular Prevention & Rehabilitation, 17*(1), 1-17. doi:10.1097/HJR.0b013e3283313592

Pinna, G. D., Opasich, C., Mazza, A., Tangenti, A., Maestri, R., & Sanarico, M. (2000). Reproducibility of the six-minute walking test in chronic heart failure patients. *Statistics in Medicine, 19*(22), 3087-3094. doi: 10.1002/1097-0258(20001130)19:22<3087::AID-SIM628>3.0.CO;2-G

Polit, D. F., & Yang, F. M. (2016). *Measurement and the measurement of change: A primer for the health professional*. Philadelphia: Wolters Kluwer.

Porter, M. E. (2010). What is value in health care? *New England Journal of Medicine, 363*(26), 2477-2481. doi:10.1056/NEJMp1011024

Portney, L. G., & Watkins, M. P. (2000). *Foundations of clinical research: Applications to practice* (2 ed.). Upper Saddle River, New Jersey: Prentice Hall Health.

Portney, L. G., & Watkins, M. P. (2009). *Foundations of clinical research: Applications to practice* (3 ed.). Upper Saddle River, New Jersey: Pearson Prentice Hall.

Price, K. J., Gordon, B. A., Bird, S. R., & Benson, A. C. (2016). A review of guidelines for cardiac rehabilitation exercise programmes: Is there an international consensus? *European Journal of Preventive Cardiology*. doi:10.1177/2047487316657669

Pulz, C., Diniz, R. V., Alves, A. N., Tebexreni, A. S., Carvalho, A. C., de Paola, A. A., & Almeida, D. R. (2008). Incremental shuttle and six minute walking tests in the assessment of functional capacity in chronic heart failure. *Canadian Journal of Cardiology, 24*, 131-135. doi:10.1016/S0828-282X(08)70569-5

Quittan, M., Sturm, B., Wiesinger, G. F., Pacher, R., & Fialka-Moser, V. (1999). Quality of life in patients with chronic heart failure: A randomized controlled trial of changes induced by a regular exercise program. *Scandinavian Journal of Rehabilitation Medicine, 31*, 223-228.

Rankin, G., & Stokes, M. (1998). Reliability of assessment tools in rehabilitation: an illustration of appropriate statistical analyses. *Clinical Rehabilitation, 12*, 187-199. doi:10.1191/026921598672178340

Rasekaba, T., Lee, A. L., Naughton, M. T., Williams, T. J., & Holland, A. E. (2009). The six minute walk test: a useful metric for the cardiopulmonary patient. *Internal Medicine Journal, 39*, 495-501. doi:10.1111/j.1445-5994.2008.01880.x

Redfern, J., Hyun, K., Chew, D. P., Astley, C., Chow, C., Aliprandi-Costa, B., . . . Briffa, T. (2014). Prescription of secondary prevention medications, lifestyle advice, and referral to rehabilitation among acute coronary syndrome

inpatients: results from a large prospective audit in Australia and New Zealand. *Heart, 100*, 1281-1288. doi:10.1136/heartjnl-2013-305296

Reeves, G. R., Gupta, S., & Forman, D. E. (2016). Evolving role of exercise testing in contemporary cardiac rehabilitation. *Journal of Cardiopulmonary Rehabilitation and Prevention, 36*, 309-319. doi:10.1097/HCR.0000000000000176

Reeves, G. R., & Whellan, D. J. (2010). Recent advances in cardiac rehabilitation. *Current Opinion in Cardiology, 25*(6), 589-596. doi:10.1097/HCO.0b013e32833f0208

Rejeski, W. J., Foy, C. G., Brawley, L. R., Brubaker, P. H., Focht, B. C., Norris III, J. L., & Smith, M. L. (2002). Older adults in cardiac rehabilitation: A new strategy for enhancing physical function. *Medicine and Science in Sports and Exercise, 34*(11), 1705-1713. doi:10.1097/00005768-200211000-00003

Revill, S. M., Morgan, M. D. L., Singh, S. J., Williams, J., & Hardman, A. E. (1999). The endurance shuttle walk: a new field test for the assessment of endurance capacity in chronic obstructive pulmonary disease. *Thorax, 54*, 213-222.

Riley, M., McParland, J., Stanford, C. F., & Nicholls, D. P. (1992). Oxygen consumption during corridor walk testing in chronic cardiac failure. *European Heart Journal, 13*(6), 789-793. doi:10.1136/thx.54.3.213

Roberts, E., Li, F. K. W., & Sykes, K. (2006). Validity of the 6-minute walk test for assessing heart rate recovery after an exercise-based cardiac rehabilitation programme. *Physiotherapy, 92*(2), 116-121. doi:10.1016/j.physio.2005.06.005

Robinson, H. J., Samani, N. J., & Singh, S. J. (2011). Can low risk cardiac patients be 'fast tracked' to Phase IV community exercise schemes for cardiac

rehabilitation? A randomised controlled trial. *International Journal Cardiology, 146*, 159-163. doi:10.1016/j.ijcard.2009.06.027

Rosenbaum, A. N., Kremers, W. K., Schirger, J. A., Thomas, R. J., Squires, R. W., Allison, T. G., . . . Edwards, B. S. (2016). Association between early cardiac rehabilitation and long-term survival in cardiac transplant recipients. *Mayo Clinic Proceedings, 91*(2), 149-156. doi:10.1016/j.mayocp.2015.12.002

Rothstein, J. M. (1985). Measurement and clinical practice: Theory and application. In J. M. Rothstein (Ed.), *Measurement in physical therapy* (pp. 1-8). New York: Churchill Livingstone.

Roul, G., Germain, P., & Bareiss, P. (1998). Does the 6-minute walk test predict the prognosis in patients with NYHA class II or III chronic heart failure? *American Heart Journal, 136*(3), 449-457. doi:10.1016/S0002-8703(98)70219-4

Ruo, B., Rumsfeld, J. S., Pipkin, S., & Whooley, M. A. (2004). Relation between depressive symptoms and treadmill exercise capacity in the Heart and Soul study. *American Journal of Cardiology, 94*(1), 96-99. doi:10.1016/j.amjcard.2004.03.035

Sanderson, B. K., Southard, D., & Oldridge, N. (2004). Outcomes evaluation in cardiac rehabilitation/secondary prevention programs: Improving patient care and program effectiveness. *Journal of Cardiopulmonary Rehabilitation, 24*(2), 68-79.

Scholtes, V. A., Terwee, C. B., & Poolman, R. W. (2011). What makes a measurement instrument valid and reliable? *Injury, 42*, 236-240. doi:10.1016/j.injury.2010.11.042

Scott, I. A. (2011). Cautionary tales in the clinical interpretation of trials assessing therapy-induced changes in health status. *The International Journal of Clinical Practice, 65*(5), 536-546. doi:10.1111/j.1742-1241.2011.02654.x

Scottish Intercollegiate Guidelines Network. (2002). *Cardiac rehabilitation: A national clinical guideline*. Retrieved from Edingburgh: http://www.scotphn.net/wp-content/uploads/2015/11/Cardiac_Rehabilitation.pdf

Shen, B.-J., McCreary, C. P., & Myers, H. F. (2004). Independent and mediated contributions of personality, coping, social support, and depressive symptoms to physical functioning outcome among patients in cardiac rehabilitation. *Journal of Behavioural Medicine, 27*(1), 39-62. doi:10.1023/B:JOBM.0000013643.36767.22

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin, 86*(2), 420-428. doi:10.1037/0033-2909.86.2.420

Simms, K., Myers, C., Adams, J., Hartment, J., Lindsey, C., Doler, M., & Suhr, J. (2007). Exercise tolerance testing in a cardiac rehabilitation setting: An exploratory study of its safety and practicality for exercise prescription and outcome data collection. *Baylor University Medical Centre Proceedings, 20*(4), 344-347.

Singh, S. J., Morgan, M. D., Scott, S., Walters, D., & Hardman, A. E. (1992). Development of a shuttle walking test of disability in patients with chronic airways obstruction. *Thorax, 47*, 1019-1024. doi:10.1136/thx.47.12.1019

Singh, S. J., Puhan, M. A., Andrianopoulos, V., Hernandes, N. A., Mitchell, K. E., Hill, C. J., . . . Holland, A. E. (2014). An official systematic review of the

European Respiratory Society/American Thoracic Society: Measurement properties of field walking tests in chronic respiratory disease. *European Respiratory Journal, 44*(6), 1447-1478. doi:10.1183/09031936.00150414

Singh, S. J., Spruit, M. A., Troosters, T., & Holland, A. E. (2015). The 6-min walk test in patients with COPD: walk this way! *Thorax, 70*(86), 86. doi:10.1136/thoraxjnl-2014-205928

Smith, S. C., Benjamin, E. J., Bonow, R. O., Braun, L. T., Creager, M. A., Franklin, B. A., . . . Jones, D. W. (2011). AHA/ACCF secondary prevention and risk reduction therapy for patients with coronary and other atherosclerotic vascular disease: 2011 update. *Circulation, 124*(22), 2458-2473. doi:10.1161/CIR.0b013e318235eb4d

Solway, S., Brooks, D., Lacasse, Y., & Thomas, S. (2001). A qualitative systematic overview of the measurement properties of functional walk tests used in the cardiorespiratory domain. *Chest, 119*, 256-270. doi:10.1378/chest.119.1.256

St John of God Health Care. (2010a). Cardiac Rehabilitation. Retrieved from St John of God Health Care, http://www.sjog.org.au/hospitals/bendigo_hospital/hospital_services/cardiac_services/cardiac_rehabilitation.aspx

St John of God Health Care. (2010b). Welcome to Bendigo. Retrieved from St John of God Health Care, http://www.sjog.org.au/hospitals/bendigo.aspx

Steele, B. (1996). Timed walking tests of exercise capacity in chronic cardiopulmonary illness. *Journal of Cardpulmonary Rehabilitation and Prevention, 16*, 25-33.

Stelken, A. M., Younis, L. T., Jennison, S. H., Miller, D. D., Miller, L. W., Shaw, L. J., . . . Chaitman, B. R. (1996). Prognostic value of cardiopulmonary exercise

testing using percent achieved of predicted peak oxygen uptake for patients with ischemic and dilated cardiomyopathy. *Journal of the American College of Cardiology, 27*(2), 345-352. doi:10.1016/0735-1097(95)00464-5

Stevens, D., Elpern, E., Sharma, K., Szidon, P., Ankin, M., & Kesten, S. (1999). Comparisons of hallway and treadmill six-minute walk tests. *American Journal of Respiratory and Critical Care Medicine, 160*(5), 1540-1543. doi:10.1164/ajrccm.160.5.9808139

Stewart, A. L., Hays, R. D., Wells, K. B., Rogers, W. H., Spritzer, K. L., & Greenfield, S. (1994). Long-term functioning and well-being outcomes associated with physical activity and exercise in patients with chronic conditions in the medical outcomes study. *Journal of Clinical Epidemiology, 47*(7), 719-730. doi:10.1016/0895-4356(94)90169-4

Stewart, A. L., King, A. C., & Haskell, W. L. (1993). Endurance exercise and health-related quality of life in 50-65 year-old adults. *The Gerontologist, 33*(6), 782-789. doi:10.1093/geront/33.6.782

Strath, S. J., Kaminsky, L. A., Ainsworth, B. E., Ekelund, U., Freedson, P. S., Gary, R. A., . . . Swartz, A. M. (2013). Guide to the assessment of physical activity: clinical and research applications: a scientific statement from the American Heart Association. *Circulation, 128*(20), 2259-2279. doi:10.1161/01.cir.0000435708.67487.da

Streiner, D. L., & Norman, G. R. (2008). *Health measurement scales* (4th edition ed.). Oxford: Oxford University Press.

Suaya, J. A., Shepard, D. S., Normand, S.-L. T., Ades, P. A., Prottas, J., & Stason, W. B. (2007). Use of cardiac rehabilitation by Medicare beneficiaries after

myocardial infarction or coronary bypass surgery. *Circulation, 116*(15), 1653-1662. doi:10.1161/CIRCULATIONAHA.107.701466

Suaya, J. A., Stason, W. B., Ades, P. A., Normand, S.-L. T., & Shepard, D. S. (2009). Cardiac rehabilitation and survival in older coronary patients. *Journal of the American College of Cardiology, 54*(1), 25-33. doi:10.1016/j.jacc.2009.01.078

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (6th ed.). Boston: Pearson Education.

Tallaj, T. A., Sanderson, B., Breland, J., Adams, C., Schumann, C., & Bittner, V. (2001). Assessment of functional outcomes using the 6-minute walk test in cardiac rehabilitation: comparisons of patients with and without left ventricular dysfunction. *Journal of Cardiopulmonary Rehabilitation, 21*, 221-224.

Taylor, A., Bell, J., & Lough, F. (2010). Cardiac rehabilitation and secondary prevention. In J. A. Pryor & S. A. Prasad (Eds.), *Physiotherapy for respiratory and cardiac problems* (4th ed., pp. 470-494). Edinburgh: Churchill Livingstone.

Taylor, N. F., Dodd, K. J., & Graham, H. K. (2004). Test-retest reliability of hand-held dynamometric strength testing in young people with cerebral palsy. *Archives of Physical Medicine and Rehabilitation, 85*(1), 77-80. doi:10.1016/S0003-9993(03)00379-4

Taylor, R. S., Brown, A., Ebrahim, S., Jolliffe, J., Noorani, H., Rees, K., . . . Oldridge, N. (2004). Exercise-based rehabilitation for patients with coronary heart disease: systematic review and meta-analysis of randomized controlled trials. *The American journal of medicine, 116*(10), 682-692. doi:10.1016/j.amjmed.2004.01.009

Terluin, B., Eekhout, I., Terwee, C. B., & De Vet, H. C. W. (2015). Minimal important change (MIC) based on a predictive modeling approach was more precise than MIC based on ROC analysis. *Journal of Clinical Epidemiology, 68*(12), 1288-1296. doi: 10.1016/j.jclinepi.2015.03.015

Terwee, C. B., Bot, S. D. M., de Boer, M. R., van der Windt, D. A. W. M., Knol, D. L., Dekker, J., . . . de Vet, H. C. W. (2007). Quality criteria were proposed for measurement properties of health status questionnaires. *Journal of Clinical Epidemiology, 60*, 34-42. doi:10.1016/j.jclinepi.2006.03.012

Terwee, C. B., Jansma, E. P., Riphagen, I. I., & de Vet, H. C. W. (2009). Development of a methodological PubMed search filter for finding studies on measurement properties of measurement instruments. *Quality Life Research, 18*(8), 115-1123. doi:10.1007/sl 1136-009-9528-5

Terwee, C. B., Mokkink, L. B., Knol, D. L., Ostelo, R. W. J. G., Bouter, L. M., & de Vet, H. C. W. (2012). Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. *Quality of Life Research, 21*(4), 651-617. doi:10.1007/s11136-011-9960-1

Terwee, C. B., Prinsen, C. A. C., Garotti, M. G. R., Suman, A., de Vet, H. C. W., & Mokkink, L. B. (2016). The quality of systematic reviews of health-related outcome measurement instruments. *Quality of Life Research, 25*(4), 767-779. doi:10.1007/s11136-015-1122-4

Testa, M., & Simonson, D. C. (1996). Assessment of quality-of-life outcomes. *New England Journal of Medicine, 334*(13), 835-840. doi:10.1056/NEJM199603283341306

Thomas, R. J., King, M., Lui, K., Oldridge, N., Piña, I. L., & Spertus, J. (2010). AACVPR/ACCF/AHA 2010 update: performance measures on cardiac rehabilitation for referral to cardiac rehabilitation/secondary prevention services. *Journal of the American College of Cardiology, 56*(14), 1159-1167. doi:10.1016/j.jacc.2010.06.006

Thomas, R. J., King, M., Lui, K., Oldridge, N., Piña, I. L., Spertus, J., . . . Goff, D. C. (2007). AACVPR/ACC/AHA 2007 performance measures on cardiac rehabilitation for referral to and delivery of cardiac rehabilitation/secondary prevention services : Endorsed by the American College of Chest Physicians, American College of Sports Medicine, American Physical Therapy Association, Canadian Association of Cardiac Rehabilitation, European Association for Cardiovascular Prevention and Rehabilitation, Inter-American Heart Foundation, National Association of Clinical Nurse Specialists, Preventive Cardiovascular Nurses Association, and the Society of Thoracic Surgeons. *Journal of American College of Cardiology, 50*(14), 1400-1433. doi: 10.1016/j.jacc.2007.04.033

Thompson, P. D., Arena, R., Riebe, D., & Pescatello, L. S. (2013). ACSM's new preparticipation health screening recommendations from ACSM's guidelines for exercise testing and prescription. *Current Sports Medicine Reports, 12*(4), 215-217. doi:10.1249/JSR.0b013e31829a68cf

Tilloston, K. M., & Burton, A. K. (1991). Noninvasive measurement of lumbar sagittal mobility in assessment of the flexicurve technique. *Spine, 16*, 29-33.

Tobin, D., & Thow, M. K. (1999). The 10 m shuttle walk test with holter monitoring: An objective outcome measure for cardiac rehabilitation. *Coronary Health Care, 3*(1), 3-17. doi:10.1016/S1362-3265(99)80028-5

Tremblay, M. S., Warburton, D. E. R., Janssen, I., Paterson, D. H., Latimer, A. E., Rhodes, R. E., . . . Duggan, M. (2011). New Canadian Physical Activity Guidelines. *Applied Physiology, Nutrition, and Metabolism, 36*(1), 36-46. doi:10.1139/H11-009

Troosters, T., Gosselink, R., & Decramer, M. (1999). Six minute walking distance in health elderly subjects. *European Respiratory Journal, 14*(2), 270-274. doi:10.1034/j.1399-3003.1999.14bo6.x

United Kingdom Department of Health. (2000). *National service framework for coronary heart disease*. England: Department of Health Retrieved from https://www.gov.uk/government/publications/quality-standards-for-coronary-heart-disease-care.

Vague, J. (1956). The degree of masculine differentiation of obesities: A factor determining predisposition to diabetes, atherosclerosis, gout, and uric calculous disease. *American Journal of Clinical Nutrition, 4*, 20-34.

Van de Werf, V., Ardissino, D., Betriu, A., Cokkinos, D. V., Falk, E., Fox, K. A. A., . . . Wijns, W. (2003). Management of acute myocardial infarction in patients presenting with ST-segment elevation. *European Heart Journal, 24*, 28-66. doi:10.1016/S0195-668X(02)00618-8

Verrill, D., Barton, C., Beasley, W., Lippard, M., & King, C. (2003). Six-minute walk performance and quality of life comparisons in North Carolina cardiac rehabilitation programs. *Heart Lung, 32*, 41-51. doi: 10.1067/mhl.2003.7

Walter, S., Eliasziw, M., & Donner, A. (1998). Sample size and optimal designs for reliability studies. *Statistics in Medicine, 17*, 101-110. doi:10.1002/(SICI)1097-0258(19980115)17:1<101::AID-SIM727>3.0.CO;2-E

Ware, J. E. J., & Sherbourne, C. D. (1992). The MOS 36-item short-form health survey (SF-36). *Medical Care, 30*(6), 473-483.

Weir, J. P. (2005). Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. *Journal of Strength and Condition Research, 19*(1), 231-240.

Weir, N. A., Brown, A. W., Shlobin, O. A., Smith, M. A., Reffett, T., Battle, E., . . . Nathan, S. D. (2013). The influence of alternative instruction of 6-min walk test distance. *Chest, 144*(6), 1900-1905. doi: 10.1378/chest.13-0287

Wenger, H. A., & Bell, G. J. (1986). The interactions of intensity, frequency and duration of exercise training in altering cardiorespiratory fitness. *Sports Medicine, 3*(5), 346-356. doi:10.2165/00007256-198603050-00004

Westlake, C., Dracup, K., Fonarow, G., & Hamilton, M. (2005). Depression in patients with heart failure. *Journal of Cardiac Failure, 11*(1), 30-35. doi:10.1016/j.cardfail.2004.03.007

Williams, J. I., & Naylor, C. D. (1992). How should health status measures be assessed? Cautionary notes on procrustean frameworks. *Journal of Clinical Epidemiology, 45*(12), 1347-1351. doi:10.1016/0895-4356(92)90195-S

Woodgate, J., & Brawley, L. (2008). Self-efficacy for exercise in cardiac rehabilitation. *Journal of Health Psychology, 13*(3), 366-387.

Woodgate, J., Brawley, L. R., & Weston, Z. J. (2005). Maintenance cardiac rehabilitation exercise adherence: Effects of task and self-regulatory self-efficacy. *Journal of Applied Social Psychology, 35*(1), 183-197. doi:10.1177/1359105307088141

Woodruffe, S., Neubeck, L., Clark, R. A., Gray, K., Ferry, C., Finan, J., . . . Briffa, T. G. (2015). Australian Cardiovascular Health and Rehabilitation Association

(ACRA) core components of cardiovascular disease secondary prevention and cardiac rehabilitation 2014. *Heart, Lung and Circulation, 24*(5), 430-441. doi: 10.1016/j.hlc.2014.12.008

World Health Organization. (1964). *Rehabilitation of patients with cardiovascular diseases: report of a WHO Expert Committee [meeting held in Geneva from 23 to 29 July 1963).* (Technical Report series, 831). Retrieved from Geneva: Switzerland.

World Health Organization. (2010). *Global recommendations on physical activity for health*. Geneva, Switzerland Retrieved from http://whqlibdoc.who.int/publications/2010/9789241599979_eng.pdf?ua=1.

World Health Organization Expert Committee. (1993). *Rehabilitation after cardiovascular diseases, with special emphasis on developing countries: report of a WHO expert committee [meeting health in Geneva from 21 to 28 October 1991]* (Technical Report Series No. 831). Geneva, Switzerland.

Wright, D. J., Khan, K. M., Gossage, E. M., & Saltissi, S. (2001). Assessment of a low-intensity cardiac rehabilitation programme using the six-minute walk test. *Clinical Rehabilitation, 15*, 119-124. doi:10.1191/026921501671364449

Wu, G., Sanderson, B., & Bittner, C. (2003). The 6-minute walk test: how important is the learning effect? *American Heart Journal, 146*, 119-124. doi:10.1034/j.1399-3003.1999.14b06.x

Wyrwich, K. W. (2004). Minimal important difference thresholds and the standard error of measurement: Is there a connection? *Journal of Biopharmaceutical Statistics, 14*, 97-110. doi:10.1081/BIP-120028508

Zainuldin, R., Mackey, M., & Alison, J. A. (2012). Prescription of walking exercise intensity from the incremental shuttle walk test in people with chronic

obstructive pulmonary disease. *American Journal of Physical Medicine and Rehabilitation, 91*(7), 592-600. doi:10.1097/PHM.0b013e31824660bd

Zugck, C., Kruger, C., Durr, S., Gerber, D., Haunstetter, A., Horning, K., . . . Haass, M. (2000). Is the 6-minute walk test a reliable substitute for peak oxygen uptake in patients with dilated cardiomyopathy? *European Heart Journal, 21*(7), 540-549. doi:10.1053/euhj.1999.1861