©Journal of Sports Science and Medicine (2012) **11,** 690-694
http://www.jssm.org

Research article

Test-retest reliability of the Aerobic Power Index Submaximal Exercise Test in cancer patients

Bonnie J. Furzer [1]✉, Karen E. Wallman [1], Timothy R. Ackland [1] and David J.L. Joske [1 2 3]
[1] School of Sport Science, Exercise & Health, University of Western Australia, Western Australia, Australia, [2] Solaris-Care Foundation, Collaborative Research Team, Western Australia [3] Sir Charles Gairdner Hospital, Western Australia

Abstract
The purpose of this study was to investigate the reliability of the Aerobic Power Index (API) submaximal cardiorespiratory exercise test, as well as associated variables of oxygen uptake (ml kg^{-1} min^{-1}) and ratings of perceived exertion (RPE) in cancer patients who are generally unable to complete maximal or lengthy aerobic fitness tests. Twenty male and female participants (11 male; 9 female) aged between 18 and 70 y (mean = 53.28 ± 11.82 y) were recruited with medical consent within 4 weeks of completing chemotherapy treatment for a lymphohaematopoietic cancer (LHC). Of the twenty recruited participants' 2 were excluded from analysis due to disease relapse or complications unrelated to testing occurring within the month following testing. Intra-class correlation coefficient (ICC) scores for power output ($W \cdot kg^{-1}$) and oxygen uptake ($ml \cdot kg^{-1} \cdot min^{-1}$) were highly reliable ($R_1$ = 0.96 and 0.96, respectively) and the ICC for RPE was moderately reliable (R_1 = 0.83). Technical error of measurement results for power output ($W \cdot kg^{-1}$), oxygen uptake (ml kg^{-1} min^{-1}) and RPE were 0.11$W.kg^{-1}$, 1.18 $ml \cdot kg^{-1} \cdot min^{-1}$ and 1.0 respectively. A Pearson's product-moment correlation demonstrated a strong relationship between power output ($W \cdot kg^{-1}$) and oxygen uptake ($ml \cdot kg^{-1} \cdot min^{-1}$) for both trials (r = 0.93 and 0.89, respectively). Results demonstrate that the API test is a highly reliable protocol for use with a LHC population and can be considered a clinically feasible, safe and tolerable exercise test.

Key words: Exercise test, non-Hodgkin's lymphoma, physical fitness, aerobic fitness, lymphohaematopoietic cancers

Introduction

Lymphohaematopoietic cancers (LHC) are cancers of the blood and lymphatic systems and include Hodgkin lymphoma (HL), non-Hodgkin's lymphoma (NHL) and myeloma. The most common LHC, NHL has seen a gradual increase in diagnosis due to improved detection and screening (Australian Institute of Health and Welfare, 2008a). However, prevalence is well countered by significant increases in 5-year survival rates in Australia from 46% in 1982-1986 to 63% in 1998-2004 in both males and females (Australian Institute of Health and Welfare, 2008b). An increase in survival brings with it a focus on the rehabilitation and long term health of cancer patients, with exercise emerging as a key strategy to recovery, preventing recurrence and managing comorbidities. The growth in research and clinical presentation has lead to a crucial need to examine relevant, tolerable, safe and reliable cardiorespiratory fitness tests that are applicable across the entire cancer continuum, and importantly, can be applied in a clinical setting, as well as in a laboratory (Courneya, 2003).

The gold standard measurement of cardiorespiratory function is maximal oxygen consumption testing. However, maximal testing in a population that exhibits high levels of fatigue and deconditioning (Curt et al., 2000), which is typically exacerbated by exercise, may be deleterious to physical functioning and deter potential participants. Consequently, submaximal testing presents a more appropriate means to assess cardiorespiratory fitness in cancer patients during and immediately post treatment, and in turn may encourage more patients to participate in the testing and ongoing exercise.

With regards to submaximal testing there are a variety of testing modalities available including step tests, treadmill tests and cycling based tests. Two key advantages of cycling tests compared to step or treadmill tests are that they are weight supported, thereby reducing stress and strain on joints, and that they do not require dynamic balance. These are important considerations, as the physical manifestations of chemotherapy treatment induce a number of side effects which may compromise a patient's ability to safely perform maximal or submaximal exercise testing requiring dynamic balance, muscle control, muscle strength, and proprioception. Furthermore, side effects as a result of chemotherapy and associated pharmacological agents used to treat LHC can also result in muscle weakness, numbness in limbs and paresthesia (Michael and Stubblefield, 2009), which can impact a patient's ability to safely perform treadmill exercise, with stationary cycling representing a well-tolerated and safer option. Additionally, cycle ergometers are generally less expensive and more portable than treadmills, allowing for utilisation in a clinical or hospital setting.

A submaximal cycle test that may be appropriate for a cancer population is the Aerobic Power Index (API), a component of the Tri-Level Fitness Profile developed by Telford and colleagues (Telford et al., 1987). The advantages of the API over other submaximal cycling tests include its short duration (~2-5 min depending on fitness levels), use of set work loads (which prevents the need to predict workloads in order to achieve a target heart rate), the use of a standard predetermined attainable heart rate (75% HR_{max}: (220 – age) X 0.75), and the availability of established normative values. In addition to participant or pathology considerations, the reliability of the API test is crucial both in research and clinical work. To date, the API has been shown to be reliable in a number of clinical populations which share some common

Received: 26 July 2012 / Accepted: 17 September 2012 / Published (online): 01 December 2012

factors with cancer including chronic fatigue syndrome (intra-class correlation coefficient: ICC = 0.97: Wallman et al., 2003b), an obese population (ICC = 0.95: Wallman and Campbell, 2007) and a sedentary population (ICC = 0.98: Wallman et al., 2003a). Whilst there are some common clinical characteristics between these populations and a LHC population, it is important for research purposes that the reliability and applicability of this test is established. Therefore, the main purpose of this study was to determine the test-retest reliability of the submaximal API test in a LHC population. Additionally, we aimed to determine the reliability of oxygen uptake (VO_2: $ml \cdot kg^{-1} \cdot min^{-1}$) and ratings of perceived exertion (RPE: Borg, 1982) during the API.

Methods

Twenty participants (11 males and 9 females) aged between 18 and 70 y were recruited from metropolitan hospitals in Perth, Western Australia, as part of a larger study involving LHC patients being conducted collaboratively between the University of Western Australia and Solaris-Care Foundation. All participants had been diagnosed with lymphoma or myeloma and had completed chemotherapy treatment within the previous 4 weeks. All participants provided informed consent prior to participation and the University of Western Australia, Sir Charles Gairdner Hospital (WA) and other metropolitan hospitals granted ethics approval. Participants were asked to attend two testing sessions that were held at the same time of day, one week apart and were instructed to follow the same daily routine with regards to diet and incidental activity and not to participate in any exercise in the 24 h prior to testing.

Prior to testing participants completed the revised physical activity readiness questionnaire (r-PAR-Q: Thomas et al., 1992) and were asked about current medications that may interfere with heart rate during testing. During the testing sessions, participants initially performed the API and following the first testing session completed the International Physical Activity Questionnaire (IPAQ)(Craig et al., 2003), the Symptom Distress Score questionnaire (Joske, 2004), and provided medical characteristics information. The IPAQ was used in order to determine current activity levels, while the Symptom Distress Score is a questionnaire that assesses cancer patients' level of distress across seven dimensions on a scale from 'not at all' at 0 to the 'worst possible' at 10. Both questionnaires have been shown to be reliable in a LHC population (Craig et al., 2003; Joske, 2004). This data was collected in order to provide a comprehensive 'snapshot' of participants' abilities in respect to their health and aerobic conditioning.

Upon arrival at the School of Sport Science, Exercise & Health physiology laboratory, all participants were weighed using Sauter scales (August Sauter GmBH, D-7470 Albstadt 1 Ebingen, West Germany). In addition, height was measured (stadiometer) and target heart rate (THR: Telford et al., 1989) was calculated, based on the formula: (220-age) x 0.75. The same exercise physiologist attended both sessions, and the temperature in the lab was kept consistent between 22°C and 23°C. During both testing sessions, participants wore Polar heart rate (HR) monitors (Polar Electro Oy, Kempele, Finland). During the API, participants were seated on a front access (back wheel only design) Exertech Ex-10 cycle ergometer (Repco Cycle Company, Huntingdale, Victoria, Australia), with the seat positioned so that their knee was slightly flexed when the foot was placed on the pedal at its lowest point. Individual seat positions were recorded and replicated each testing session. Participants were positioned so that they could easily sight an attached computer that displayed watts (W) achieved during cycling. Resting HR was recorded and participants were familiarised with the testing protocol, which included the use of respiratory breathing and gas analysis equipment, as well as the reporting of their rating of perceived exertion (RPE) associated with the exercise. Participants practiced breathing into the gas analyses equipment for 1 min in order to familiarise themselves with this form of testing. Oxygen consumption was measured throughout the test by a metabolic cart consisting of a computerised on-line system. Inspired air was measured by a Morgan Ventilometer Mark II 225A (P.K Morgan, UK), while expired air was continuously sampled and recorded every 15 s by Applied Electrochemistry S-3A O_2 and CD-3A CO_2 anaylisers (Pittsburgh, Pennsylvania, USA). The O_2 and CO_2 sensors were calibrated prior to and after each test using reference gases that had been gravimetrically determined on a previous occasion. The Morgan ventilator was calibrated prior to and after each test according to the manufacturer's instructions. Additionally, after each exercise test was completed, the reference gases were continuously sampled by the O_2 and CO_2 sensors for a period of 2 min. The last minute of data was averaged and assessed for analyser drift, with no ventilatory drift occurring during testing.

Participants then commenced the API cycle test at a very low staring point of 25 W for 1 min, with the power output increasing by a further 25 W every subsequent minute until the participant reached their THR or choses to stop the test prematurely. The test is terminated at the end of the minute that THR is reached. Additionally, RPE was recorded 55 s into each min, while HR was recorded at the end of each min.

The power output (W) that occurred when THR was reached was determined through previously used interpolation techniques (Telford et al., 1989). This was performed by calculating the difference between the HR recorded at the second last workload and the THR, as well as the difference in HR scores recorded for the last two workloads. The two outcomes were then represented as a fraction and applied to the workload increment of 25 W. The result was added to the power output achieved during the second last workload, and then divided by the participant's body mass in kilograms ($W \cdot kg^{-1}$). This same interpolation method was applied to VO_2 and RPE data in order to determine the equivalent of these values at the THR.

As the Ex-10 cycle ergometer is an air-braked system that depends on air resistance for the absorption of energy, corrections were made to the $W \cdot kg^{-1}$ value in

Table 1. Descriptive statistics of test-retest participants.

	Male (n = 9)	Female (n = 9)	Total (n = 18)
Symptom Distress Score, mean(SD)	6.54 (8.50)	8.33 (5.89)	7.83 (7.36)
Disease Type			
Hodgkin's Lymphoma	0	2	2
Non-Hodgkin's Lymphoma	9	6	15
Myeloma	0	1	1
Incidence of comorbid pathologies			
Affecting lower limb & trunk	7	1	8
Affecting upper limb	1	3	4
Excess body weight *	5	5	10
Activity Classification Level #			
Low	0	3	3
Moderate	5	5	10
High	4	1	5

According to International Physical Activity Questionnaire with 7 day recall
* Based on body mass index

order to account for changes in daily temperature and atmospheric pressure. The correction factor is represented by the formula: P / 760 x 295 / (273 + T), where P is atmospheric pressure in mmHg, and T is room temperature in C°.

Statistical analysis was carried out under the guidance of a statistician and involved a number of methods including paired t-tests to determine significant differences between trials for all endpoint data, Pearson's correlations, absolute and relative technical error of measurements (TEM) (Knapp, 1992), intra-class correlations (ICC)(Vincent, 1995) and limits of agreement (Atkinson and Nevill, 1998; Bland and Altman, 1986), which have been detailed in an earlier publication (Wallman and Campbell, 2007). Classification of reliability for physiological measures followed the guidelines proposed by Vincent, with ICC scores above 0.90 categorised at highly reliable, values between 0.80 and 0.89 considered as moderately reliable, and values below 0.80 categorised to be of questionable reliability (Vincent, 1995). Correlation values were interpreted following guidelines established by Cohen (Cohen,1988). Data analysis was performed using Excel (Microsoft Office, 2011 edition) and Statistical Package for the Social Sciences, version 19 (SPSS, IBM inc., US)(Vincent, 1995).

Results

Of the twenty recruited participants, two were excluded from analysis due to disease relapse or complications unrelated to testing occurring within the month following the initial test. The eighteen remaining participants (9 males and 9 females) ranged in age from 27-68 y (mean ± SD: age 53.00 ± 11.00 y, height 1.72 ± 0.07 m, body mass 80.22 ± 7.03 kg, body mass index 25.82 ± 1.62)(Table 1). The majority of participants were diagnosed with Non-Hodgkin's lymphoma (n = 15), in addition to Hodgkin's lymphoma (n = 2) and myeloma (n = 1)(Table 1). No participants exhibited co-morbidities that would limit testing or were taking medications known to influence heart rate. Individual power output results recorded during the trials were compared to established normative tables which account for age and gender, with overall rankings ranging from 1 (poor) to 4 (above average) out of 6 categories and averaged 2.22 (fair) for both trials (Telford et al., 1987; 1989). This score highlights below average fitness levels for patients post treatment. The mean time for the tests was 4.36 min (SD = 1.15). Levels of symptom distress, physical activity and pathologies incidence are highlighted in Table 1. Symptom distress showed no significant change following completion of the test (p = 0.18), with a mean score of 8.17 and 7.5, respectively for the two sessions.

Intra-class correlation (ICC) analysis indicated high reliability for power output and VO_2 and moderate reliability for RPE (Vincent, 1995). Absolute technical error of measurement and relative technical error of measurement results are reported in Table 2 along with average results for $W \cdot kg^{-1}$, VO_2, and RPE.

Pearson product-moment correlation coefficient scores indicated a strong relationship between power

Table 2. Additional test-retest results for power output, oxygen uptake (VO_2), and Ratings of Perceived Exertion (RPE) during the Aerobic Power Index submaximal aerobic exercise test (n = 18).

	Power output ($W \cdot kg^{-1}$)		VO_2 ($ml \cdot kg^{-1} \cdot min^{-1}$)		RPE	
	Test 1	Test 2	Test 1	Test 2	Test 1	Test 2
Mean (± SD)	1.31 (.43)	1.32 (.41)	14.19 (4.00)	14.69 (3.82)	12.39 (2.07)	12.40 (2.37)
Confidence intervals						
Lower bound	-.92		-1.25		-.85	
Upper bound	.06		.24		.81	
Paired *t*-test (two-tailed)	p = .75		p = .18		p = .96	
TEM	.11		1.18		1.0	
Relative TEM (%)	8.65%		7.37%		9.57%	
Intra-class Correlation	0.96		0.96		0.83	
LOG (bias ± random error)	-.01 ± .32		-.02 ± 3.29		-.50 ± 2.95	

TEM: technical error of measurement, LOG: Limits of agreement

output ($W \cdot kg^{-1}$) and VO_2 ($ml \cdot kg^{-1} \cdot min^{-1}$) for trial 1 (r = 0.93) and trial 2 (r = 0.89). However, RPE was moderately positively correlated with power output ($W \cdot kg^{-1}$) and VO_2 ($ml \cdot kg^{-1} \cdot min^{-1}$) in trial 1(r = 0.36 and r = 0.35 respectively) showing a tendency towards a strong positive relationship in trial 2 (r = 0.58 and r = 0.45).

Discussion

A significant side effect for LHC patients post chemotherapy is persistent mental and physical fatigue, with long duration or onerous testing protocols presenting a significant burden for patients. Currently, there are numerous submaximal cycling or walking exercise tests that have and continue to be used clinically and in research, however the API is unique in its short duration and relative ease of administration.

Due to the multifactorial nature of cancer and its treatment side effects, both research and clinical assessment often requires the testing of multiple factors during one session. Well documented side effects such as cognitive impairments (Robert and Raffa, 2010) ('chemo brain' or 'chemo fog') and cancer related fatigue (Curt et al., 2000) may significantly impair patients' abilities to engage in and concentrate during long testing sessions and may compromise the validity of their results. The average RPE values during the trials in the current study were 12.39 and 12.40 respectively, indicating a perceived exertion just below 'somewhat hard'. This result confirms that participants were required to exert some effort, yet did not consider the protocol to be overly demanding. Additionally, all participants were able to achieve their THR, and returned for the second testing session.

Despite a number of participants reporting musculoskeletal and/or neuromuscular pathologies prior to the testing, all were able to successfully complete the stationary bike API without limitation or exacerbation of symptoms, with this evidenced by no significant difference in symptom distress scores pre and post testing. This, combined with the final average RPE scores noted above, suggests that the API protocol was well tolerated. Finally, the mean time of the test was 4.36 min, confirming the short duration of the protocol. The results are not confounded by the varied diagnoses as analysis protocols conducted were within individuals.

The selection of an exercise test to be used to assess cardiovascular fitness in a deconditioned or diseased state, as well as to monitor fitness over time presents a number of challenges. It must allow for ease of administration, tolerance at a wide variety of fitness levels and be reliable in respect to results. Importantly, the API has established normative values that account for gender and age, with outcomes ranging from 'poor' to 'elite', allowing for the evaluation of fitness along an extensive continuum. This is of particular importance in a LHC population, as the varying diseases do not have an established lifestyle cause, so patients may exhibit varying levels of fitness and activity prior to diagnosis. The IPAQ results of the study population showed the spread of physical activity levels with the majority of participants (55%) having moderate levels of activity.

While the original published protocol for the API does not include assessment of oxygen uptake this information can provide additional information to researchers and clinicians. The correlation of end-point results for power output and VO_{2peak} demonstrated a strong relationship between these two variables during both trials, suggesting that the API may be used to predict a participant's maximal oxygen uptake. Future research should consider assessing this aspect.

Conclusion

Test-retest of the API submaximal exercise test in a cancer population demonstrated high reliability for the main outcome measure of power output ($W \cdot kg^{-1}$), while associated variables relating to RPE and VO_2 were also found to be reliable. All participants were able to reach and complete the given workload for their THR and none reported any limitations during testing or exacerbation of symptoms. When comparing submaximal exercise tests, the API test represents a simple protocol that is easy to administer, is relatively short in duration, requires no need for estimation of workloads, is weight assisted and has normative values. As the test can be performed without the assessment of ventilatory variables it also requires minimal specialised equipment, making it ideally suited to both a clinical and research setting. These factors combine to make the API a highly valid and clinically applicable test of cardiorespiratory fitness for cancer patients during and post treatment.

Acknowledgements

This study was funded in part by a research grant from the SolarisCare Foundation, Western Australia.

References

Atkinson, G. and Nevill, A.M. (1998) Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. *Sports Medicine* **26(4),** 217-238.

Australian Institute of Health & Welfare. (2008a) *Cancer in Australia: an overview 2008,* Cancer series no. 46. Cat. no. CAN 42. Canberra: AIHW. Available from URL: http://www.aihw.gov.au/publication-detail/?id=6442468196.

Australian Institute of Health & Welfare. (2008b) *Cancer survival and prevalence in Australia: cancers diagnosed from 1982 to 2004,* Cat. no. CAN 38. Canberra: AIHW. Available from URL: http://www.aihw.gov.au/publication-detail/?id=6442468141.

Borg, G. (1982) 'Psychophysical basis of perceived exertion', *Medicine and Science in Sport and Exercise* **14(5),** 377-381.

Bland, J.M. and Altman, D.G. (1986) Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* **8,** 307-310.

Borg, G. (1982) Psychophysical basis of perceived exertion. *Medicine and Science in Sport and Exercise* **14(5),** 377-381.

Cohen, J.W. (1988) *Statistical power analysis for the behavioral sciences.* 2nd edition. Lawrence Erlbaum Associates, New Jersey.

Courneya, K. (2003) Exercise in cancer survivors: an overview of research., *Medicine and Science in Sports and Exercise* **35(11),** 1846-1852.

Craig, C., Marshall, A., Sjostrom, M., Bauman, A., Booth, M., Ainsworth, B., Pratt, M., Ekelund, U., Yngve, A., Sallis J. and Oja P. (2003) International physical activity questionnaire: 12-country reliability and validity. *Medicine and Science in Sports and Exercise* **35(8),** 1381-1395.

Curt, G., Bretibart, W., Cella, D., Groopman, J., Horning, S., Itri, L.,

Johnson, D., Miaskowski, C., Scherr, S., Portenoy, R. and Vogelzang, N. (2000) Impact of cancer related fatigue on the lives of patients: new findings from the fatigue coalition. *Oncologist* **5,** 353-360.

Joske, D. (2004) *Brownes Cancer Support Centre Patient Care Report.* Solariscare Foundation. Available from URL http://www.reikiaustralia.com.au/Resources/ReikiAustralia/PDFs/brownes_cancer_support_centre_report_090305.pdf

Knapp, T. (1992) Technical error of measurement: a methodological critique. *American Journal of Physical Anthropology* **87,** 235-236.

Michael, S. and Stubblefield, M. (2009) *Cancer Rehabilitation Principles and Practice.* Demos Medical Publishing, New York.

Robert, B. and Raffa, R. (2010) *Chemo Fog.* Springer, Dordrecht.

Telford, R., Minikin, B. and Hahn, A. (1989) A simple method for the assessment of general fitness: the tri-level profile. *Australian Journal of Science and Medicine in Sport* **21,** 6-9.

Telford, R., Minikin, B. and Hooper, L. (1987) The tri-level fitness profile. *EXCEL* **4,** 11-13.

Thomas, S., Reading, J. and Shephard, R. (1992) Revision of the physical activity readiness questionnaire (PAR-Q). *Canadian Journal of Sport Science* **17(4),** 338-345.

Vincent, W. (1995) *Statistics in Kinesiology,* Human Kinetics, Champaign, Illinois.

Wallman, K. and Campbell, L. (2007) Test-retest reliability of the aerobic power index submaximal exercise test in an obese population. *Journal of Science and Medicine in Sport* **10,** 141-146.

Wallman, K., Goodman, C., Morton A., Grove, R. and Dawson, B. (2003a) Test-retest reliability of the aerobic power index component of the tri-level fitness profile in a sedentary population. *Journal of Science and Medicine in Sport* 6(4), 443-454.

Wallman, K., Goodman, C., Morton A., Grove, R. and Dawson, B. (2003b) Test-retest reliability of the Aerobic Power Index test in patients with chronic fatigue syndrome. *Journal of Chronic Fatigue Syndrome* **11(4),** 19-32.

Key points

- The API test is a highly reliable protocol for use within a haematological cancer population.
- The API test of cardiovascular fitness can be considered a clinically feasible, safe and tolerable exercise test in cancer patients.
- Intra-class correlation coefficient (ICC) scores for power output ($W \cdot kg^{-1}$) and oxygen uptake ($ml \cdot kg^{-1} \cdot min^{-1}$) were highly reliable and a correlations demonstrated a strong relationship between power output ($W \cdot kg^{-1}$) and oxygen uptake ($ml \cdot kg^{-1} \cdot min^{-1}$).

AUTHORS BIOGRAPHY

Bonnie J. FURZER
Employment
PhD Candidate at the University of Western Australia. Accredited Exercise Physiologist – Clinic Director UWA Health & Rehabilitation Clinic.
Degrees
Bachelor of Arts, Bachelor of Science
Research interests
Chronic disease and injury, implementing multi-faceted exercise programs incorporating aerobic exercise, resistance training and functional training.
E-mail: bonnie.furzer@uwa.edu.au

Karen E. WALLMAN
Employment
Associate Professor, University of Western Australia
Degree
PhD
Research interests
Elite athletes, chronic illnesses, exercise and cancer, exercise and obesity and fatigue syndrome.
E-mail: karen.wallman@uwa.edu.au

Timothy R. ACKLAND
Employment
University of Western Australia - Head of School (Sport Science, Exercise & Health) – Director of UWA Active Life Institute
Degree
PhD
Research interests
The mechanics of human movement with themes spanning exercise rehabilitation, high performance sport and human performance in industry.
E-mail: tim.ackland@uwa.edu.au

David J. L. JOSKE
Employment
SolarisCare Foundation – Founder & Chairman, University of Western Australia – Clinical and Adjunct Staff (Medicine & Pharmacology). Sir Charles Gairdner Hospital - Haematologist
Degree
MD
Research interests
Cancer related illness.
E-mail: david.joske@health.wa.gov.au

✉ **Bonnie J. Furzer**
School of Sport Science, Exercise & Health, University of Western Australia, 35 Stirling Hwy, Crawley, Western Australia 6009, Australia