NIH-PA Author Manuscript
NIH-PA Author Manuscript

Published in final edited form as:
Int J Neurosci. 2011 August ; 121(8): 450–456. doi:10.3109/00207454.2011.574762.

Reliability of Peak Treadmill Exercise Tests in Mild Alzheimer Disease

Heather S. Anderson, MD[1], **Patricia M. Kluding, PT, PhD**[2], **Byron J. Gajewski, PhD**[3], **Joseph E. Donnelly, EdD**[4], and **Jeffrey M. Burns, MD**[1]

[1] University of Kansas School of Medicine, Kansas City, KS, USA

[2] University of Kansas School of Allied Health, Kansas City, KS, USA

[3] University of Kansas Department of Biostatistics, Kansas City, KS, USA

[4] University of Kansas, Lawrence, KS, USA

The prevalence of Alzheimer disease (AD) doubles every 5 years after the age of 65, reaching nearly 50% after age 85 (Evans et al., 1989). This, along with an unprecedented growth in the elderly population, is leading to dramatic increases in the incidence of AD. Thus, effective strategies for promoting healthy brain aging and preventing AD are increasingly important. One strategy that appears promising in promoting healthy brain aging is exercise and physical activity. Evidence is accumulating that endurance exercise is beneficial to brain health (Laurin, Verreault, Lindsay, MacPherson, & Rockwood, 2001), and increased cardiorespiratory fitness is associated with increased brain volume in subjects with very mild to mild AD (Burns et al., 2008).

While enhancing cardiorespiratory fitness may be a strategy for preventing cognitive decline in AD, there is limited information available on the validity and reliability of cardiorespiratory fitness measures in this population. The gold standard measure of cardiorespiratory fitness is maximum oxygen consumption (VO_{2max}) (Frankin, 2001), the highest rate of oxygen uptake attainable during maximal or exhaustive exercise (American College of Sports Medicine, 2005). If the subject becomes exhausted and ends the test prior to reaching the physiologic VO_{2max}, the end of the test is called peak oxygen consumption (VO_{2peak}). It is unknown if advanced age and cognitive difficulties in people with AD would limit their ability to fully participate in a standard graded exercise test to reliably assess VO_{2max} or VO_{2peak}.

Treadmill exercise testing has been found to be reliable in subjects with traumatic brain injury and mental retardation, although these subjects were very young (Fernhall, Millar, Tymeson, & Burkett, 1990; Mossberg & Greene, 2005). Traumatic brain injury and mental retardation are different disease processes than AD and would be expected to result in static rather than progressive cognitive symptoms. With AD, memory is impaired as is the ability to follow commands, however patients in the earliest stages of AD would be expected to respond to prompting and reminders to follow testing procedures. To our knowledge, no studies have assessed the reliability of peak treadmill exercise testing in subjects with AD. In our previous research on patients with very mild to mild AD (Burns, et al., 2008), we have found them to be capable of ambulating on a treadmill and completing peak treadmill

Corresponding Author: Heather S. Anderson, MD, Assistant Professor, Department of Neurology, University of Kansas School of Medicine, 3901 Rainbow Blvd, MSN 1063, Kansas City, KS 66160, (913) 588-6970, (913) 945-5035 (fax), handerson3@kumc.edu.

Declaration of Interest

The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

exercise testing with 3 participants out of 74 (126 total peak exercise tests) identified as having EKG changes during testing. All 3 participants had negative follow-up testing in cardiology. The purpose of this study was to investigate the reliability of a graded peak treadmill exercise test in elderly people with early AD.

Method

Study Design

A sample of convenience was used to investigate test-retest reliability of VO_{2peak} assessment using a graded peak exercise test on a treadmill. The VO_{2peak} test was performed twice within a 14-day period.

Participants

Sixteen subjects (11 men, 5 women) enrolled in the University of Kansas Brain Aging Project participated in this study. This program, which is a longitudinal study of brain health in community dwelling older adults with and without dementia has been previously described (Burns et al., 2007). Subject qualifications for participating in the current study included 1) age 60 years or older, 2) diagnosis of Probable AD according to the National Institute of Neurological and Communicative Disorders and Stroke-Alzheimer Disease and Related Disorders Association (NINDS-ADRDA),(McKhann et al., 1984) and 3) Clinical Dementia Rating (CDR)(Hughes, Berg, Danziger, Coben, & Martin, 1982; Morris, 1993) of 0.5 (very mild) or 1 (mild dementia). The stringent clinical diagnostic criteria for AD utilized in this study has a diagnostic accuracy for AD of 93% (Berg et al., 1998). Subjects were excluded if they had myocardial infarction or symptoms of coronary artery disease (e.g., angina) in the last two-years or uncontrolled hypertension within the last 6 months. All subjects or their legally acceptable representative gave written, informed consent to participate in the study. The institutional review board of the University of Kansas Medical Center approved all procedures.

Exercise Test Protocol

Each subject completed two identical symptom-limited incremental exercise tests to voluntary fatigue over a 14-day period in the presence of a physician, an exercise physiologist, a research assistant, and a nurse. The instrumentation and procedure for treadmill testing described by Mossberg (Mossberg & Greene, 2005) was modified to include use of the Trackmaster TMX425 treadmill (Full Vision, Inc, Newton KS), 12-lead electrocardiograph system (Cardio-Card, Nasiff Associates, Brewerton, NY), and 2-way valve mouthpiece (Hans Rudolph, Inc., Shawnee, KS) attached to an automated metabolic measurement cart (TrueOne 2400, ParvoMedics, Sandy, UT) where respiratory gas exchange data for oxygen consumption (VO_2) and RER were averaged over 15-second intervals. Prior to beginning peak exercise testing, subjects and their spouses were administered a modified Physical Activity Readiness Questionnaire (PAR-Q) (S. Thomas, Reading, & Shephard, 1992) to assess possible safety concerns before beginning the graded exercise testing. Resting data were collected for at least 30 seconds with the patient standing on the treadmill to allow CO_2 levels to stabilize and to familiarize the subject with the mouthpiece, then the patient began walking at a pace of 1.7 mph at 0% grade. The Modified Bruce protocol was used for the peak exercise test, (Okin, Ameisen, & Kligfield, 1986) and the grade was increased by 5% every 2 minutes up to 10% grade at which point speed was increased by 0.4 mph and grade by 1% every 2 minutes. Borg Rating of Perceived Exertion (RPE) and blood pressure were recorded at the end of each stage.

American College of Sports Medicine (ACSM) guidelines (American College of Sports Medicine, 2005) were used to determine whether the exercise test should be terminated early and included ST-segment depression of more than 2mm, increasing nervous system symptoms (ie, ataxia, dizziness), sustained ventricular tachycardia, and chest discomfort. Exercise was continued until (1) objective signs required termination of the test (e.g., electrocardiogram abnormality, (2) the cardiorespiratory responses failed to increase normally with increasing workload or (3) patients could not continue to exercise at the prescribed workload despite encouragement to do so.

Patients were verbally encouraged to give their maximal effort as they approached the end of the test. On termination of exercise, patients continued walking at a speed less than or equal to 1.7 mph and 0% grade until blood pressure and heart rate returned to near pre-exercise values. After the testing was completed, a blinded exercise physiologist reviewed the metabolic cart readings and made the determination for VO_{2peak}, the highest VO_2 level recorded within the last 60 seconds of testing. RER and peak heart rate values were taken from the time of the corresponding VO_{2peak}.

Tests and retests were generally performed in the morning between 2 and 14 days apart, with the two testing sessions performed at approximately the same time of day (within one hour). The retest was performed by the same staff members and using the same treadmill as the initial test. Subjects (and their spouses) were instructed to undergo the same activities before the second test as they had undertaken before the first test to ensure as much pretest standardization as possible. This included physical activity, medication administration, and smoking. Subjects were encouraged to not initiate a new diet or exercise routine between the test and retest. Subjects fasted for at least 6 hours prior to the test and retest, but were allowed to hydrate with water ad libitum. Additionally, participants were asked to avoid caffeine and vigorous physical activity for at least 6 and 24 hours, respectively, prior to testing.

Data Analysis

Analyses were conducted using SPSS, version 16.0 (SPSS Inc., Chicago, IL). Descriptive statistics were performed on all measures. The frequency and percentage of participants who achieved the criteria for VO_{2max} were determined. Intraclass correlation coefficients (ICC [3,1]) were calculated to determine test-retest reliability, and reliability was calculated using a two-way random-effects model (the between-time variance is excluded from the denominator variance). ICC [2,1] (where the between-time variance is included in the denominator variance) was also calculated. Paired *t* tests were used to determine significant differences between trials with respect to VO_{2peak}, RER, test duration, and peak heart rate. Standard error of measurement (SEM), the average error (Hopkins, 2000) or mean deviation from the mean, complements ICCs in studies of reliability (Shrout, 1998) where ICCs provide information on relative reliability and SEM on absolute reliability. SEM was calculated as the standard deviation of the sample pooled between subjects × $\sqrt{(1\text{-ICC})}$. Continuous variables are summarized by means and standard deviations, whereas categorical variables are summarized by frequency and percent.

Although the purpose of this study was to determine the reliability of exercise testing in subjects with early AD, we assessed the participants' ability to reach maximum aerobic capacity (ie, VO_{2max}) during the graded peak exercise test. To determine if VO_{2max} was reached, participants had to meet three criteria including (1) a leveling off (plateau) or no change in VO_2 with increasing workload, (2) heart rate within 10 beats of age-predicted maximum heart rate (calculated as 208-(0.7 × age), (Tanaka, Monahan, & Seals, 2001) and (3) elevated respiratory exchange ratio (RER). The ACSM, 8th edition (ACSM, 2010) recommends RER > 1.1, however the criteria for establishing maximal exercise testing in older adults is inconsistent in the literature, ranging from greater than 1.0 (da Cunha-Filho et al., 2001) to 1.15 (Eng, Dawson, & Chu, 2004; Howley, Bassett, & Welch, 1995). We therefore applied an RER cut-off value of >1.0 as a criterion for determining VO_{2max}. β-blockers may reduce heart rate by 25% to 30% (Tesch, 1985), so as in Tang, et al,(Tang, Sibley, Thomas, McIlroy, & Brooks, 2006) for the subjects on β-blockers, the age-predicted maximal heart rate was adjusted according to the equation (0.7 × (220-age)).

Results

Patients' physical characteristics and severity of dementia as represented by the Clinical Dementia Rating Scale (CDR) are summarized in Table 1. CDR scores were determined an average of 5.7 months prior to entering the study, and the dementia severity is expected to have declined very little over the 5.7 months (Aisen et al., 2008; Rogers, Farlow, Doody, Mohs, & Friedhoff, 1998; Tariot et al., 2001). The graded treadmill test and re-test were performed an average of 3.2 days apart (range: 2 to 7 days).

Sixteen subjects were contacted to participate in the study, and all 16 individuals met inclusion criteria and agreed to participate in the study. All 16 participants felt comfortable walking on the treadmill, and none of the participants required the use of the familiarization session. No additional personnel were required beyond the standard testing personnel, and no additional safety precautions were necessary. Two subjects (12.5%) did not complete the first of the two treadmill exercise tests. One woman (CDR 0.5, age 73 years) was withdrawn due to excessive sweating and possible ST segment depression, however evaluation during a formal cardiology consult revealed a normal exercise EKG. The other woman (CDR 1, age 78 years) was withdrawn due to ST segment depression.

Fourteen (87.5%) participants completed both graded peak exercise tests without incident. Of these, exercise testing was terminated by all subjects due to subjective report of fatigue. Three of the subjects (21%) reached the predicted maximal heart rate, and 5 additional subjects (total of 57%) reached a peak heart rate within 10 beats per minute of the age-predicted maximal heart rate. With respect to the other 2 criteria for VO_{2max}, 14 (100%) participants showed a plateau in VO_2 at the end of the initial test and the retest, and 13 (93%) reached an RER value greater than 1.0 (1 participant reached an RER of 1.0). Data from these 14 subjects (11 men, 3 women) were included in the reliability analysis.

Peak Cardiorespiratory Responses During Incremental Exercise

As noted above, two of the patients enrolled had signs and symptoms that required termination of the first test. The remaining 14 subjects continued to exercise to voluntary fatigue. The descriptive characteristics of the exercise testing and reliability results are presented in Table 2. Figure 1 illustrates the relationship between test and retest for VO_{2peak}.

Reliability of the Peak Cardiorespiratory Responses

Reliability was excellent with total VO_{2peak} (mL/kg/min) highly correlated across the two tests (r=0.94, p<0.001) with an intraclass correlation coefficient (ICC [3,1]) of 0.92. Other measures such as exercise duration (r=0.93, p<0.001; ICC [3,1] =0.92) and maximum heart rate (r=0.87, p<0.001; ICC [3,1] = 0.87) had excellent reliability. Low ICC values were noted for RER, possibly because of the low variability in the measure. There was a statistically significant difference between test and retest for exercise duration. The SEM data are presented in Table 2. Noted is the point estimates for ICC [2,1] were very similar to ICC [3,1].

Discussion

To our knowledge, the reliability of peak exercise testing has not been previously reported in patients with mild AD. Our data demonstrate that this testing procedure is reliable in this population. Although two subjects (12.5%) were not able to complete the test because of cardiac concerns, our previous experience with exercise testing in subjects in the earliest stages of AD (Burns, et al., 2008) has shown a very low incidence of cardiac concerns (2%) during peak treadmill exercise testing.

Although the ACSM criteria for a maximal exercise test include plateau in VO_2, elevated RER, and peak heart rate within 10 beats per minute of age-predicted maximal heart rate (ACSM, 2010), these criteria are not specific to older adults or those individuals with dementia. Because peak heart rate can be influenced by the use of β-blockers and determination of plateau in VO_2 is subjective, we feel that the most important physiologic criteria is RER. Among the subjects who did complete our exercise protocol, 11 (78.6%) met all three of our criteria for maximal effort as described in Data Analysis above (plateau in VO2, RER > 1.0, and peak heart rate within 10 beats per minute of age-predicted maximal heart rate), and 13 (92.9%) achieved RER > 1.0. The fourteenth subject achieved RER = 1.0. In a study of maximal exercise testing in subacute stroke patients, only 2.9% reached criteria for maximal exercise testing using criteria which included reaching an RER >1.0 (Tang, et al., 2006). Information regarding medical comorbidities such as cardiovascular or pulmonary conditions was not provided for the subacute stroke patients, however the subjects enrolled in our study are generally healthy with no diabetes, heart disease, or stroke. Perhaps underlying medical conditions along with limb fatigue from hemiparesis following stroke in Tang et al (Tang, et al., 2006) influenced their capacity to reach maximal exercise testing. Four of our patients were taking β-blockers during the testing period, and the age-predicted heart rate for these subjects was adjusted to 70%. Using this adjustment, three of these four subjects reached predicted heart rates. The average duration of our test was within the recommended guidelines for graded maximal exercise testing of 8–12 minutes (ACSM, 2010).

Reliability studies help researchers understand the measurement error "inherent in any diagnosis, score, or measurement (Kottner et al., 2011)." ICC's are "for measuring the reliability of continuous scales (Kottner, et al., 2011)." Low reliability will result in an unacceptable amount of noise and can affect the power to detect differences in clinical trials or cause bias when using such variables as predictors (e.g. in a regression analysis). We found the variables exercise duration, heart rate, and VO_2 (ml/kg/min) to have moderate reliability, VO_2 (l/min) has substantial reliability, and RER is inconclusive, but lower than VO_2 (l/min). More work is needed to refine the understanding of the absolute reliability of RER. Our results were identical to the reliability of maximal exercise testing in individuals with traumatic brain injury (Mossberg & Greene, 2005) and comparable to those in individuals with mental retardation (Fernhall, et al., 1990) with respect to ICCs for VO_2 (ml/kg/min), and our results showed higher levels of reliability compared to a study in older women (Foster, Hume, Dickinson, Chatfield, & Byrnes, 1986) and in older men (S. G. Thomas, Cunningham, Rechnitzer, Donner, & Howard, 1987) with respect to Pearson correlations for VO_{2peak} (ml/kg/min). Average VO_{2peak} (ml/kg/min) in our study was similar compared to a group of men (20.9 vs. 20.5) and women (17.0 vs. 16.6) (Paterson, Cunningham, Koval, & St Croix, 1999) age 75–79.

Although we found excellent consistency of most measures between the tests, we did find a statistically significant difference in duration between tests. Although duration is frequently reported as an outcome measure (Lyerly et al., 2008; Pollock et al., 1976; Sui et al., 2009) this is not a true measure of cardiorespiratory fitness and may reflect patient comfort with the test and willingness to continue beyond the point of VO_2 plateau. There was also an increase in performance from trial one to trial two which might be considered a limiting factor on the accuracy of the absolute reliability of the procedure, however there was no statistically significant change in VO_{2peak} between trial one and trial two. Although the mean test duration was within the recommended guidelines for graded maximal exercise testing of 8–12 minutes (ACSM, 2010), 10 of 28 tests were less than 8 minutes in duration. Therefore a different protocol with more gradual workload increases or a bicycle ergometer may be useful for future research studies.

A limitation to the study is the small sample size which limits the study's generalizability. There are other limitations to this study as well. Eleven men but only 3 women completed the test and retest however the numbers are too small to perform subgroup analysis to detect differences between men and women on results. Clinical Dementia Rating (CDR) scoring was performed an average of 5.7 months prior to the exercise tests. The CDR was not a primary outcome measure for this study and was only used to characterize participants as having early-stage Alzheimer disease. Dementia severity is expected to have declined very little over the 5.7 months (Aisen, et al., 2008; Rogers, et al., 1998; Tariot, et al., 2001). The sample was limited to early-stage AD and is not clearly generalizable to more advanced AD. Subjects with more advanced dementia may have greater variability in fatigue tolerance or may have more difficulty understanding directions for exercise testing. Finally, these individuals were cardiovascularly healthy and did not have significant orthopedic conditions limiting their ability to participate in fitness testing.

Nevertheless these data provides the first evidence that peak exercise testing on the treadmill is a reliable method of investigating cardiorespiratory fitness in individuals with early stage AD. Future work to explore physical fitness in patients with AD should utilize this testing protocol with confidence to identify changes as a result of an exercise intervention. This study is an important first step to establishing appropriate testing protocols for exercise testing in subjects with mild AD.

Acknowledgments

This study was supported by grants R03AG026374 and R21 AG029615 from the National Institutes of Aging and K23NS058252 from the National Institute on Neurological Disorders and Stroke and by support from the University of Kansas Endowment Association and the Fraternal Order of Eagles. The University of Kansas General Clinical Research Center (M01RR023940) provided essential space, expertise, and nursing support. This study was also supported by Hal Oppenheimer and the Oppenheimer Foundation.

Abbreviations

AD	Alzheimer disease
VO_{2max}	maximum oxygen consumption
VO_{2peak}	peak oxygen consumption
RER	respiratory exchange ratio
NINDS-ADRDA	National Institute of Neurological and Communicative Disorders and Stroke-Alzheimer Disease and Related Disorders Association
CDR	Clinical Dementia Rating
RPE	Rating of Perceived Exertion
PAR-Q	Physical Activity Readiness Questionnaire
VO_2	oxygen consumption
ACSM	American College of Sports Medicine
ICC	intraclass correlation coefficient
SEM	standard error of measurement

NIH-PA Author Manuscript

NIH-PA Author Manuscript

SD	standard deviation
CI	confidence interval

References

ACSM, A. C. o. S. M. ACSM's guidelines for exercise testing and prescription. 8. Philladelphia: Lippincott Williams & Wilkins; 2010.

Aisen PS, Schneider LS, Sano M, Diaz-Arrastia R, van Dyck CH, Weiner MF, et al. High-dose B vitamin supplementation and cognitive decline in Alzheimer disease: a randomized controlled trial. JAMA. 2008; 300(15):1774–1783. [PubMed: 18854539]

American College of Sports Medicine. ACSM's Guidelines for Exercise Testing and Prescription. 7. Lippincott Williams & Wilkins; 2005.

Berg L, McKeel DW Jr, Miller JP, Storandt M, Rubin EH, Morris JC, et al. Clinicopathologic Studies in Cognitively Healthy Aging and Alzheimer Disease: Relation of Histologic Markers to Dementia Severity, Age, Sex, and Apolipoprotein E Genotype. Archives of Neurology. 1998; 55(3):326–335. [PubMed: 9520006]

Burns JM, Cronk BB, Anderson HS, Donnelly JE, Thomas GP, Harsha A, et al. Cardiorespiratory fitness and brain atrophy in early Alzheimer disease. Neurology. 2008; 71(3):210–216. [PubMed: 18625967]

Burns JM, Donnelly JE, Anderson HS, Mayo MS, Spencer-Gardner L, Thomas G, et al. Peripheral insulin and brain structure in early Alzheimer disease. Neurology. 2007; 69(11):1094–1104. [PubMed: 17846409]

da Cunha-Filho I, Lim P, HQ, Henson H, Monga T, Protas E. A comparison of regular rehabilitation and regular rehabilitation with supported treadmill ambulation training for acute stroke patients. Journal of Rehabilitation Research and Development. 2001; 38:245–255. [PubMed: 11392657]

Eng JJ, Dawson AS, Chu KS. Submaximal exercise in persons with stroke: test-retest reliability and concurrent validity with maximal oxygen consumption. Archives of Physical Medicine and Rehabilitation. 2004; 85(1):113–118. [PubMed: 14970978]

Evans DA, Funkenstein HH, Albert MS, Scherr PA, Cook NR, Chown MJ, et al. Prevalence of Alzheimer's disease in a community population of older persons: Higher than previously reported. Journal of the American Medical Association. 1989; 262:2551–2556. [PubMed: 2810583]

Fernhall B, Millar AL, Tymeson GT, Burkett LN. Maximal exercise testing of mentally retarded adolescents and adults: reliability study. Archives of Physical Medicine and Rehabilitation. 1990; 71(13):1065–1068. [PubMed: 2256807]

Foster VL, Hume GJ, Dickinson AL, Chatfield SJ, Byrnes WC. The reproducibility of VO2max, ventilatory, and lactate thresholds in elderly women. Medicine & Science in Sports & Exercise. 1986; 18(4):425–430. [PubMed: 3747803]

Frankin, B. Normal cardiorespiratory responses to acute aerobic exercise. In: RJL, editor. ASCM's resource manual for guidelines for exercise testing and prescription. 4. Philadelphia: Lippincott Williams & Wilkins; 2001. p. 141-149.

Hopkins WG. Measures of reliability in sports medicine and science. Sports Medicine. 2000; 30(1):1–15. [PubMed: 10907753]

Howley ET, Bassett DR Jr, Welch HG. Criteria for maximal oxygen uptake: review and commentary. Med Sci Sports Exerc. 1995; 27(9):1292–1301. [PubMed: 8531628]

Hughes CP, Berg L, Danziger WL, Coben LA, Martin RL. A new clinical scale for the staging of dementia. British Journal of Psychiatry. 1982; 140:566–572. [PubMed: 7104545]

Kottner J, Audigé L, Brorson S, Donner A, Gajewski BJ, Hróbjartsson A, et al. Proposed Guidelines for Reporting Reliability and Agreement Studies (GRRAS). Journal of Clinical Epidemiology. 2011; 64(1):96–106. [PubMed: 21130355]

Laurin D, Verreault R, Lindsay J, MacPherson K, Rockwood K. Physical activity and risk of cognitive impairment and dementia in elderly persons. Archives of Neurology. 2001; 58(3):498–504. [PubMed: 11255456]

Lyerly GW, Sui X, Church TS, Lavie CJ, Hand GA, Blair SN. Maximal exercise electrocardiography responses and coronary heart disease mortality among men with diabetes mellitus. Circulation. 2008; 117(21):2734–2742. [PubMed: 18490521]

McKhann G, Drachman D, Folstein M, Katzman R, Price D, Stadlan EM. Clinical diagnosis of Alzheimer's disease: Report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's disease. Neurology. 1984; 34:939–944. [PubMed: 6610841]

Morris JC. The Clinical Dementia Rating (CDR): current version and scoring rules. Neurology. 1993; 43(11):2412b–2414. [PubMed: 8232972]

Mossberg KA, Greene BP. Reliability of graded exercise testing after traumatic brain injury: submaximal and peak responses. [Research Support, Non-U.S. Gov't]. American Journal of Physical Medicine & Rehabilitation. 2005; 84(7):492–500. [PubMed: 15973085]

Okin PM, Ameisen O, Kligfield P. A modified treadmill exercise protocol for computer-assisted analysis of the ST segment/heart rate slope: methods and reproducibility. J Electrocardiol. 1986; 19(4):311–318. [PubMed: 3540174]

Paterson DH, Cunningham DA, Koval JJ, St Croix CM. Aerobic fitness in a population of independently living men and women aged 55–86 years. Med Sci Sports Exerc. 1999; 31(12): 1813–1820. [PubMed: 10613433]

Pollock ML, Bohannon RL, Cooper KH, Ayres JJ, Ward A, White SR, et al. A comparative analysis of four protocols for maximal treadmill stress testing. American Heart Journal. 1976; 92(1):39–46. [PubMed: 961576]

Rogers SL, Farlow MR, Doody RS, Mohs R, Friedhoff LT. A 24-week, double-blind, placebo-controlled trial of donepezil in patients with Alzheimer's disease. Donepezil Study Group. Neurology. 1998; 50(1):136–145. [PubMed: 9443470]

Shrout PE. Measurement reliability and agreement in psychiatry. Stat Methods Med Res. 1998; 7(3): 301–317. [PubMed: 9803527]

Sui X, Laditka JN, Church TS, Hardin JW, Chase N, Davis K, et al. Prospective study of cardiorespiratory fitness and depressive symptoms in women and men. Journal of Psychiatric Research. 2009; 43(5):546–552. [PubMed: 18845305]

Tanaka H, Monahan KD, Seals DR. Age-predicted maximal heart rate revisited. J Am Coll Cardiol. 2001; 37(1):153–156. [PubMed: 11153730]

Tang A, Sibley KM, Thomas SG, McIlroy WE, Brooks D. Maximal exercise test results in subacute stroke. Archives of Physical Medicine and Rehabilitation. 2006; 87(8):1100–1105. [PubMed: 16876556]

Tariot PN, Cummings JL, Katz IR, Mintzer J, Perdomo CA, Schwam EM, et al. A randomized, double-blind, placebo-controlled study of the efficacy and safety of donepezil in patients with Alzheimer's disease in the nursing home setting. Journal of the American Geriatric Society. 2001; 49(12):1590–1599.

Tesch PA. Exercise performance and beta-blockade. Sports Medicine. 1985; 2(6):389–412. [PubMed: 2866577]

Thomas S, Reading J, Shephard RJ. Revision of the Physical Activity Readiness Questionnaire (PAR-Q). Canadian Journal of Sport Science. 1992; 17(4):338–345.

Thomas SG, Cunningham DA, Rechnitzer PA, Donner AP, Howard JH. Protocols and reliability of maximal oxygen uptake in the elderly. Canadian Journal of Sport Science. 1987; 12:144–151.

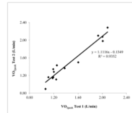

Figure 1.

Scatterplot showing relationship between test and retest peak exercise responses in patients with very mild to mild AD (N=14) measured in mL/kg/min.

Table 1

Descriptive characteristics of subjects with very mild to mild Alzheimer disease who completed peak treadmill test and retest (n=14)

Variable	Values	Range
Men/women	9/5	
Clinical Dementia Rating		0.5–1
CDR 0.5	9	
CDR 1	5	
Sum of Boxes	3.9 ± 1.5	2.0–7.0
Age at test, yrs	77 ± 4.8	70–84
β-blockers (%)	4 (29%)	
Height, cm	171.6 ± 9.7	157.0–188.0
Weight, kg	72.6 ± 9.4	63.4–95.9
Body mass index, kg/m^2	24.7 ± 3.1	20.0–29.5

NOTE. Values are n or mean ± standard deviation (SD) (range) unless noted otherwise.

NIH-PA Author Manuscript
NIH-PA Author Manuscript

TABLE 2

Peak responses for test and retest treadmill graded exercise tests

	Ex Duration minutes		HR beats/min		VO_2 mL/kg/min		VO_2 liters/min		RER VCO_2/VO_2	
Subject	Test	Retest	Test	Retest	Test	Retest	Test	Retest	Test	Retest
1	6.5	7.0	156	160	19.1	21.9	1.25	1.43	1.12	1.05
2	9.1	9.8	140	144	20.9	20.8	2.01	1.98	1.09	1.19
3	7.1	10.4	154	135	17.7	20.0	1.19	1.34	1.17	1.11
4	8.0	8.2	137	151	16.6	17.5	1.11	1.16	1.12	1.09
5	5.4	5.9	133	140	16.6	17.7	1.20	1.28	1.15	1.11
6*	6.9	6.8	123	121	15.9	15.9	1.38	1.36	1.07	1.05
7*	9.0	9.9	88	100	25.3	25.9	2.00	2.08	1.09	1.11
8	13.6	14.9	139	151	27.6	30.2	2.09	2.28	1.04	1.06
9	10.6	10.9	135	130	25.3	23.3	1.60	1.50	1.00	1.01
10	8.8	8.4	110	91	17.5	15.4	1.24	1.11	1.02	1.05
11	8.6	8.3	143	126	17.2	16.2	1.19	1.16	1.03	1.05
12*	5.3	6.1	94	98	16.8	14.0	1.06	0.89	1.13	1.10
13	11.6	14.0	137	138	26.1	28.2	1.93	2.10	1.04	1.06
14*	7.8	8.2	100	100	18.1	17.6	1.18	1.14	1.07	1.13
Mean	8.4[a]	9.2[a]	128	128	20.0	20.3	1.46	1.49	1.08	1.08
SD	2.3	2.7	21.6	22.4	4.2	5.0	0.38	0.44	0.05	0.05
ICC [3,1]	0.92		0.87		0.92		0.96		0.54	
95% CI	(0.77, 0.97)		(0.65,0.96)		(0.78,0.97)		(0.88,0.99)		(0.035, 0.825)	
ICC [2,1]	0.88		0.88		0.93		0.96		0.57	
95% CI	(0.56, 0.97)		(0.66,0.96)		(0.79,0.98)		(0.88,0.99)		(0.038,0.83)	
SEM	0.72		7.9		1.29		0.08		0.03	
95% CI[†]	(0.49, 1.06)		(5.38, 11.60)		(0.88, 1.89)		(0.06, 0.12)		(0.02, 0.05)	
%SEM of mean	8.2%		6.2%		6.4%		5.7%		3.1%	
95% CI	(5.6%, 12.0%)		(4.2%, 9.1%)		(4.3%, 9.4%)		(3.9%, 8.4%)		(2.1%, 4.5%)	

Ex Duration, total exercise duration; HR, heart rate; VO_2, oxygen consumption; VCO_2, carbon dioxide production; RER, respiratory exchange ratio; SD, standard deviation; ICC, single measure intraclass correlation coefficient; CI, confidence interval; SEM, standard error of measurement

[a] $p = 0.02$

[*] Subjects on β-blocker

[†] Using SPSS mixed command, we calculated the restricted maximum likelihood (REML) estimates for SEM and associated CI. Asymmetric intervals are due to chi-squared based intervals.