Maximal Fitness Testing in Sedentary Elderly at Substantial Risk of Disability:
LIFE-P Study Experience

Timothy S. Church, MD, M.P.H., Ph.D.[1]
Thomas M. Gill, MD[2]
Anne B. Newman, MD, MPH[3]
Steven N. Blair, PED[4]
Conrad P. Earnest, PhD[1]
Marco Pahor, MD[5]

[1]Pennington Biomedical Research Center
Louisiana State University System
Baton Rouge, LA

[2]Yale University School of Medicine, New Haven, CT

[3]University of Pittsburgh, Pittsburgh, PA

[4] Arnold School of Public Health
University of South Carolina
Columbia, SC

[5]University of Florida, Gainesville, FL

Address for correspondence and reprints:
Timothy S. Church
Pennington Biomedical Research Center
6400 Perkins Road
Baton Rouge, LA 70810

Email: tim.church@pbrc.edu

Short Title: Maximal Fitness Testing in sedentary Elderly

Abstract

Background

We sought to evaluate the acceptability and feasibility of maximal fitness testing in sedentary older individuals at risk of mobility disability.

Methods

Maximal cycle ergometer testing was performed at baseline, 6 and 12 months in a sub-set of LIFE-P study participants at the Cooper Institute site. We used the following criteria to determine if participants achieved maximal effort: 1) respiratory exchange ratio (RER) ≥ 1.1, 2) heart rate within 10 beats$\cdot$min^{-1} of the maximal level predicted by age and 3) rating of perceived exertion (RPE) > 17.

Results

The mean age of the 20 participants (80% female) tested was 74.7±3.4 years. The mean peak VO_2 was 12.1 (3.7) mL/kg/min. At baseline testing, only 20% of participants obtained a respiratory exchange ratio $\geq$ 1.10, only 35% achieved a peak heart rate within 10 beats of their age-predicted maximum heart rate and 18% had an RPE of >17. Subsequent testing at 6 and 12 months produced similar results.

Conclusions

In this pilot study of sedentary older persons at risk for mobility disability, we found that very few participants were able to achieve maximal effort during graded cycle ergometer testing.

Introduction

With the well-documented aging of the American population, research evaluating interventions that promote the maintenance of function and quality of life in the older individuals are of great public health importance. Physical activity has been promoted as an intervention to enhance physical function and attenuate the co-morbidities associated with the aging process. A hallmark means of assessing the success of aerobic exercise interventions is maximal graded aerobic exercise testing. However, few studies have evaluated the feasibility of maximal graded exercise testing in sedentary older persons who are at risk of disability, despite potential concerns related to the burden of chronic disease, participant tolerability, and quality of data.

The Lifestyle Interventions and Independence for Elders Pilot Study (LIFE-P) was a multicenter pilot study to test feasibility and develop plans for a definitive, randomized, single-blind, controlled trial to evaluate the efficacy of a physical activity intervention to reduce the incidence of major mobility disability in at-risk older adults. (Rejeski et al., 2005) The goal of LIFE-P was to provide key benchmarks to inform the design and implementation of a definitive trial. In a pilot study to the LIFE-P trial, we sought to evaluate the acceptability and feasibility of maximal graded exercise testing in older individuals at risk of disability in order to determine whether this might be a viable outcome measure in future intervention trials in this population.

Methods

LIFE-P Study and Participants

The LIFE-P study was conducted at four field centers (The Cooper Institute, Dallas, TX; Stanford University, Palo Alto, CA; University of Pittsburgh, Pittsburgh, PA; and Wake Forest

University, Winston-Salem, NC). The design and rationale of the LIFE-P study have been presented elsewhere.(Rejeski et al., 2005) Briefly, sedentary older adults were randomized to participate in either a walking-based physical activity program or a "successful aging" health information/education program lasting from 12 to 18 months depending on the month of randomization. The eligibility criteria were aimed at identifying persons aged 70-89 years who were at high-risk of mobility disability but who had not yet developed disability. High risk was defined as a score on the Short Physical Performance Battery (SPPB) less than 10. The SPPB includes assessments of balance, gait speed and the ability to rise from a chair and stratifies persons according to their disability risk on a 0-12 scale, with the risk of mobility disability rising sharply for scores less than 10. (Guralnik, Ferrucci, Simonsick, Salive, & Wallace, 1995)

Rational for maximal graded exercise testing in sub-group

Since LIFE-P was a pilot study, the Steering Committee decided that maximal graded exercise testing should be performed on a subset of participants to determine the need and/or feasibility of performing maximal graded exercise testing in the full-scale trial. The Cooper Institute site volunteered to take the lead on performing maximal graded exercise testing in a subset of participants. The primary substudy goals were to: 1) evaluate participant burden and acceptability of testing, 2) determine staff burden, and 3) determine whether potential changes in exercise in response to the physical activity-training program could be adequately assessed even though improvements in maximal aerobic fitness was not a goal of the intervention.

IRB Approval and Recruitment

The Cooper Institute IRB was concerned about the risk-benefit ratio of the maximal fitness testing, both in terms of participant burden and safety, especially because fitness was not specified as a study outcome in LIFE-P. The IRB approved the maximal graded exercise testing protocol with two conditions: 1) a physician must be in the room during testing, and 2) testing should not be performed in very high-risk individuals, such as those with diabetes or a previous history of cardiovascular disease (CVD).

Eligibility and enrollment for maximal graded exercise testing:

Eligibility for the exercise testing was based on clinical criteria for medical stability and on the IRB review of the risk/benefit ratio (as noted above). Of the 103 persons randomized at the Cooper Institute center, nearly half were ineligible, including those with diabetes (n= 20), previous CVD (n=21), or other limitations (primarily orthopedic) that made testing unsafe (n = 9). Many of these medically ineligible individuals met multiple exclusion criteria. Eligible individuals (n= 53) were informed about the opportunity to participate in the ancillary exercise testing only after all their baseline testing for the parent study was completed and they had been randomized. This was done to ensure that the ancillary exercise testing did not interfere with recruitment or other baseline assessments of the parent study. The details of the fitness testing pilot study were presented in a one-on-one session in a quiet, private room. While it was emphasized that participation in the pilot study had no bearing or impact on participation in the parent study, individuals were passively encouraged to partake in the pilot study. Further, individuals were informed that a small monetary incentive ($20) would be provided at each fitness test. Of the eligible individuals, less than half were interested in performing the maximal

graded exercise testing (n=24). Of the 24, 4 declined the day of testing (2 due to musculoskeletal issues and 2 due to stories from friends about exercise testing).

Exercise testing

Maximal graded exercise testing was performed at baseline, 6 months and 12 months. During each test, participants were monitored at rest and throughout exercise using a 12-lead electrocardiogram (ECG) system. All exercise tests were conducted using a Lode Excalibur Sport cycle ergometer (Groningen, Netherlands), an electronic, rate-independent ergometer, thereby allowing participants to self-select pedal rate without compromising data outcome. Participants were allowed to practice on the cycle ergometer before exercise testing as well as familiarize themselves with the mouth piece and nose clip. Exercise tests started at a low work load (15 W) and proceeded in 2 minute stages until volitional fatigue was reached. Each subsequent stage following test initiation proceeded in 15 W increments. Participants were asked to keep a constant cycling cadence within the range of 50-80 RPM and were strongly encouraged throughout the test. We chose to use a cycle ergometer instead of a treadmill because of safety issues, reduced participant burden and previous work demonstrating fitness testing with a cycle ergometer to be sensitive to change in response to low intensity exercise training. (Church, Earnest, Skinner, & Blair, 2007) Respiratory gases were measured using a Parvomedics True Max 2400 Metabolic Measurement Cart and gas exchange variables (VO_2, CO_2 production, ventilation, and respiratory exchange ratio [RER]) were recorded every 15 sec throughout the entire protocol. Ratings of perceived exertion (RPE) were obtained using the 20-point Borg scale during the last minute of each stage. As summarized in the most recent American of Sports Medicine Guidelines for Exercise Testing and Prescription, a variety of objective and subjective

indicators can be used to assess maximal effort during graded exercise testing. (American

College of Sports Medicine, 2006) Based on the available data and testing protocol we used the

following criteria for assessing if participants achieved maximal effort:

1) RER $\geq$ 1.1

2) Heart rate within 10 beats·min^{-1} of the maximal level predicted by age.

3) RPE > 17

Statistical Analysis

Group descriptive statistics were compared between LIFE-P participants who did and did

not participate in exercise testing at the Cooper Institute site. Between group differences were

tested using analysis of variance with adjustment for gender. The mean VO_{2max}, maximum heart

rate, RER and RPE were calculated for the group for each testing time point. We calculated the

percentage of participants that obtained an RER of $\geq$ 1.1, came within 10 or fewer heart beats of

their predicted maximum heart rate (220-age) and those that had an RER > 17 up for each testing

time point. All reported P values are two-sided. All analyses were performed using SAS

version 9.0 (Cary, NC).

Results

The study population consisted of 16 women and 4 men (Table 1). The mean age was

74.7 (3.4) years and mean SPPB score was 7.6 (1.4). The participants who participated in

exercise testing weighed less, had a smaller mean waist circumference and a faster mean 400

meter walking speed compared with the participants who did not participate in the exercise test

($P < 0.05$ for each in gender adjusted analyses). Twenty participants were tested at baseline, 13

at the six-month follow-up, and 15 at the 12-month follow-up. At both the 6 month and 12 months testing the staff went to great efforts to get participants back for testing. No adverse events were noted during the testing and none of the tested participants dropped out of the parent study.

Table 2 provides the results of the exercise testing, among all participants and subsequently among only those who had test data at all three time points. Whether considered in relative or absolute terms, the peak VO_2 values were very low at each of the time points. (American College of Sports Medicine, 2006) At baseline, for example, the mean absolute VO_2 was 0.9 (.3) L/min and mean relative VO_2 was 12.1 (3.7) ml/kg/min among all tested participants. Values for the mean RER attained were also low at all three time points. At baseline, for example, the mean RER was 1.05 (0.07), suggesting as a group that maximal exercise effort was not achieved. Similarly, mean RPE at each testing point was low. For example, at baseline mean RPE was 16.2 (1.7) and at all three testing points mean RPE was < 17.

The number (%) of participants meeting conventional benchmarks for maximal exercise effort are provided in Table 3. At each time point, relatively few participants (33-46% for all data) reached the predicted maximum heart rate. Similarly, relatively few (20-62%) achieved an RER $\geq$ 1.1 across the three time points or achieved an RPE > 17 (15.4-20%). Further there was a great variability in terms of which individuals met the criteria for max across the testing time points. For example, only 3 of the 11 (27%) met the maximum heart rate criteria, 2 of 11 (18%) met RER criteria and none met RPE criteria at all three time points.

Discussion

The goal of this study was to evaluate the acceptability and feasibility of maximal graded exercise testing in sedentary older individuals who are at risk of disability. Almost half of the potential participants had a high-risk medical condition that made the testing inadvisable, according to the local IRB. Of the eligible participants, less than half were interested in the testing.   At each of the three time points, relatively few participants met accepted criteria for defining a successful maximal test, making the data difficult to interpret. The staff that performed the testing were highly experienced and found that despite strong verbal encouragement pushing participants to maximal effort was quite challenging. This may be due to an age-related decrease in chronotropic function, as well as the anecdotal observation that participants often abruptly stopped the test for reasons other that cardiopulmonary fatigue, such knee pain or non-specific complaints of discomfort. This is supported by the mean group SPPB of 7.6 which suggests that these were individuals at high risk for disability and that many were starting to develop function limitations.

Other investigators have questioned the value of maximal testing in older adults to measure fitness because many older adults fall below the level of fitness that these tests were designed to assess. (Gill, DiPietro, & Krumholz, 2000) Walking tests such as the 6-minute walk, or the long distance or 400 meter walk have been shown to correlate well with maximal testing in those who can do both. (Enright et al., 2003; Meyerhardt et al., 2006; Simonsick, Fan, & Fleg, 2006) These prior studies have also reported that many older adults and chronically ill people with congestive heart failure or chronic obstructive pulmonary disease fall below the level of walking speed that is needed for the slowest starting level of treadmill based tests. (Peeters & Mets, 1996; Swerts, Mostert, & Wouters, 1990; Newman, Haggerty, Kritchevsky, Nevitt, &

Simonsick, 2003) Our experience confirms these prior observations and provides data on a specific subgroup who are at high risk for disability, but able to walk the distance required for one of these modified tests.

Based on this experience, there are a number of reasons why we recommend that maximal graded exercise testing not be considered as a proxy outcome for trials in older adults at high risk for mobility disability. Given the low percentage of participants who achieved conventional benchmarks of maximal testing at any of the testing points combined with the great individual variation in regard to achieving max over the testing points, the data are of little value for assessing change in maximal fitness. Though not specifically quantified for the purposes of this pilot study, the costs associated with the equipment, training and staff time needed to conduct maximal graded exercise testing are substantial. While some participants were initially excited about participating in the exercise testing, this enthusiasm waned with each subsequent testing. This not only poses a problem for the exercise testing data, but more importantly could compromise the integrity of the whole study should participants dropout in order to avoid repeat maximal exercise tests.

Our experience with maximal graded exercise testing should not dissuade investigators from using other forms of exercise or functional testing in older adults at high risk for disability, such as the 6-minute walk, 400-meter walk or sub-maximal exercise testing. (Enright et al., 2003; Meyerhardt et al., 2006; Simonsick et al., 2006) For example, the 400 meter walk has low participant and researcher burden, is highly reproducible and is predictive of mortality and mobility disability. (Newman et al., 2003; Simonsick, Montgomery, Newman, Bauer, & Harris, 2001; Bittner et al., 1993)

Conclusion

In this pilot study of sedentary older persons at risk for mobility disability, we found that very few participants were able to achieve maximal effort during graded cycle ergometer testing. These findings diminish the potential value of this type of testing for assessing changes in fitness in sedentary older persons.

Reference List

American College of Sports Medicine (2006). *ACSM's Guidelines for Exercise Testing and Prescription*. (Seventh ed.) Philadelphia: Lippincott Williams & Wilkins.

Bittner, V., Weiner, D. H., Yusuf, S., Rogers, W. J., McIntyre, K. M., Bangdiwala, S. I. et al. (1993). Prediction of mortality and morbidity with a 6-minute walk test in patients with left ventricular dysfunction. SOLVD Investigators. *Journal of the American Medical Association, 270,* 1702-1707.

Church, T. S., Earnest, C. P., Skinner, J. S., & Blair, S. N. (2007). Effects of different doses of physical activity on cardiorespiratory fitness among sedentary, overweight or obese postmenopausal women with elevated blood pressure: a randomized controlled trial. *Journal of the American Medical Association, 297,* 2081-2091.

Enright, P. L., McBurnie, M. A., Bittner, V., Tracy, R. P., McNamara, R., Arnold, A. et al. (2003). The 6-min walk test: a quick measure of functional status in elderly adults. *Chest, 123,* 387-398.

Gill, T. M., DiPietro, L., & Krumholz, H. M. (2000). Exercise Stress Testing for Older Persons Starting an Exercise Program. *Journal of the American Medical Association, 284,* 2591.

Guralnik, J. M., Ferrucci, L., Simonsick, E. M., Salive, M. E., & Wallace, R. B. (1995). Lower-extremity function in persons over the age of 70 years as a predictor of subsequent disability. *N.Engl.J.Med., 332,* 556-561.

Meyerhardt, J. A., Heseltine, D., Niedzwiecki, D., Hollis, D., Saltz, L. B., Mayer, R. J. et al. (2006). Impact of physical activity on cancer recurrence and survival in patients with stage III colon cancer: findings from CALGB 89803. *J Clin.Oncol., 24,* 3535-3541.

Newman, A. B., Haggerty, C. L., Kritchevsky, S. B., Nevitt, M. C., & Simonsick, E. M. (2003). Walking performance and cardiovascular response: associations with age and morbidity--the Health, Aging and Body Composition Study. *J.Gerontol.A Biol.Sci.Med.Sci., 58,* 715-720.

Peeters, P. & Mets, T. (1996). The 6-minute walk as an appropriate exercise test in elderly patients with chronic heart failure. *J.Gerontol.A Biol.Sci.Med.Sci., 51,* M147-M151.

Rejeski, W. J., Fielding R.A., Blair, S. N., Guralnik, J. M., Gill, T. M., Hadley, E. C. et al. (2005). The lifestyle interventions and independence for elders (LIFE) pilot study: design and methods. *Contemporary Clinical Trials, 26,* 141-154.

Simonsick, E. M., Fan, E., & Fleg, J. L. (2006). Estimating cardiorespiratory fitness in well-functioning older adults: treadmill validation of the long distance corridor walk. *J.Am.Geriatr.Soc., 54,* 127-132.

Simonsick, E. M., Montgomery, P. S., Newman, A. B., Bauer, D. C., & Harris, T. (2001). Measuring fitness in healthy older adults: the Health ABC Long Distance Corridor Walk. *J.Am.Geriatr.Soc., 49,* 1544-1548.

Swerts, P. M., Mostert, R., & Wouters, E. F. (1990). Comparison of corridor and treadmill walking in patients with severe chronic obstructive pulmonary disease. *Phys.Ther., 70,* 439-442.

Table 1. Characteristics of Participants at Cooper Institute Field Site

	Participants Tested N=20 Mean (SD)	Participants Not Tested N = 83 Mean (SD)
Age, years	74.7 (3.4)	76.0 (4.5)
Female, %	80%	65%
Weight, kg	74.8 (11.5)	84.9 (17.2)*
Waist circumference, cm	93.4 (14.5)	101.0 (18.0)*
Systolic BP, mmHg	132.2 (19.1)	130.8 (23.7)
Diastolic BP, mmHg	72.6 (9.29)	67.6 (13.3)
SPPB score	7.6 (1.4)	7.6 (1.4)
400 m walk speed, m/s	1.03 (0.19)	0.93 (0.15)*

* p< 0.05 for between group differences using gender adjusted analyses

Table 2. Exercise Testing Data Across the Three Testing Time Points

Time	N	Absolute VO_{2max} (L/min)	Relative VO_{2max} (ml/kg/min)	Max Heart Rate	Respiratory Exchange Ratio	Rating of Perceived Exertion
All Available Data						
Baseline	20	0.9 (0.3)	12.1 (3.7)	125.7 (19.3)	1.05 (0.07)	16.2 (1.7)*
6 Months	13	1.1 (0.3)	12.9 (5.1)	130.5 (13.8)	1.12 (0.07)	15.7 (1.9)
12 Months	15	1.0 (0.3)	13.3 (3.3)	127.8 (15.6)	1.09 (0.07)	15.9 (2.2)
Participants Tested At All Three Times Points						
Baseline	11	1.0 (0.4)	12.9 (4.1)	127.5 (19.8)	1.08 (0.07)	16.4 (1.9)**
6 Months	11	1.1 (0.3)	14.5 (3.5)	131.7 (11.7)	1.13 (0.06)	15.6 (2.0)**
12 Months	11	1.1 (0.4)	13.7 (3.7)	128.5 (16.4)	1.11 (0.05)	15.2 (1.9)**

* n =17, ** n =10

Table 3. Prevalence of Meeting Measures of Maximal Effort Across the Three Testing Points

Time	N	Max HR > Age Predicted HR - 10	Respiratory Exchange Ratio ≥1.1	Rating of Perceived Exertion > 17
All Available Data				
Baseline	20	35%	20%	18%*
6 Months	13	46%	62%	15%
12 Months	15	33%	60%	20%
Participants Tested At All Three Times Points				
Baseline	11	45%	36%	30%**
6 Months	11	55%	73%	10%**
12 Months	11	36%	73%	20%**

HR = Heart Rate

* n =17, ** n =10