Maximal Aerobic Capacity Testing of Older Adults: A Critical Review

Deanna L. Huggett, Denise M. Connelly, and Tom J. Overend

School of Physical Therapy, University of Western Ontario, London, Canada.

Most of the data that describe maximal oxygen uptake (VO2max) and the requirements for its attainment have been developed using young adults as subjects. Many older adults are unable to satisfactorily complete a maximal exercise effort in a standard exercise stress test. This review describes exercise tests currently available to measure VO2max in older adults. PubMed and CINAHL databases were searched for studies including healthy individuals older than 65 years with reproducible descriptions of the testing protocol. The research on VO2max testing in healthy individuals older than 65 years is limited, does not describe the protocols in detail, and/or lacks information on the psychometric properties of the exercise tests. There is a need for refinement of the few existing protocols for testing aerobic capacity in older adults, as well as the development of new protocols specifically applicable to older adults. Consensus on the criteria defining VO2max attainment during exercise in older adults is required, as well as agreement on the most appropriate exercise protocols and equipment, specific to older adults, to successfully fulfill these criteria.

MAXIMAL oxygen uptake (VO2max), also known as functional aerobic capacity, represents the maximal rate of oxygen utilization by exercising muscles and is considered the gold standard measure of the functional limit of the cardiorespiratory system (1,2). Functional aerobic capacity can be objectively measured with an exercise stress test and is useful in assessing physical fitness, diagnosing and defining the prognosis of ischemic heart disease, developing an exercise prescription, and guiding cardiac rehabilitation (3). Assessment of VO2max also allows clinicians to identify individuals who may have difficulty performing various functional tasks, as the energy requirements (oxygen uptake) for many physical tasks have been determined and are now readily available (4,5).

A maximal exercise test with direct measurement of oxygen uptake is considered the most accurate method of assessing aerobic power, but different values may be obtained for the same individual when different testing modalities are used (6,7). Exercise stress testing is most commonly carried out using a treadmill or bicycle ergometer. The treadmill provides a more familiar and functional exercise modality, namely walking, and has been shown to have greater diagnostic sensitivity than the bicycle ergometer (8). However, exercise on a treadmill may not be appropriate for many older individuals with balance deficits or conditions, such as arthritis, in which weight-bearing exercise may cause excessive joint pain. VO2max values during cycle ergometer exercise average 11% lower (range, 8%–15%) than those during treadmill exercise, primarily because of the smaller volume of exercising muscle mass (9–12). Regardless of the modality used, an appropriate exercise protocol should be followed.

VO2max is known to decline with age (13–15). The age-related rate of decline in VO2max differs among investigations and is highly variable among individuals within a population (16). The most commonly cited estimation is that VO2max declines at a rate of 1% per year after the third decade of life (17). The decreased cardiorespiratory function and muscular performance associated with advancing age and/or inactivity can significantly diminish an individual’s functional capacity (18–20). Many sedentary older adults are of an age that theoretically places them very close to an aerobic capacity threshold. Any further decline may render them unable to perform their activities of daily living and thus impair their ability to live independently (19,21,22).

The technical requirements of measuring VO2max and concerns regarding patient safety often limit direct measurement of VO2max (23). Thus, it is often calculated indirectly through linear regression equations relating exercise performance to oxygen uptake (24,25). Such prediction equations, however, tend to be population specific and valid only through a limited age range. The prevalence of asymptomatic coronary artery disease and the coexistence of other chronic conditions and physical limitations result in distinct differences between younger and older adults (26). Many older persons have at least one, if not multiple, chronic medical conditions, such as arthritis, hypertension, or diabetes (27). The involvement of a variety of medical conditions, many of which may be asymptomatic, may make prediction equations inaccurate for older adults.

The burden of chronic conditions and physical limitations increases substantially with age and contributes to the difficulty of measuring VO2max in older persons (26,28). Measurement of VO2max is particularly useful for fit and highly motivated persons, but it depends heavily on subjective factors such as muscle fatigue, perceived exhaustion, and level of motivation, as well as the clinician’s willingness to allow participants to exercise to exhaustion (29).

The requirements for attaining VO2max are variable in the literature and include a measurement of the ratio between...
carbon dioxide production and oxygen consumption \((VCO_2/\ VO_2)\) (30), or respiratory exchange ratio (RER), of \(>1.0\) (31,32) to \(>1.15\) (33); a peak exercise heart rate (HR) of at least 85% of age-predicted maximum (34) or a HR within 10 beats per minute (bpm) of age-predicted maximum (35); and a plateau in oxygen consumption, defined variously as an increase in oxygen uptake of \(<1.5 \text{ ml/min per kg}\) (36) to \(<2.0 \text{ ml/min per kg}\) (17,37). Borg’s Rating of Perceived Exertion scale is now frequently used as a marker as well (38,39). However, most of the data that describe \(V_O_2\max\) and the requirements for its attainment have been developed using young adults as subjects (40). These data have subsequently been adapted or extrapolated for use in subjects of various ages. Those studies that do describe maximal aerobic capacity in older adults or declines in fitness with age are based on samples of subjects that included either relatively few older subjects, or older subjects who were selected on the basis of known physical activity patterns (13,41,42). Moreover, most of these studies included virtually no subjects older than 70 years of age (30,43).

Many older adults are unable to satisfactorily complete a maximal exercise effort in a standard exercise stress test (26). The increment of \(V_O_2\) accepted as a plateau in younger subjects could represent 10% or more of the total \(V_O_2\max\) of an older adult or unfit individual (17,37). The high initial exercise rates and the long duration of exercise in the various protocols often prevent successful completion of exercise stress tests in older adults (40). The effort of the older adults is not necessarily limited by the power of the heart and lungs, but more often by factors such as dyspnea (44,45), fear of overexertion (46,47), muscular weakness (44,45), poor motivation (44), or the appearance of electrocardiographic abnormalities (48). Currently, there is a debate regarding (a) the need for attainment of a plateau in \(V_O_2\), or (b) whether the peak \(V_O_2\) (\(V_O_2\peak\)) value obtained is sufficient to report as \(V_O_2\max\). \(V_O_2\peak\) represents the level of oxygen consumption recorded at the peak workload (30). Thus agreement is needed on what is acceptable as a maximal or near-maximal effort in older adults (30).

This critical review describes exercise tests currently available to measure the maximal aerobic capacity of older adults. A summary of the protocols used with healthy participants older than 65 years of age is included, and recommendations are made for use of these tests in clinical and research settings. Potential areas for future research are also outlined.

**METHODS**

**Literature Search**

A computer search was conducted of the English language medical literature using the databases PubMed (1966 through February 2004) and CINAHL (1982 through February 2004). Combinations of the following key words were used: “exercise test,” “maximal aerobic capacity,” “\(V_O_2\max\),” “maximal oxygen consumption,” “exercise capacity testing,” “older adults,” and “aging.” We also conducted a manual search of the references of retrieved articles to identify additional potential studies. Investigations published only in abstract form were not included, and no attempt was made to contact experts in the field to uncover unpublished material.

**Inclusion and Exclusion Criteria**

The following criteria for inclusion were used: (a) English language studies published in peer-reviewed journals; (b) older adult subjects (\(\geq 65\) years); (c) healthy (e.g., non-ischemic electrocardiographic response) subjects; and (d) inclusion of a reproducible description of the exercise testing protocol used. Healthy subjects were deemed to have no known medical conditions. Only data related to the protocols used and the results of \(V_O_2\max\) testing were included in this review. Using the above criteria, we reviewed the search results from PubMed and CINAHL and decided which articles to retrieve. Retrieved articles fulfilling the inclusion criteria were reviewed, and data were extracted and summarized. Quality of the studies was not used as an exclusion criterion but has been addressed in the discussion of the articles included in this review.

**RESULTS OF LITERATURE SEARCH**

A compilation of the available literature on protocols for \(V_O_2\max\) testing in older adults follows in alphabetical order. An overview of the available tests and their relevant studies is provided, followed by an evaluation of the protocols and discussion. Maximal treadmill test protocols are summarized in Table 1, and maximal cycle ergometer test protocols are summarized in Table 2.

**Maximal Treadmill Tests**

Foster and colleagues (36) tested the reproducibility of \(V_O_2\max\) values in eight older women (80.6 ± 3.7 years). The researchers used a continuous incremental treadmill test, starting at a 0% grade and a speed of 40 m/min (1.5 mph, 2.40 km/h). The treadmill speed was held constant throughout the test, and the elevation was increased by 1% every 3 minutes. Each subject was tested on three occasions. The results indicated that \(V_O_2\max\) is reproducible for older women, with significant intra-individual correlation coefficients for all comparisons (\(r = 0.70–0.84\)).

Hollenberg and colleagues (30) tested 1101 subjects between 55 and 94 years of age (438 men, mean age 67.8 ± 7.8 years; 663 women, mean age 67.8 ± 7.9 years), and used the Cornell modification of the Bruce treadmill protocol. This modification consists of eleven 2-minute stages (49). The test starts at 1.7 mph (2.72 km/h) and a 0% grade and increases in a step-like manner to a final stage at 5 mph (8 km/h) with an 18% grade (see Table 1 for details). Hollenberg and colleagues suggest that the Cornell exercise protocol might be appropriate for use in large epidemiologic studies of healthy adults older than age 65 because of the small increments in work rate that produced average test durations of 10 minutes, close to the recommended target for such tests.

Katzel and colleagues (50) performed a longitudinal follow-up on male athletes compared to sedentary counterparts. Subjects included 42 athletes with a baseline age of 63.4 ± 1.0 years and a follow-up age of 71.4 ± 1.1 years, as well as 47 sedentary men with a baseline age of 61.1 ± 0.9 years and a follow-up age of 70.4 ± 0.9 years. \(V_O_2\max\)
Table 1. Maximal Treadmill Protocols for Older Adults

<table>
<thead>
<tr>
<th>Authors (Ref.)</th>
<th>Subjects</th>
<th>Protocol</th>
</tr>
</thead>
<tbody>
<tr>
<td>Foster et al., 1986 (36)</td>
<td>8 older women (80.6 ± 3.7 y)</td>
<td>Continuous incremental treadmill test</td>
</tr>
<tr>
<td>Hollenberg et al., 1998 (30)</td>
<td>1101 subjects 55–94 y of age</td>
<td>Cornell modification of the Bruce protocol (eleven 2-min stages):</td>
</tr>
<tr>
<td></td>
<td>438 men (67.8 ± 7.8 y)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>663 women (67.8 ± 7.9 y)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Katzel et al., 2001 (50)</td>
<td>Longitudinal follow-up of male athletes vs sedentary counterparts</td>
<td>Modified Balke protocol:</td>
</tr>
<tr>
<td></td>
<td>42 athletes; baseline = 63.4 ± 1.0 y, follow-up = 71.4 ± 1.1 y</td>
<td></td>
</tr>
<tr>
<td></td>
<td>47 sedentary men; baseline = 61.1 ± 0.9 y, follow-up = 70.4 ± 0.9 y</td>
<td></td>
</tr>
<tr>
<td>Morey et al., 1998 (4)</td>
<td>161 community-dwelling adults, ages 65–90 y (71.6 ± 5.1 y)</td>
<td>Subjects exercised to symptom-limited maximum</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Paterson et al., 1999 (58)</td>
<td>298 subjects (152 men, 146 women), representing the non-institutionalized population aged 55–85 y</td>
<td>4-min warm-up, speed 45.5 m/min (1.7 mph, 2.72 km/h, 0% grade)</td>
</tr>
<tr>
<td></td>
<td>Six 5-y age groups starting with age 55; 35 men and women in each stratum</td>
<td></td>
</tr>
<tr>
<td>Sidney &amp; Shephard, 1977 (59)</td>
<td>Men and women aged 60–83 y</td>
<td>Set speed of 2.5–3.5 mph (4.0–5.6 km/h)</td>
</tr>
<tr>
<td></td>
<td>26 men (63.7 ± 2.6 y)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>29 women (63.4 ± 3.6 y)</td>
<td></td>
</tr>
<tr>
<td>Tanaka et al., 1997 (38)</td>
<td>156 women (age 20–75 y): 84 endurance-trained runners including 16 subjects &gt;60 y (64 ± 1) and 72 sedentary subjects including 13 subjects &gt;60 y (66 ± 1)</td>
<td>6- to 10-min warm-up</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thomas et al., 1987 (60)</td>
<td>224 men ages 55 and 68 y, mean 62.3 y</td>
<td>Protocol 1: continuous, progressive test to maximum</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 1. Maximal Treadmill Protocols for Older Adults (Continued)

<table>
<thead>
<tr>
<th>Authors (Ref.)</th>
<th>Subjects</th>
<th>Protocol</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tonino &amp; Driscoll, 1988 (61)</td>
<td>9 healthy older persons (5 males, 4 females, aged 62–79 y, mean 70.4 y)</td>
<td>Continuous modified Balke protocol:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Four 3-min stages of 2 MET increments, followed by 2-min stages of 1 MET increments until subjective fatigue</td>
</tr>
</tbody>
</table>

Note: HR = heart rate; MET = metabolic equivalent (3.5 ml O2/kg per min).

was determined using a treadmill and the modified Balke protocol (51). In this procedure, the initial treadmill speed varied according to the individual, and was set to produce approximately 70% of the subject’s peak HR—determined from an initial screening test on the treadmill at baseline by using the Bruce treadmill protocol. Elevation of the treadmill increased from 0% to 4% after 2 minutes, to 6% after 4 minutes, and then by 2% every minute until the subject reached voluntary exhaustion. Criteria for VO2max included meeting two of the following three criteria: (a) HR at maximal exercise $>95\%$ of age-predicted maximal HR, (b) RER $\geq$ 1.10, and (c) increase in VO2 of $<$0.2 l/min during the final increase in workload (i.e., VO2max plateau). The rate of decline of VO2max was significantly greater in the endurance athletes as compared with the sedentary men, particularly those athletes who were unable to maintain their high level of training. The authors suggest that much of the observed decline in VO2max in the athletes was due to lifestyle changes rather than aging itself.

Morey and colleagues (4) studied 161 community-dwelling older adults between the ages of 65 and 90 years (mean 71.6 ± 5.1 years). They used a treadmill and exercised the subjects to a symptom-limited maximum. Their protocol was as follows: minutes 1–3 at 1.5 mph (2.4 km/h) at 0%, 2%, and 4% grades; minutes 4–9 at 2.0 mph (3.2 km/h) at 3%, 5%–9% grades; minutes 10–13 at 2.5 mph (4 km/h) at 8%, 10%, 12%, and 14% grades; minutes 14–16 at 3.0 mph (4.8 km/h) at 12%–14% grades; and minutes 17–20 at 3.3 mph (5.3 km/h) at 14%, 16%, 18%, and 20% grades. VO2peak was strongly associated with physical function, as measured by a summary score combining scores from five functional scales ($p = .004$) including: (a) the seven-item instrumental activities of daily living subscale of the Older Americans Resources and Services (OARS) Multidimensional Functional Assessment questionnaire (52); (b) five dichotomous questions of self-reported difficulty in performing instrumental activities of daily living derived from the Nagi study of disability (53); (c) a three-item self-report of advanced activities of daily living (54); (d) the Physical Functioning Scale of the Medical Outcomes Study short form health survey SF-36PF (55); and (e) the Falls Efficacy Scale (56,57).

Paterson and colleagues (58) used a noninstitutionalized population of older adults between 55 and 85 years of age to develop normative data on VO2max by 5-year increments. Subjects were given a 4-minute adaptation period and warmed up on a treadmill at a speed of 45.5 m/min (1.7 mph, 2.72 km/h) at a 0% grade. Following the warm-up period, speed and grade were increased in small increments in a ramp-like manner. The increases were individually based, and were designed to elicit VO2 increases each minute ranging from 1 to 3 ml/kg per min, with each subject’s test duration between 8 and 12 minutes. Criteria for determining VO2max included: (a) the occurrence of a 15-second plateau in VO2 despite a continuing increase in power output; or (b) an RER $>1.0$ and an HR within 5 bpm of the age-specific maximal HR (HRmax = 220 bpm − age). Results from this study indicate that a linear model provides the best fit for age-related declines in VO2max, although age explained at most 37% of the variance across the age range.

Sidney and Shephard (59) compared submaximal cycle ergometer and maximal treadmill tests in men and women aged 60–83 years (men: n = 26, 63.7 ± 2.6 years; women, n = 29, 63.4 ± 3.6 years). Their treadmill protocol had

Table 2. Maximal Cycle Ergometer Protocols for Older Adults

<table>
<thead>
<tr>
<th>Authors (Ref.)</th>
<th>Subjects</th>
<th>Protocol</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cress &amp; Meyer, 2003 (62)</td>
<td>Men and women ($N = 192; 76 ± 7 y$) from single-family community dwellings or retirement communities with multiple levels of care</td>
<td>Familiarization period</td>
</tr>
<tr>
<td>Fairbarn et al., 1994 (64)</td>
<td>231 men and women equally divided within decades between 20 and 80 y (no breakdown of specific ages within decades provided)</td>
<td>Ramp test</td>
</tr>
<tr>
<td>Jones et al., 1985 (65)</td>
<td>100 healthy subjects (50 male and 50 female), selected to provide an even distribution of age (15–71 y) and height 10 men and 10 women between 55 and 71 y</td>
<td>Initial power output at 16 or 32 W</td>
</tr>
<tr>
<td>Pochlman &amp; Danforth, 1991 (66)</td>
<td>19 older adults (13 men, 6 women) (64 ± 1.6 y)</td>
<td>Initial power setting 16.3 W</td>
</tr>
</tbody>
</table>

- Power increased by 16 W/min until subject reached symptom-limited maximal power output
- Power increased by 16.3 W/min until a symptom-limited power output reached

- Initial workload for men: 50 W at 50 rpm
- Initial workload sustained for 3 min
- Increased by 25 W every 2 min until exhaustion or until subjects unable to maintain 50 rpm.
a fixed speed of 2.5–3.5 mph (4.0–5.6 km/h), and the slope was individually adjusted every 3 minutes to achieve an HR of 75%–85% of the maximum HR during the 9th minute of exercise. At the beginning of the 10th minute, the slope was increased to achieve an estimated 90%–100% of the individual’s VO₂max. Additional 1%–2% increases in slope were made every 2 minutes until subjective exhaustion, or the presence of clinical indications to stop the test. Results from this study showed that only 75% of the subjects reached the plateau definition proposed for younger subjects. This study also indicated that valid data cannot be obtained without working subjects to exhaustion, as VO₂max was higher in subjects making a “good” exhausting effort (evidenced by high lactate levels) than in subjects making a “fair” effort (lower lactate levels).

Tanaka and colleagues (38) compared VO₂max in 72 sedentary women to 84 endurance-trained women, aged 20–75 years (16 endurance-trained athletes >60 years, mean 64 ± 1 years; 13 sedentary women >60 years, mean 66 ± 1 years). Following a 6- to 10-minute warm-up period on the treadmill, each subject continued to either run or walk on the treadmill at a comfortable speed that elicited 70%–80% of age-predicted maximal HR (equation to determine maximal HR not provided). Grade was increased by 2.5% every 2 minutes until volitional exhaustion. The Borg Rating of Perceived Exertion scale was used at the end of each stage. Age alone accounted for 69% of the total variance in VO₂max in endurance-trained women and 57% in sedentary women. The absolute rate of decline in VO₂max was greater in the endurance-trained women (−5.7 ml/kg/min/decade) compared with the sedentary women (−3.2 ml/kg/min/decade; p < .01), but the relative (%) rate of decline was similar (−9.7% vs −9.1% per decade). The greater absolute rate of decline in VO₂max in endurance-trained women compared with sedentary women was not associated with a greater rate of decline in maximal HR, nor was it related to training factors.

Thomas and colleagues (60) used three different treadmill test protocols to assess VO₂max in 224 men between 55 and 68 years of age (mean 62.3 years). Protocol 1 was a continuous, progressive test starting at 1.33 m/s (3.0 mph, 4.80 km/h) and 0% grade for 4 minutes. The grade was then increased by 2.5% every 2 minutes up to 20%, at which point the speed was increased by 0.133 m/s (0.3 mph, 0.48 km/h) every 2 minutes until a maximal effort was reached. A Modified Protocol 1 started at 1.33 m/s (3.0 mph, 4.8 km/h) and 0% grade for 4 minutes, at which point the grade was increased by 5% every 2 minutes up to 20%, then the speed was increased by 0.267 m/s (0.6 mph, 0.96 km/h) until exhaustion. Protocol 2 was a discontinuous test involving 5–6 minutes at each of three submaximal workloads of 50%, 65%, and 85% of HR reserve [(maximum HR – resting HR) × 60% + resting HR]. Following the third stage and 3–5 minutes of rest, the grade of the treadmill was increased by 2.5% every 2 minutes up to 20%, at which point the speed was increased by 0.133 m/s (0.3 mph, 0.48 km/h) until exhaustion. All subjects completed Protocol 1, and then were randomly assigned to groups to repeat Protocol 1 (n = 24), complete Protocol 2 (n = 145), or complete the Modified Protocol 1 (n = 38). In all protocols, a plateau in VO₂ was defined as an increase of ≤2 ml/kg per min with a change in work rate. The results of this study indicate that reliability of repeated VO₂max measures in older men is acceptable (r = .67–.90), despite the finding that only one-third of the men demonstrated a plateau in VO₂. More specifically, the correlation between the first and second tests for VO₂max was high for Protocol 1 (0.90) and between Protocol 1 and 2 (0.87), but was lower for the Modified Protocol 1 test (0.67). Most of the older adults demonstrated a higher VO₂max on repeated testing using the same or similar test protocol, suggesting that at least two maximal tests are required to precisely quantify the cardiorespiratory capacity of older men.

Tonino and Driscoll (61) studied nine healthy older adults (5 men and 4 women, mean age 70.4 years). They used a continuous modified Balke protocol on the treadmill, whereby four 3-minute stages of 2-MET increments were followed by 2-minute stages of 1-MET increments until the subject reached subjective fatigue. This study also found that VO₂max is a reliable measure in older adults (test–retest reliability: p < .05), whereas maximal duration of exercise, submaximal HR, and submaximal VO₂ demonstrated inconsistent results (p > .05). Unfortunately, correlation coefficients and actual error are not reported. Contrary to the results of Thomas and colleagues (60), Tonino and Driscoll (61) found that VO₂max was reproducible on repeat testing. However, their results must be viewed with caution, as Tonino and Driscoll (61) had a very small sample size of nine persons.

**Maximal Cycle Ergometer Tests**

Cress and Meyer (62) collected VO₂peak data on a sample of 192 men and women (76 ± 7 years, no gender breakdown provided) using a cycle ergometer. Their protocol began with a familiarization period followed by a ramp test with power output increased at a rate of 8–16 W/min. No description was provided as to the initial power output or criteria for the rate of increase. The test was terminated when the participant demonstrated signs and symptoms of fatigue or exercise intolerance, electrocardiogram changes, or abnormal blood pressure or respiratory responses. VO₂peak data were reported so as to include data on individuals not meeting the typical criteria for maximal effort. Cress and Meyer also used two functional scales to measure functional limitation, including the SF-36PF (55) and the Continuous-Scale Physical Function Performance Test [CS-PFP] (63). Although Cress and Meyer used VO₂peak data for their analysis, they did note that 80.2% of their participants met two or more of the criteria for reaching VO₂max. Their criteria for VO₂max included: (a) an HRmax within 10 bpm of age-predicted maximum, (b) an RER >1.0, or (c) a rating of perceived exertion of at least 18 on the Borg Rating of Perceived Exertion Scale. They noted that by using only VO₂peak data physical reserve may have been underestimated for some subjects, but they felt that the data presented may be more representative of the actual physical ability of this population.

Fairbarn and colleagues (64) assessed 231 men and women equally divided within the 6 decades between the ages of 20 and 80 years. Using a cycle ergometer to test
VO\textsubscript{2\text{max}}. they set the initial power output either at 16 or 32 W. Regardless of initial output level, power was increased by 16 W/min until the participant reached a symptom-limited maximal power output. The VO\textsubscript{2\text{max}} data were used to develop prediction equations. Age alone accounted for 64% of the variability in VO\textsubscript{2\text{max}} in women and 35% of the variability in men. There was no information provided on the tolerance of the protocol by the older versus the younger participants.

Jones and coworkers (65) incorporated a sample of 50 men and 50 women between the ages of 15 and 71 years (10 men and 10 women between 55 and 71 years of age). A cycle ergometer was used to measure VO\textsubscript{2\text{max}} with an initial power setting of 16.3 W and an increase of 16.3 W/min until a symptom-limited power output was reached. They discovered that the effect of age alone on VO\textsubscript{2\text{max}} ($r = 0.47$ in both instances).

Poehlman and Danforth (66) studied a small sample of 19 older adult participants (13 men and 6 women, mean age 64 ± 1.6 years). Initially, the workload was set at 25 W for women and 50 W for men. After 3 minutes, workload was increased by 25 W every 2 minutes until exhaustion, or until subjects were no longer able to maintain a pedal speed of 50 rpm. The purpose of this study was to investigate the effects of endurance training on resting metabolic rate and plasma norepinephrine kinetics in older adults. No information was provided as to the ability of the older adults to tolerate the maximal testing protocol.

**DISCUSSION**

Several of the studies reviewed do not describe their protocols in explicit detail, and only a select few provide data on psychometric properties, thus making comparisons difficult. Also of note is that four of the nine treadmill protocols described included either only men or only women as subjects, making it difficult to compare results across test protocols. This discussion will attempt to briefly summarize similarities and differences between protocols as well as the ability of subjects to satisfy predetermined criteria for reaching VO\textsubscript{2\text{max}}.

**Speed and Grade Increases**

The studies by Foster and colleagues (36), Hollenberg and associates (30), Morey and coworkers (4), and Paterson and colleagues (58) all initiated their treadmill protocols at a 0% grade with a speed of either 1.5 or 1.7 mph (2.40 or 2.72 km/h). Although the initial work rates were similar, the subsequent ramp-like stages were quite varied, with some increasing both speed and grade [Hollenberg et al. (30), Morey et al. (4), and Paterson et al. (58)], whereas Foster and colleagues (36) kept speed constant throughout and only increased grade. A wide range in grade increases was also found among the studies. For example: Morey and colleagues (4), 1%–2% every minute; Hollenberg and coworkers (30), 1% every 30 seconds after the initial 6 minutes; Foster and associates (36), 1% every 3 minutes; and Tanaka and colleagues (38), 2.5% every 2 minutes.

Sidney and Shephard (59) and Thomas and colleagues (60) initiated their protocols at speeds >2.5 mph (4.0 km/h), whereas Sidney and Shephard (59) maintained a constant speed throughout the test and only increased grade, all three protocols described by Thomas and colleagues (60) increased speed after a certain grade level was achieved.

Thomas and colleagues (60) compared three different test protocols, two that incorporated differences in speed and grade increases, and one discontinuous protocol. They suggested that their Protocol 1, which involved a steady but slow increase in both speed and grade, was the most appropriate choice, if one did not require data on submaximal steady-state conditions, as would be available from a discontinuous protocol (60). The study by Thomas and colleagues (60) is unique in that the authors compared different protocols using the same group of subjects; the results thus provide insight on the appropriateness of the protocols for older adults. However, the study sample involved only men, and as such the results cannot be generalized to the greater population of all older adults.

**Individualization**

Katzel and associates (50), Paterson and coworkers (58), Sidney and Shephard (59), and Tanaka and colleagues (38) all included individualized components in their protocols. Paterson and colleagues (58) based individualization on increases in VO\textsubscript{2}, whereas the other studies individualized the increases in work rate based on percentage of maximum HR.

**Attainment of VO\textsubscript{2\text{max}} Criteria and Exercise Test Duration**

A plateau in oxygen consumption occurred in 75% of test sessions 1 and 2 and in 62% of test session 3 in the study by Foster and colleagues (36). Their step increase protocol resulted in longer tests (mean duration 20 minutes; range, 6–51 minutes) than did several of the other protocols. The authors did comment that the testing protocol may have been too lengthy for some individuals, resulting in decreased motivation both within a test, and following repeated testing (36).

In the study by Hollenberg and associates (30), 32.9% of women and 52.7% of men achieved an RER ≥ 1.10, which was their criterion for having achieved maximal exercise. The mean duration of exercise in this study was 10.0 ± 4.0 minutes for women and 13.2 ± 4.6 minutes for men. Hollenberg and associates (30) and Morey colleagues (4) describe VO\textsubscript{2peak} data, so as to include as many individuals as possible, rather than excluding data based on the criteria for maximal effort. Hollenberg and coworkers (30) found that only 41% of women and 36% of men who exercised for less than 8 minutes rated their maximal exercise effort as “hard” to “very, very hard” (Borg score of 15–20) compared with 71% and 84% of women and men, respectively, who exercised for longer than 13 minutes; this finding suggests that motivation influences VO\textsubscript{2peak} data.

The average time on the treadmill in the study by Morey and associates (4) was 10.6 ± 4.2 minutes. While Morey and associates (4) did report on VO\textsubscript{2peak} data, they also noted 78.8% of all subjects met two of the three criteria for attaining VO\textsubscript{2max}, although the three criteria were not described.

Paterson and coworkers (58) excluded 20.5% of their test results from data analysis because the criteria for VO\textsubscript{2max}
were not met. The individualized ramp protocol in their study allowed them to target a test termination between 8 and 12 minutes, which may have sustained subject motivation (58). The individualized protocol, as well as the shorter test duration, may have contributed to the finding that 79.5% of their subjects were able to obtain the criteria for VO2max.

In the study by Sidney and Shephard (59), 69.2% of men and 65.5% of women were able to achieve a plateau in VO2 of ≤0.2 ml/kg per min, while 96.2% of men and 72.4% of women were able to achieve an RER ≥1.00. Most of their subjects reached exhaustion after approximately 15 minutes of exercise. In Protocol 1 in the study by Thomas and colleagues (60), 33.6% of the 224 subjects reached a plateau in oxygen consumption. VO2max was not significantly different between those exhibiting a levelling-off of VO2max and those who did not reach a plateau, but the number of subjects reaching a plateau did rise significantly with repetition of the same or similar test protocol.

Katzel and colleagues (50) noted that all VO2max tests fulfilled at least two of the three following criteria: (a) the HR at maximal exercise was >95% of the age-predicted maximal HR, (b) the RER was ≥1.10, and (c) the VO2max reached a plateau during the final stage of exercise—that is, the increase in VO2 was <0.2 l/min during the final increase in work rate.

Tanaka and associates (38) had similarly strict criteria for attainment of VO2max in their study, and reported that all subjects fulfilled at least three of these four criteria: (a) a plateau in VO2 with increasing exercise intensity, (b) RER ≥ 1.15, (c) achievement of age-predicted maximal HR (±10 bpm), and (d) rating of perceived exertion ≥18 units. Treadmill tests lasted between 8 and 12 minutes in this study.

Although a plateau in VO2 has long been considered the classical requirement for attaining maximal cardiopulmonary effort, there has been considerable disagreement with respect to the definition and criteria for a plateau (67). The results of this review suggest that many older adults are unable to reach a plateau in VO2. This could be related to the fact that the criteria for a plateau in VO2 have been defined using data from younger adults. Of particular note is the criterion for maximum HR, defined in most studies as 220 – age. Tanaka and coworkers (68) performed both a meta-analysis and a laboratory-based study investigating HRmax in older adults. Their findings suggest the formula [208 – (0.7 × age)] is a more acceptable equation for both older men and women and that the use of the traditional equation (220 – age) may in fact underestimate the actual level of physical stress imposed during exercise testing. It would be interesting to evaluate the effect of this change in criteria for attaining VO2max on the previously described protocols to determine if a greater percentage of older adults met the new criteria.

In addition, there is sufficient evidence available to suggest that reaching a plateau is a highly variable outcome, even in younger adults (67, 69, 70). Myers and associates (67, 69) found significant variability in the slope of change in VO2 during either a constant or progressive change in external work, regardless of the sampling interval used; this finding suggests that the plateau concept has limitations for general application during standard exercise testing. As previously noted, Thomas and colleagues (60) found that reliability of repeated VO2max measures in older men is acceptable (r = .67–.90), despite the finding that only one-third of the men demonstrated a plateau in VO2. However, they also reported that reliability was higher for those individuals reaching a plateau in two of their three protocols (r = 0.92 vs 0.82, r = 0.86 vs 0.60) (55). Recently, Day and coworkers (70), using several different protocols, specifically investigated whether a VO2 plateau was a consistent outcome of maximal aerobic exercise. They discovered a plateau in the actual VO2 response is not a necessary consequence of incremental exercise testing. As well, in constant load tests, the peak value attained was not different from the plateau obtained in the plot of VO2 versus work rate. Therefore, they suggest that the VO2peak attained on a maximum effort incremental test is likely to be a valid index of VO2max, even though there may be no evidence of a plateau in the data themselves.

**Cycle Ergometer Tests**

Of the three studies that included VO2max measurement of older adults using cycle ergometers, none described the ability of the older adults to tolerate the test. In their study, Cress and Meyer (62) did not describe their protocol in detail, and did not provide the initial power output. They did, however, report that 80.2% of their subjects met two or more criteria for attaining VO2max (62). Both Fairbarn and colleagues (64) and Jones and coworkers (65) used an initial power output of 16 W [although an undisclosed number of subjects in the study by Fairbarn and associates (64) started their test at 32 W], with an increase of 16 W/min. Subjects exercised to a symptom-limited maximal power output, with no other indicators that maximal aerobic power was reached. The subjects in the Poehlman and Danforth study (66) started with a much higher workload (25 W for women and 50 W for men) and increased by a larger workload (25 W every 2 minutes). Information on individualization, attainment of VO2max criteria, and exercise test duration was not provided in the literature for the cycle ergometer tests.

**Functional Cut Point**

Morey and colleagues (4), Paterson and coworkers (58), and Cress and Meyer (62) all described a VO2 cut point for optimal functional performance. Morey and colleagues (4) described the optimal cut point between low physical function (requiring assistance with activities of daily living) and high physical function (independent living) to be 18.3 ml/kg per min (p < .0001), although no evidence of a levelling-off, indicative of a true threshold distinguishing individuals with low from those with high physical function, was reported. Paterson and associates (58) reported the minimum level of VO2max that corresponded to a fully independent life at age 85 in their study sample to be 17.7 ml/kg per min for men and 15.4 ml/kg per min for women. The major finding of the Cress and Meyer study (62) was the existence of a threshold of aerobic capacity (20 ml/kg per min) below which physical function was most affected by low fitness levels. All three of these studies used a ramp-like protocol, with Morey and colleagues (4) and Paterson and coworkers (58) using the
treadmill and Cress and Meyer (62) using the cycle ergometer for testing VO₂max. Both Morey and associates (4) and Cress and Meyer (62) used the 36-item SF-36PF (55) as one of the measures of functional limitation in their samples. Subjects in the study by Paterson and colleagues (58) were all living independently in the community at the time of testing. It is interesting to note that the study by Cress and Meyer (62), using the cycle ergometer, had a higher value for the threshold of aerobic capacity, despite the fact that VO₂max is typically lower using a cycle ergometer than a treadmill (9–12). With 80.2% of individuals able to meet the criteria for VO₂max in the study by Cress and Meyer (62), as compared to 74% of the individuals in the study by Paterson and colleagues (58), it might be possible that the individuals in the study by Cress and Meyer (62) were a higher functioning group. Both Morey and associates (4) and Cress and Meyer (62) used VO₂peak for their data analysis.

Clinical Use of Exercise Tests

The treadmill has been lauded for its use as a testing device because of the familiarity of subjects with walking, and for the purpose of data comparison, because there is a great deal more research using the treadmill than using any other type of equipment (36). As well, many studies have found that VO₂max is lower using a cycle ergometer than a treadmill (7,10,71). However, the difference in VO₂max noted between treadmill and cycle ergometer testing may be exaggerated in older adults, because they are forced to use muscle groups they are unaccustomed to using, during a form of exercise with which they are unfamiliar (16). The lack of well-described cycle ergometer protocols for older adults, as well as the lack of descriptive data provided on tolerance to the protocols, does not facilitate their use in a clinical setting.

However, there are distinct advantages to using a cycle ergometer, including greater ease of obtaining blood pressure, electrocardiogram, and other physiological measures; relative expense of a cycle ergometer (typically much less than a motorized treadmill); and practical considerations (a cycle ergometer occupies less space than a treadmill, is easier to move, and does not necessarily require electricity) (72). As well, there is a safety concern for older individuals on a treadmill, with respect to balance (72). In some instances, frail older adults may not be able to ambulate safely at the lowest speed possible on the treadmill, typically 1.0 mph (1.6 km/h) (62). The many practical advantages of the bicycle ergometer certainly justify further research using this modality for VO₂max testing in older adults.

Another possible avenue for future research addressing VO₂max testing in older adults is the investigation of an all-extremity ergometer, such as a recumbent cross-trainer or an elliptical trainer. Hass and coworkers (73) compared nine non-exercising control subjects with 17 individuals who participated in a thrice weekly progressive exercise program for 12 weeks using a total body recumbent stepper (mean age = 48.4 ± 6.4 years). The researchers found a significant 18% increase in VO₂max in the exercising group with no significant change in the control group. The use of an all-extremity ergometer might be particularly useful in the older adult population to maximize the muscle mass involved in exercise and to allow for compensations for either unilateral or bilateral extremity weakness or pathology (74).

Buchfuhrer and associates (75) suggested work rate increments should be selected appropriately to attain VO₂max in approximately 10 minutes (±2 min) with all exercise stress tests. Longer tests not only waste time and supply no additional information (76), but also risk a subject’s inability to achieve VO₂max criteria due to a decrease in motivation, increase in discomfort (complaints of low back pain on the treadmill and seat discomfort on the cycle ergometer), increase in body temperature, greater dehydration, or ventilatory muscle fatigue (75). Prolonged stage durations and/or rapid stage increases, typically seen in some of the protocols for younger individuals, may cause muscles to fatigue prior to the older subject obtaining a plateau in VO₂ (75,77). The protocols of Foster and colleagues (36) and Sidney and Shephard (59), having average test durations of 20 minutes and 15 minutes, respectively, may have yielded a lower percentage of individuals attaining VO₂max criteria than if the duration of exercise had been shorter.

Prior to developing their test protocol, Foster and colleagues (36) conducted a pilot study on older women and determined that a walking speed of 4.8 km/h at a 0% grade produced a negative psychological response. Although no further information regarding the exact nature of the response was provided, they did note that arrhythmias and HR near age-predicted maximum were elicited in apparently healthy individuals at the above-mentioned speed, even with a prior warm-up period. Although the limited details and the small number of participants in this study provide only a weak basis for a recommendation, these data, along with the concern for balance deficiencies and lack of familiarity with treadmill ambulation, may support the use of the protocols with slower initial speeds on the treadmill.

Limitations of the Critical Review

This review is limited to published data in the English language and is focused solely on studies of healthy individuals. However, as previously noted, many older adults have at least one, if not multiple, chronic medical conditions (27). Smaller work rate increases and a lower starting point may be particularly necessary for individuals with one or more health conditions. Future investigations could benefit from review of the available literature on VO₂max testing of individuals with various medical conditions.

Conclusion

There is extensive literature available on VO₂max testing in younger adults; however, this review has indicated that the research on VO₂max testing in individuals older than 65 years of age is limited. Additionally, the literature available on VO₂max testing in healthy older adults is flawed by a lack of detail with respect to protocol description as well as a lack of data on psychometric properties. The study by Thomas and coworkers (60) is the only one to compare three different VO₂max testing protocols using the same group of subjects. Their results suggest that a continuous ramp-like protocol with small increases in both speed and grade is perhaps most appropriate for healthy older men. This study provides
a strong basis for comparison of the appropriateness of different test protocols, however, there is limited generalizability to the greater population of older adults due to the exclusion of women. Based on this critical review, no specific recommendations could be made regarding an optimal protocol or modality for VO2max testing in healthy older adults. There is a pressing need for refinement of the few existing protocols for testing aerobic capacity in the older adult population, as well as the development of new protocols specifically applicable to older adults. Future investigation of studies comparing different testing modalities using similar protocols as well as comparing different protocols in the same group of older adults would be beneficial. This review was limited to describing VO2max testing protocols in healthy older adults. Further review of the literature is required to determine the appropriateness of maximal exercise testing protocols in older adults with various medical conditions. Finally, there is a need for agreement on whether the criteria for reaching VO2max should be redefined for older adults, and if the most appropriate test, for successful fulfillment of the criteria by all subjects, still needs to be developed. Perhaps a different definition of the VO2max criteria, such as using the (0.7 x age) equation for maximum HR described by Tanaka and associates (68), would allow more individuals to meet the criteria, or perhaps VO2peak data provide sufficient information for our current purposes in using maximal aerobic capacity data.

ACKNOWLEDGMENTS

We acknowledge with thanks financial support provided by the Ontario Provincial Rehabilitation Program and the Physiotherapy Foundation of Canada.

Address correspondence to Dr. Tom Overend, School of Physical Therapy, University of Western Ontario, London, Ontario, Canada N6G 1H1. E-mail: toverend@uwo.ca

REFERENCES


