New Data-based Cutoffs for Maximal Exercise Criteria across the Lifespan

JONATHAN WAGNER[1], MAX NIEMEYER[2], DENIS INFANGER[1], TIMO HINRICHS[1], LUKAS STREESE[1], HENNER HANSSEN[1], JONATHAN MYERS[3], ARNO SCHMIDT-TRUCKSÄSS[1], and RAPHAEL KNAIER[1]

[1]*Division Sports and Exercise Medicine, Department of Sport, Exercise and Health, University of Basel, Basel, SWITZERLAND;* [2]*Department Medicine, Training and Health, Institute of Sport Science and Motology, Philipps-University Marburg, Marburg, GERMANY; and* [3]*Cardiology Division, VA Palo Alto Health Care System and Stanford University, Palo Alto, CA*

ABSTRACT

WAGNER, J., M. NIEMEYER, D. INFANGER, T. HINRICHS, L. STREESE, H. HANSSEN, J. MYERS, A. SCHMIDT-TRUCKSÄSS, and R. KNAIER. New Data-based Cutoffs for Maximal Exercise Criteria across the Lifespan. *Med. Sci. Sports Exerc.*, Vol. 52, No. 9, pp. 1915–1923, 2020. **Purpose:** To determine age-dependent cutoff values for secondary exhaustion criteria for a general population free of exercise limiting chronic conditions; to describe the percentage of participants reaching commonly used exhaustion criteria during a cardiopulmonary exercise test (CPET); and to analyze their oxygen uptake at the respective criteria to quantify the impact of a given criterion on the respective oxygen uptake ($\dot{V}O_2$) values. **Methods:** Data from the COmPLETE-Health Study were analyzed involving participants from 20 to 91 yr of age. All underwent a CPET to maximal voluntary exertion using a cycle ergometer. To determine new exhaustion criteria, based on maximal respiratory exchange ratio (RER_{max}) and age-predicted maximal HR (APMHR), one-sided lower tolerance intervals for the tests confirming $\dot{V}O_2$ plateau status were calculated using a confidence level of 95% and a coverage of 90%. **Results:** A total of 274 men and 252 women participated in the study. Participants were nearly equally distributed across age decades from 20 to >80 yr. A $\dot{V}O_2$ plateau was present in 32%. There were only minor differences in secondary exhaustion criteria between participants exhibiting a $\dot{V}O_2$ plateau and participants not showing a $\dot{V}O_2$ plateau. New exhaustion criteria according to the tolerance intervals for the age group of 20 to 39 yr were: $RER_{max} \geq 1.13$, $APMHR_{210}$ − age ≥ 96%, and $APMHR_{208}$ × 0.7 age ≥ 93%; for the age group of 40 to 59 yr: $RER_{max} \geq 1.10$, $APMHR_{210}$ − age ≥ 99%, and $APMHR_{208}$ × 0.7 age ≥ 92%; and, for the age group of 60 to 69 yr: $RER_{max} \geq 1.06$, $APMHR_{210}$ − age ≥ 99%, and $APMHR_{208}$ × 0.7 age ≥ 89%. **Conclusions:** The proposed cutoff values for secondary criteria reduce the risk of underestimating $\dot{V}O_{2max}$. Lower values would increase false-positive results, assuming participants are exhausted although, in fact, they are not. **Key Words:** $\dot{V}O_{2max}$, $\dot{V}O_{2peak}$, EXHAUSTION CRITERIA, $\dot{V}O_2$ PLATEAU, CRF

The maximal volume of oxygen uptake per minute ($\dot{V}O_{2max}$) measured by cardiopulmonary exercise testing (CPET) is a strong risk factor for mortality and morbidity and outperforms other traditional risk factors (1,2). $\dot{V}O_{2max}$ is an important measurement that should be considered a vital sign as outlined in a recent statement by the American Heart Association (3). Achieving physical exhaustion and a predefined physiological limit is necessary to determine $\dot{V}O_{2max}$, but making the distinction between those participants who have reached this limit and those who have not remains a challenge (4). To maximize the signal-to-noise ratio, it is crucial to measure $\dot{V}O_{2max}$ with sufficient rigor. Therefore, an accurate determination of an individual's physiological limits is important to: 1) utilize $\dot{V}O_{2max}$ as primary outcome in randomized controlled trials; 2) apply $\dot{V}O_{2max}$ in the clinical setting as a vital sign, to stratify risk, or to guide therapeutic strategies; or 3) as a criterion for clinical decision making, for example, for heart transplantation (5,6).

Historically, the criterion standard to distinguish between $\dot{V}O_{2peak}$ and $\dot{V}O_{2max}$ is the occurrence of a $\dot{V}O_2$ plateau. This criterion, however, has several limitations. First, it is not straightforward to apply as relatively complex data analyses are required and, second, numerous definitions have been proposed which has led to a great deal of controversy (7). Most importantly, the occurrence of a $\dot{V}O_2$ plateau in a healthy non-athletic population or among patients with disease is rather low even though participants have performed an exercise test with maximal effort (8,9) The occurrence of a $\dot{V}O_2$ plateau varies widely depending on the definition used but it is clearly below 50% in the general population (10,11).

Another suggested method from the field of performance sports that has arisen in recent years is the use of verification tests. These tests, however, are not feasible in daily clinical practice or in large-scale studies because they are very

Address for correspondence: Arno Schmidt-Trucksäss, M.D., M.A., Division of Sports and Exercise Medicine, Department of Sport, Exercise and Health, University of Basel, Birsstrasse 320B, 4052 Basel, Switzerland; E-mail: arno.schmidt-trucksaess@unibas.ch.
A.S.-T. and R.K. share senior authorship. J.M. and T.H. participated in the writing, review and editing, and participated in the methodology.
Submitted for publication January 2020.
Accepted for publication March 2020.

0195-9131/20/5209-1915/0
MEDICINE & SCIENCE IN SPORTS & EXERCISE®

DOI: 10.1249/MSS.0000000000002344

APPLIED SCIENCES

time-consuming and seem to be of limited benefit (12,13) (Wagner et al., unpublished data). Therefore, for the large proportion of individuals not reaching a $\dot{V}O_2$ plateau, secondary criteria are needed to minimize the risk of an underestimation of the true $\dot{V}O_{2max}$ (7,14,15). Secondary criteria to minimize this bias are rarely reported or the criteria chosen conservatively with relatively low cutoff levels (4,7). The most common secondary exhaustion criteria are the maximum RER (RER_{max}), maximum heart rate (HR_{max}), maximum RPE (RPE_{max}), and maximum concentration of blood lactate (BL_{max}). Values for these parameters at which maximal physiologic effort is accepted for middle-age participants include RER values >1.00 (16), >1.05, and >1.1; 85% (17) to 100% of the age-predicted HR_{max}; ≥17 to ≥19 RPE_{max}; and from ≥4 to ≥10 $mmol \cdot L^{-1}$ for BL_{max} (4,7,8).

Defining optimal criteria for a maximal physiologic response requires a balance between assuming that participants have reached this point when they have not (i.e., low criteria, type I error) and declaring participants have not reached this point when though they have (i.e., high criteria, type II error). Therefore, this study analyzed data from the COmPLETE-Health study (18) including a large number of individuals without exercise-limiting chronic disease conditions across a broad age spectrum with several aims. The first aim was to determine age-dependent cutoff values using tolerance intervals based only on those tests where $\dot{V}O_2$ plateaued. The second aim was to establish a multiparameter score to improve the performance of a single criterion. The third aim was to provide a descriptive analysis, based on the data of all participants, of the percentage of participants reaching commonly used exhaustion criteria during a CPET. Finally, we analyzed participants' oxygen uptake values at each criterion to quantify the impact of a chosen criterion on the respective $\dot{V}O_2$ values.

METHODS

Study Design

Population and recruitment. The COmPLETE-Study is a cross-sectional single-center study performed in Switzerland. Participants were healthy men and women examined in 2018 and 2019. Participants met several inclusion criteria, such as being between 20 and 100 yr, having a body mass index <30 $kg \cdot m^{-2}$, and being nonsmokers or ex-smokers for more than 10 yr. Exclusion criteria included any kind of manifest exercise-limiting chronic disease (e.g., myocardial infarction, stroke, heart failure, lower-extremity artery disease, cancer, diabetes, clinically apparent renal failure, severe liver disease, chronic bronchitis Global Initiative for Chronic Obstructive Lung Disease stages II to IV, osteoporosis); hypertensive blood pressure >160/100 mm Hg; compromising orthopedic problems; Alzheimer's disease or any other form of dementia inability to follow the procedures of the study (e.g., due to language problems, psychological disorders, dementia of the participant); diseases regarded as an absolute contraindication for maximal exertion.

The exact recruitment procedure and the full list of inclusion and exclusion criteria can be found elsewhere (18).

APPLIED SCIENCES

Setting. The study was carried out at the Department of Sport, Exercise, and Health at the University of Basel, Switzerland, was funded by the Swiss National Science Foundation (grant 182815) and approved by the Ethics Committee of Northwestern and Central Switzerland (EKNZ 2017-01451). Written informed consent was obtained from all participants before the start of the study.

Participants in the COmPLETE-Health Study underwent a battery of tests including CPET; the detailed testing procedure can be found in the study protocol (18). The participants were requested to not perform any sporting activities 24 h before examination and to abstain from alcohol for 24 h and from caffeine for 4 h before the examination.

Acquisition of participant characteristics. Smoking status was assessed by telephone interview before the appointment, whereas physicians reviewed medical history and medications by questionnaire on site. Further, a 12-lead resting electrocardiogram was acquired and reviewed by a physician immediately before the exercise test. Maximal heart rate data of participants taking beta blockers (n = 12) were excluded from the analysis. Height and body weight were measured to the nearest 0.5 cm and 0.1 kg, respectively, and the body mass index was calculated. To measure the body fat content and lean body mass, a four-segment bioelectrical impedance analysis was conducted (Inbody 720; Inbody Co. Ltd., Seoul, South Korea). Resting systolic and diastolic blood pressures and HR were measured with the participant in the supine position using a noninvasive vascular screening system (VaSera VS-1500N; Fukuda Denshi, Tokyo, Japan).

Cardiopulmonary exercise testing. An exercise test to maximal voluntary exertion using an electromagnetically braked cycle ergometer (Ergoselect 200; Ergoline, Bitz, Germany) was performed according to one of the following five ramp protocols, depending on the subject tested: i) a 3-min warm-up either unloaded, a load of 10 or 20 W for protocols 1 to 3, or a load of 50 W for protocols 4 and 5 followed by ii) a ramp protocol with a linear workload increases of 7, 10, 15, 20, or 30 $W \cdot min^{-1}$ for protocols 1 to 5, respectively. A 3-min recovery phase was maintained at the same workload as the warm-up. The protocol was chosen to achieve a duration of 10 min, and the participant was excluded when the exercise time was not between 6 and 18 min (19,20). Pedaling cadence was freely chosen by participants but was required to be more than 60 rpm.

Gas exchange and ventilatory variables were analyzed breath-by-breath continuously using a computer-based system (MetaMax 3B; Cortex Biophysik GmbH, Leipzig, Germany). Every test was preceded by a resting period of 3 min to reach steady-state conditions. A trained and certified sports scientist continuously supervised the examination, and a physician was available upon request. In the absence of clinical symptoms or electrocardiographic abnormalities, all tests were continued until maximal exertion (i.e., volitional exertion, dyspnea, or fatigue). Heart rate was measured with a 12-lead electrocardiography (Custo med GmbH, Ottobrunn, Germany). The capillary blood lactate concentration from the earlobe was measured at rest, at maximum performance, and at 1 and

3 min after the end of the exercise test. Rating of perceived exertion (Borg Scale) was applied every 2 min during the test and immediately after termination of the ramp. The examiners were instructed to keep the following priorities upon termination of the ramp: continue to recovery phase in the testing software, collecting capillary blood samples, and applying RPE. Before and during the test, participants were verbally encouraged to reach maximal exhaustion. All tests were performed in controlled humidity and temperature conditions (21). Before each test, the equipment was calibrated in standard fashion with reference gas and known volume.

$\dot{V}_E$ (L·min^{-1}), $\dot{V}O_2$ (mL·min^{-1}), and $\dot{V}CO_2$ (mL·min^{-1}) were acquired on a breath-by-breath basis and averaged over 10-s intervals. $\dot{V}O_{2peak}$ was defined as the highest 30-s average of $\dot{V}O_2$ at any point during the test.

Blood lactate concentration (mmol·L^{-1}) was measured from 10 μL of capillary blood drawn from the ear. The analysis of blood lactate concentrations was done via the SuperGL Ambulance (Hitado Diagnostic Systems, Moehnesee, Germany) immediately after the last blood sample was drawn. Only a small number of well-trained assessment staff performed and supervised the CPET, and standardized procedures and instructions were used to ensure equal testing conditions for all participants.

Because there is no validated plateau definition for ramp tests with an increment rate of less than 20 W·min^{-1}, we determined the occurrence of a $\dot{V}O_2$ plateau with two common calculation approaches. This was done to calculate the coefficient of agreement and to ensure that the selection of participants reaching a $\dot{V}O_2$ plateau, and therefore, the $\dot{V}O_{2max}$, was not dependent on the definition. The first calculation involved an increase in $\dot{V}O_2 < 50\%$ of the expected increase between the last and the second-to-last minute of the CPET (10). The expected increase $\dot{V}O_2$ was calculated for each of the five protocols based on the assumption that $\dot{V}O_2$ increases 10.0 mL·min^{-1}·W^{-1} (22,23). The second calculation involved an increase in $\dot{V}O_2$ during the final 2 min which was <50% of the corresponding increase in the submaximal intensity domain. For the latter, we calculated the slope of the $\dot{V}O_2$–workload relationship during submaximal exercise using linear regression analysis after excluding the first and the final 2 min of exercise. The linear regression was extrapolated to the end of the ramp test, and the difference between the measured and the calculated $\dot{V}O_{2max}$ was used to determine the occurrence of a $\dot{V}O_2$ plateau as previously described (24). The second definition accounts for the individual increase in $\dot{V}O_2$ in the submaximal intensity domain, which is mainly affected by age and fitness (25,26). Therefore, the second definition is advantageous with respect to the heterogeneous cohort in the present study and was used for the subsequent analyses.

For the $\dot{V}O_{2max}$ criteria using maximal HR, two different definitions for age-predicted maximal HR (APMHR) were used: i) $APMHR_{210}$ constituting 210 minus the participant's age, where the conventional formula for age-predicted HR, 220 minus age was adapted to 210 minus age to consider the lower muscle mass involved in a cycle ergometer test, which results in a lower maximum HR in comparison to treadmill tests (27); and ii) $APMHR_{208}$, which is the formula recommended by Tanaka et al. (28) ($208 - 0.7 \times$ age in years).

Statistical Analysis

We compared the five secondary exhaustion criteria (RER_{max}, $APMHR_{210}$, $APMHR_{208}$, RPE_{max}, and BL_{max}) and several performance parameters between participants showing a $\dot{V}O_2$ plateau (according to the second definition) and those showing no $\dot{V}O_2$ plateau using an analysis of covariance with age and sex as covariates. The coefficient of agreement between the two definitions used to calculate the $\dot{V}O_2$ plateau

TABLE 1. Descriptive characteristics of the study population separated by sex

	n	Male	*n*	Female
Participants, *n* (%)		274 (52)		252 (48)
Age (yr)	274	53.5 ± 19.8 (20–91)	252	54.3 ± 19.4 (20–89)
Height (cm)	274	177.0 ± 7.3 (156.0–196.5)	252	165.6 ± 7.1 (149.0–186.0)
Body mass (kg)	274	76.6 ± 9.7 (51–104)	252	62.6 ± 8.8 (43–90)
BMI (kg·m^{-2})	274	24.4 ± 2.5 (19.0–30.0)	252	22.8 ± 2.7 (17–30)
Rest systolic BP, mm Hg	274	130 ± 12 (95–159)	252	124 ± 15 (87–160)
Rest diastolic BP, mm Hg	274	79 ± 9 (57–100)	252	76 ± 9 (48–97)
HR at rest (bpm)	272	60 ± 10 (38–129)	247	62 ± 10 (44–104)
Smoking status, *n* (%)				
Never smoked (%)	274	220 (80.3)	252	200 (79.4)
Ex-smokers >10 yr (%)	274	54 (19.7)	252	52 (20.6)
Medication, *n* (%)				
Antihypertensives (%)	274	27 (9.9)	252	19 (7.5)
Beta blockers (%)	274	7 (2.6)	252	5 (1.9)
Performance				
Pmax (W)	274	246 ± 82 (75–468)	252	157 ± 58 (45–299)
$\dot{V}O_{2max}$ absolute (L·min^{-1})	274	2.91 ± 0.80 (1.26–4.86)	252	1.88 ± 0.56 (0.70–3.55)
$\dot{V}O_{2max}$ relative (mL·kg^{-1}·min^{-1})	274	38.3 ± 10.5 (16.4–65.1)	252	31.3 ± 8.6 (13.0–53.5)
Exhaustion criteria				
RER_{max}	274	1.17 ± 0.08 (0.93–1.41)	252	1.17 ± 0.08 (0.93–1.47)
HR_{max} (bpm)	265	171 ± 22 (96–210)	242	167 ± 20 (97–203)
RPE_{max}	234	19.3 ± 0.9 (15–20)	219	19.0 ± 1.2 (15–20)
BL_{max} (mmol·L^{-1})	241	8.6 ± 2.9 (2.4–16.2)	214	7.1 ± 2.6 (1.7–13.7)

Data are presented as mean ± standard deviation (minimum, maximum) if not stated otherwise.
BMI, body mass index; BP, blood pressure; Pmax, maximal power.

TABLE 2. New recommendations for secondary exhaustion criteria.

Age Group, yr	RER_{max}	APMHR 210 − age	APMHR 208 − 0.7 × age	Multiparameter- Score* PCA, First Component (APMHR [210 − age], RER & Age)
20–39	1.13	96%	93%	−1.885
40–59	1.10	99%	92%	
60–69	1.06	99%	89%	

*To use the provided criteria of the multiparameter score, the standardized *z*-values for RER_{max}, $APMHR_{210}$, and age must be calculated. These *z* variables then must be multiplied with the loadings −0.489, 0.594, and 0.638 for RER_{max}, $APMHR_{210}$, and age, respectively. Finally, the results need to be summed and then compared to the values in the table to judge maximal exhaustion.

prevalence was calculated using Gwet's AC_1 (29). Coefficients of agreement were interpreted as follows: <0.2, poor; 0.2 to 0.4, fair; 0.4 to 0.6, moderate; 0.6 to 0.8, good; and 0.8 to 1, very good (30).

Mean differences (MD) and 95% confidence intervals (95% CI) for the four secondary exhaustion criteria were calculated between male and female participants within the four age subgroups using an analysis of covariance approach with age selected as a covariate.

To determine new exhaustion criteria, one-sided lower tolerance intervals were calculated using a confidence level of 95% and a coverage of 90%. For these analyses, only those tests were considered in which $\dot{V}O_{2max}$ was confirmed by the presence of a $\dot{V}O_2$ plateau (according to the second definition). Therefore, the newly defined secondary exhaustion criteria were unlikely to produce type II errors. The sample was divided into four age categories: 20 to 39 yr, 40 to 59 yr, 60 to 69 yr, and 70 yr or older. The rationale behind this allocation lies in the trajectories of the RER_{max}, $APMHR_{210}$ and $APMHR_{208}$ criteria over age. Because there is wider variance in these criteria after the age of 70 yr, the last group where robust criteria could be calculated was cut to 60 to 69 yr instead of the 20-yr age grouping of 60–79 yr. The calculation of the tolerance intervals for all variables except RPE_{max} assumed that the variables were reasonably modeled by a normal distribution, which were checked using Q–Q-plots. The tolerance intervals for RPE_{max} were calculated using nonparametric methods. To improve the performance of a single secondary exhaustion criterion, a principal component analysis (PCA) was performed with the two secondary exhaustion criteria RER and $APMHR_{210}$ and age using all data up to 69 yr. For the PCA, only data of participants up to 69 yr were included due to the large variation in the criteria reached in participants older than 69 yr. Because these variables are on different scales, the correlation matrix instead of the covariance matrix was used to extract the principal components. To minimize the influence of extreme values on the Pearson correlation, the robust correlation matrix for the PCA was employed.

Descriptive statistics were used to compare the number of participants reaching the following exhaustion criteria: RER (≥1.0, ≥1.05, ≥1.10, ≥1.15, and new criteria), $APMHR_{210}$ (≥85%, ≥90%, ≥95%, ≥100%, and new criteria), and $APMHR_{210}$ (≥85%, ≥90%, ≥95%, ≥100%, and new criteria) and the respective $\dot{V}O_2$ reached at that point in time. Further,

TABLE 3. Values for $\dot{V}O_2$ at the time point when exhaustion is reached based on different criteria by different age categories.

Age (yr)	20–39 (*n* = 152)			40–59 (*n* = 155)		
Criteria	Participants Reaching Criteria (%)	Mean ± SD (95% CI) $\dot{V}O_2$ (mL·kg^{-1}·min^{-1})	Mean ± SD (95% CI) % of $\dot{V}O_{2peak}$	Participants Reaching Criteria (%)	Mean ± SD (95% CI) $\dot{V}O_2$ (mL·kg^{-1}·min^{-1})	Mean ± SD (95% CI) % of $\dot{V}O_{2peak}$
RER_{max}						
1.00	99	29.4 ± 8.2 (28.0–30.7)	68 ± 12 (66–70)	100	27.4 ± 9.2 (25.9–28.8)	73 ± 14 (70–75)
1.05	99	34.2 ± 8.9 (32.735.6)	79 ± 12 (77–81)	97	30.5 ± 9.0 (29.0–32.0)	81 ± 12 (79–83)
1.10	95	38.2 ± 8.8 (36.8–39.7)	89 ± 10 (87–91)	93	33.1 ± 8.8 (31.7–34.6)	89 ± 9 (87–90)
1.15	82	40.7 ± 8.4 (39.2–42.2)	95 ± 7 (93–96)	75	34.4 ± 7.8 (33.0–35.9)	95 ± 6 (93–95)
New Criteria	91	40.0 ± 8.5 (38.6–41.4)	93 ± 8 (92–94)	93	33.1 ± 8.8 (31.7–34.6)	89 ± 9 (87–90)
HR_{max} APMHR 210 − age						
85%	99	29.1 ± 6.5 (28.1–30.2)	68 ± 9 (66–69)	99	24.6 ± 6.6 (23.6–25.7)	67 ± 12 (65–69)
90%	98	32.5 ± 7.0 (31.4–33.6)	76 ± 10 (74–77)	99	27.3 ± 7.1 (26.2–28.4)	74 ± 12 (72–76)
95%	94	35.5 ± 7.5 (34.3–36.8)	82 ± 10 (80–84)	98	29.6 ± 7.8 (28.4–30.9)	80 ± 12 (78–82)
100%	83	39.0 ± 8.3 (37.6–40.5)	90 ± 11 (88–92)	94	32.1 ± 8.4 (30.7–33.5)	85 ± 13 (83–87)
New Criteria	92	36.3 ± 7.6 (35.1–37.6)	84 ± 8 (92–94)	95	31.7 ± 8.3 (30.3–33.0)	84 ± 13 (82–86)
HR_{max} APMHR 208 × 0.7 age						
85%	99	31.4 ± 6.8 (30.0–32.2)	72 ± 10 (71–74)	99	28.0 ± 7.4 (26.8–29.1)	76 ± 13 (74–78)
90%	95	34.6 ± 7.2 (33.4–35.7)	80 ± 11 (88–92)	96	30.9 ± 7.9 (29.6–32.2)	82 ± 12 (80–84)
95%	87	38.0 ± 7.8 (36.6–39.3)	88 ± 11 (86–89)	86	33.5 ± 8.7 (32.0–34.0)	88 ± 12 (86–90)
100%	61	40.5 ± 8.5 (38.7–42.2)	92 ± 11 (90–94)	64	34.4 ± 8.7 (32.7–36.2)	91 ± 12 (89–93)
New Criteria	91	37.3 ± 7.7 (36.0–38.6)	86 ± 11 (84–88)	94	33.0 ± 8.5 (31.5–34.4)	85 ± 12 (83–87)
RPEmax						
≥17	100			100		
≥18	99			95		
≥19	94			87		
20	55			49		
BLmax						
4 mmol·L^{-1}	100			100		
6 mmol·L^{-1}	98			90		
8 mmol·L^{-1}	85			57		
10 mmol·L^{-1}	57			24		

New criteria according to Table 2 for the age group of 20 to 39 yr are: RER_{max} ≥ 1.13, $APMHR_{210}$ ≥ 96%, and $APMHR_{208}$ 93%; for the age group of 40 to 59 yr: RER_{max} ≥ 1.10, $APMHR_{210}$ ≥ 99%, and $APMHR_{208}$ 92%; and, for the age group of 60 to 69 yr: RER_{max} ≥ 1.06, $APMHR_{210}$ ≥ 99%, and $APMHR_{208}$ 89%. For the analysis of the 20–69 group, the respective criteria of the 3 subgroups were used.

RPE_{max} (≥17, ≥18, ≥19, 20) and BL_{max} (4, 6, 8, 10 mmol·L^{-1}) were assessed to evaluate whether they were reached by the participants.

Descriptive data are presented as means and standard deviations. The Statistical Package for the Social Sciences version 26 for Windows software program (IBM Corp., Armonk, NY) and R version 3.6.1 (R Foundation for Statistical Computing, Vienna, Austria) were used for all analyses. Because this analysis was not the primary aim of the COmPLETE study, no sample-size calculation for this specific research question was conducted.

RESULTS

Participant characteristics. A total of 526 participants were included in the study (274 men and 252 women). Ten of the initial 536 participants were excluded because the exercise test was not between 6 and 18 min. Participants were equally distributed across age decades from 20 to greater than 80 yr, with at least 75 participants representing every decade from 20 to 80 yr and 66 participants included in the category 80 yr or greater. Participant characteristics from medical examinations and CPET are presented in Table 1.

In 20 participants, it could not be determined if a plateau occurred or not due to insufficient data quality at the end of the test. In the remaining 506 tests, a $\dot{V}O_2$ plateau was present in 153 (30%) and 164 (32%) according to first and second definition of the $\dot{V}O_2$ plateau (defined in the methods section), respectively. The two $\dot{V}O_2$ plateau definitions showed a good level of agreement with a Gwet's AC_1 of 0.76 (95% CI, 0.70–0.82).

There were only minor differences in the secondary exhaustion criteria between participants showing a $\dot{V}O_2$ plateau (according to the second criteria) and participants not showing a $\dot{V}O_2$ plateau. The MD and 95% CI for RER_{max} were 0.012 (−0.002 to 0.027); for $APMHR_{210}$, 1.33% (−0.39 to 1.52); for $APMHR_{208}$, 1.08% (−0.21 to 2.37); for RPE_{max}, 0.03 (−0.15 to 0.22); and 0.4 mmol·L^{-1} (0.0–0.8) for BL_{max}. There was little evidence that the $\dot{V}O_{2peak}$ values of participants showing a $\dot{V}O_2$ plateau were higher relative to those not reaching a plateau (MD: 0.77 mL·kg^{-1}·min^{-1} [−0.57 to 2.13]). Peak power (W), maximal ventilation (L·min^{-1}), and maximal breathing frequency (breaths per minute) were, however, significantly higher among participants showing a $\dot{V}O_2$ plateau as compared to the other participants (MD [95% CI]: 23 W [14–32] 8.4 L·min^{-1} [3.9–12.9], and 5 breaths per minute [3–7]), respectively.

New data-based secondary exhaustion criteria. The tolerance intervals and, the resulting new suggested exhaustion criteria for the age groups 20 to 39 yr, 40 to 59 yr, and 60 to 69 yr are shown in Table 2.

Multiparameter score as an exhaustion criterion. The multiparameter exhaustion criterion score was the first principal component and explained 64.7% of the total variation in the three individual parameters RER_{max}, $APMHR_{210}$,

TABLE 3. *Continued.*

60–69 (*n* = 75)			20–69 (*n* = 382)			70+ (*n* = 144)		
Participants Reaching Criteria (%)	Mean ± SD (95% CI) $\dot{V}O_2$ (mL·kg^{-1}·min^{-1})	Mean ± SD (95% CI) % of $\dot{V}O_{2peak}$	Participants Reaching Criteria (%)	Mean ± SD (95% CI) $\dot{V}O_2$ (mL·kg^{-1}·min^{-1})	Mean ± SD (95% CI) % of $\dot{V}O_{2peak}$	Participants Reaching Criteria (%)	Mean ± SD (95% CI) $\dot{V}O_2$ (mL·kg^{-1}·min^{-1})	Mean ± SD (95% CI) % of $\dot{V}O_{2peak}$
99	24.9 ± 8.3 (23.0–26.8)	77 ± 13 (74–80)	99	27.7 ± 8.8 (26.8–28.6)	72 ± 14 (70–73)	92	22.1 ± 5.3 (21.1–23.0)	87 ± 11 (85–89)
97	27.4 ± 8.6 (25.3–29.4)	85 ± 12 (82–87)	98	31.4 ± 9.2 (30.4–32.3)	81 ± 12 (80–82)	75	23.7 ± 5.7 (22.6–24.8)	91 ± 9 (90–93)
84	29.5 ± 8.4 (27.5–31.7)	91 ± 8 (89–93)	92	34.6 ± 9.3 (33.6–35.5)	89 ± 9 (88–90)	51	25.6 ± 5.0 (24.4–26.8)	95 ± 6 (94–97)
61	30.8 ± 7.7 (28.5–33.0)	96 ± 5 (95–98)	75	36.6 ± 8.9 (35.5–37.6)	95 ± 6 (94–95)	61	25.9 ± 5.0 (24.2–27.6)	97 ± 4 (95–98)
97	28.0 ± 8.6 (26–30.0)	87 ± 10 (84–89)	93	34.7 ± 9.8 (33.7–35.8)	90 ± 9 (89–91)	—	—	—
99	21.2 ± 6.0 (19.8–22.6)	67 ± 15 (64–70)	99	25.7 ± 7.1 (25.0–26.5)	67 ± 12 (66–68)	96	16.6 ± 5.0 (15.8–17.4)	68 ± 16 (65–70)
95	23.2 ± 6.4 (21.7–24.7)	73 ± 14 (69–76)	98	28.6 ± 7.7 (27.8–29.3)	74 ± 12 (73–76)	94	18.6 ± 5.0 (17.8–19.5)	75 ± 15 (72–78)
95	25.3 ± 6.6 (23.8–26.9)	79 ± 13 (76–82)	96	31.1 ± 8.4 (30.2–32.0)	81 ± 12 (79–82)	92	20.1 ± 5.0 (19.3–21.0)	80 ± 15 (77–82)
91	27.0 ± 7.2 (25.3–28.8)	84 ± 11 (81–87)	89	33.7 ± 9.3 (32.7–34.7)	86 ± 12 (85–88)	84	21.2 ± 5.4 (20.3–22.2)	83 ± 14 (80–85)
93	26.5 ± 6.8 (24.9–28.1)	83 ± 12 (80–86)	93	32.5 ± 8.6 (31.6–33.4)	84 ± 12 (82–85)	c	—	—
94	25.1 ± 6.5 (23.5–26.6)	79 13 (76–82)	99	28.7 7.3 (28.0–29.4)	75 ± 12 (74–76)	85	20.4 5.1 (19.5–21.3)	80 15 (77–83)
84	27.6 ± 7.0 (25.8–29.3)	84 ± 11 (81–87)	93	31.8 ± 7.9 (30.0–32.6)	82 ± 11 (80–83)	77	22.5 ± 5.5 (21.5–23.6)	86 ± 12 (84–89)
76	30.0 ± 7.4 (28.0–32.0)	89 ± 9 (87–92)	84	34.7 ± 8.6 (33.8–35.7)	88 ± 11 (87–89)	65	24.0 ± 6.0 (22.8–25.3)	89 ± 14 (87–92)
64	31.3 ± 8.1 (29.0–33.6)	92 ± 8 (90–95)	63	36.1 ± 9.2 (35.0–37.3)	92 ± 11 (90–93)	51	25.7 ± 6.3 (24.2–27.2)	94 ± 10 (87–92)
91	26.6 ± 6.8 (25.0–28.2)	83 ± 12 (80–86)	92	33.5 ± 8.8 (33.5–34.4)	85 ± 12 (84–86)	—	—	—
99			100			92		
97			97			77		
85			90			64		
38			49			24		
97			99			74		
73			90			34		
41			65			9		
12			35			2		

APPLIED SCIENCES

and age. The loadings of the first component were −0.489, 0.594, and 0.638 for these three parameters, respectively.

Descriptive analysis of secondary exhaustion criteria. Table 3 shows the percentage of participants reaching the different exhaustion criteria defined in the literature and for the new calculated and proposed criteria. Among subjects between 20 and 69 yr of age, the lower criteria of RER_{max} 1.0 and 1.05, 85% and 90% $APMHR_{210}$, and 85% $APMHR_{208}$ were reached by almost all participants (≥98%), suggesting that these values were highly unlikely to cause type II errors in this population. Instead, these criteria may be more likely to produce type I errors with the mean % $\dot{V}O_{2peak}$ values being between 67% and 85% only. In the group of participants 70 yr or older, even the lowest criteria for RER_{max} 1.00, 85% $APMHR_{210}$, and 85% $APMHR_{208}$ were not reached by 8%, 4%, and 15% of these participants, respectively. The new criteria for RER_{max}, $APMHR_{210}$, and $APMHR_{208}$ were reached by 93%, 93%, and 92% of participants between 20 and 69 yr, respectively, who achieved, on average, 90%, 84%, and 85% of their individual $\dot{V}O_{2peak}$.

Figures 1 and 2 illustrate the tradeoff between choosing a criterion that is too low and accepting $\dot{V}O_2$ values that are not maximal and thus excluding participants as not reaching physiological limits (which is either correct as they have not reached their $\dot{V}O_{2max}$ or not correct, reflecting false-negative cases in that they achieved $\dot{V}O_{2max}$). All data were used to create Table 3 and Figures 1 and 2, including both those from the participants reaching a plateau and those who did not.

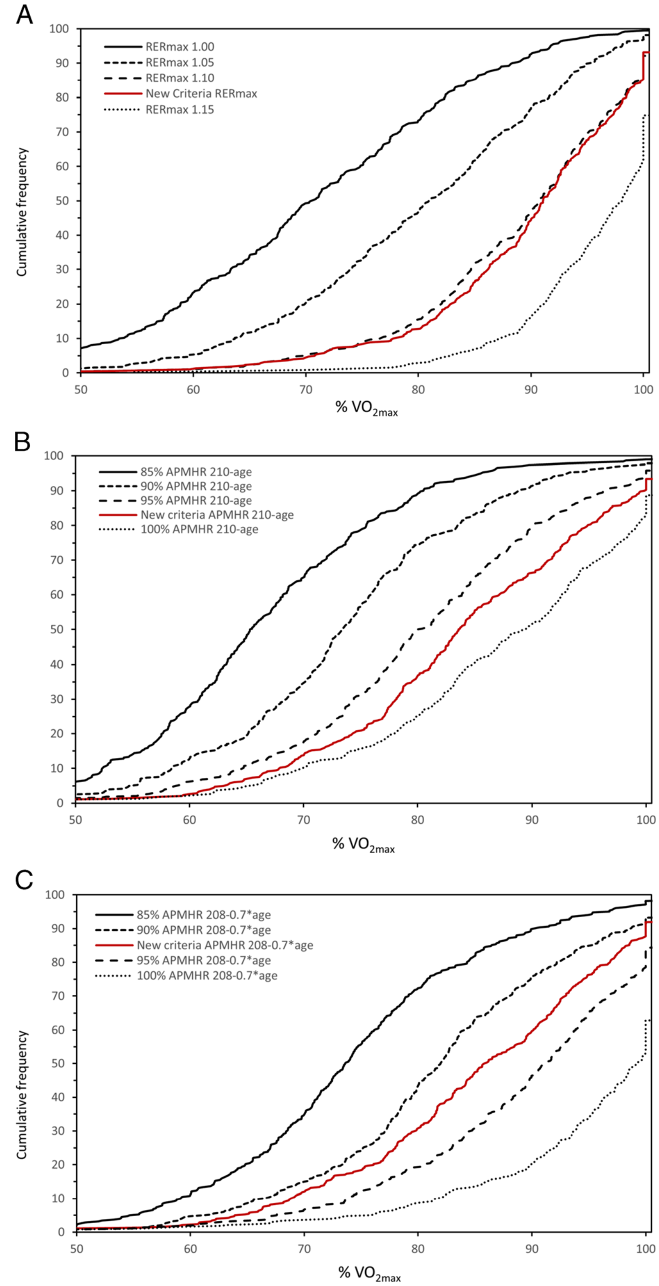

FIGURE 1—Cumulative frequency of subjects between 20 and 69 yr who satisfied the RER (A), $APMHR_{210}$ (B), and $APMHR_{208}$ (C) criteria at increasing percentages of $\dot{V}O_{2peak}$.

DISCUSSION

This comparatively large study provides data-based optimal secondary exhaustion criteria for different age groups to optimize the evaluation of $\dot{V}O_{2max}$. The suggested criteria are RER 1.13, 1.10, or 1.06; 96%, 99%, or 99% $APMHR_{210}$; or 93%, 92%, or 89% $APMHR_{208}$ for the age groups 20 to 39 yr, 40 to 59 yr, and 60 to 69 yr, respectively. These numbers differ clearly from previously used cutoffs and our results show that higher criteria need to be applied.

Plateau occurrence. The $\dot{V}O_2$ plateau prevalence of 32% in the current study is comparable to findings in previous studies. Lucia et al. reported that the $\dot{V}O_2$ plateau prevalence rates were 24% and 47% in sedentary male participants and professional cyclists, respectively (10). Another large-scale study of CPET performed on a treadmill reported a similar prevalence of $\dot{V}O_2$ plateau of 41% for women and 42% for men (8). The prevalence of $\dot{V}O_2$ plateau is dependent on training status, protocol, exercise mode, sampling method, and most importantly, on the definition used (23,31–33). To our knowledge, no previous studies have investigated the prevalence of $\dot{V}O_2$ plateau during CPET performed on a cycle ergometer in a comparable large, healthy, nonathletic population over the age spectrum of 20 to older than 80 yr.

RER and APMHR. Lower exhaustion criteria than those recommended in Table 2 can underestimate $\dot{V}O_{2max}$ by as much as 33% on a group level. In the age group of 20 to 39 yr for example, choosing an RER 1.00 would lead to an underestimation of maximal oxygen uptake by at least 20%, and the average % $\dot{V}O_{2peak}$ reached in that group was 68% ± 12%. Using the new criteria reduces the underestimation in the younger subsample of the study population (20–69 yr) to 7% on average with 93% $\dot{V}O_{2peak}$ reached. Another extreme example leading to large type I errors is using the criterion 85% $APMHR_{210}$ in the age group of 40 to 59 yr. Using this criterion resulted in a mean value of 67% $\dot{V}O_{2peak}$ compared with the new criterion of 99% $APMHR_{210}$, resulting in a mean value of 83% $\dot{V}O_{2peak}$. Similar results can be observed in the study by Knaier et al. involving 70 athletes (34). These athletes

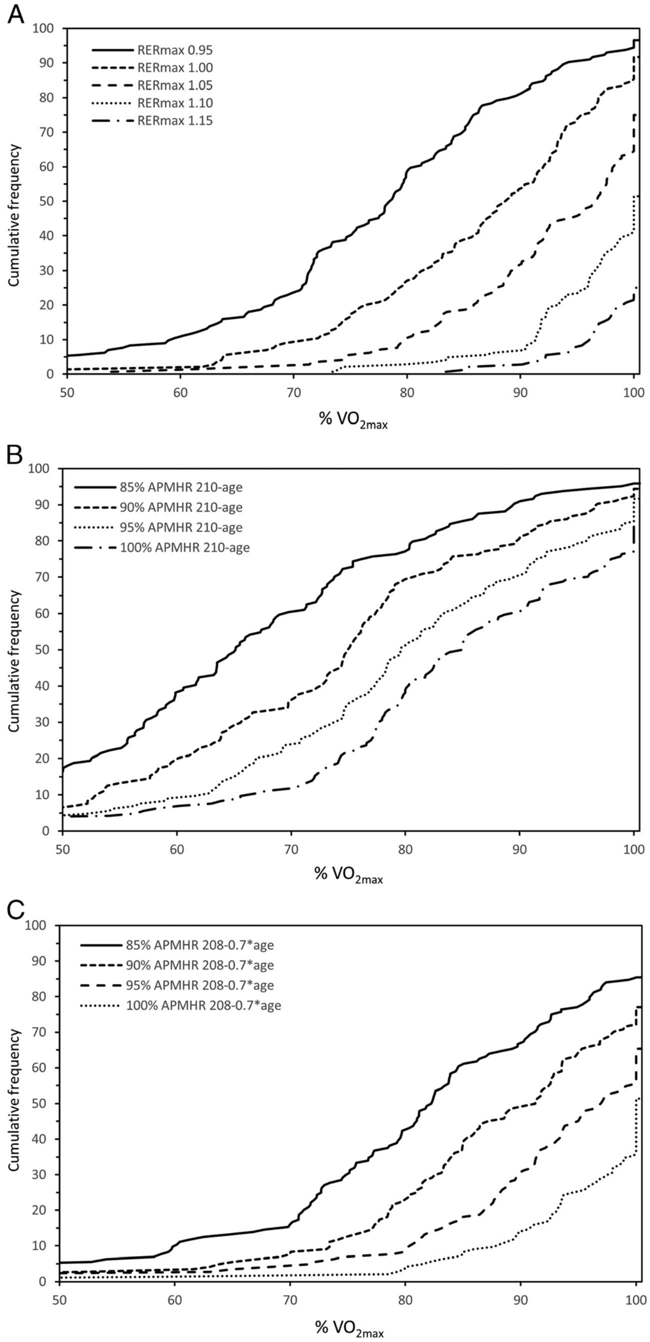

FIGURE 2—Cumulative frequency of subjects 70+ yr who satisfied the RER (A), $APMHR_{210}$ (B), and $APMHR_{208}$ (C) criteria at increasing percentages of $\dot{V}O_{2peak}$.

achieved only 93% or less of their actual $\dot{V}O_{2peak}$ according to the criterion of RER 1.05 and achieved 83% or less for the criterion of 90% $APMHR_{210}$. In older age groups, choosing a low criterion does not have as large an effect on the % $\dot{V}O_{2peak}$ reached as seen in the group age 20 to 39 yr. The underestimation is, however, still substantial (Table 3). Using the proposed criteria of RER 1.06 relative to the criterion of RER 1.0 in the age group of 60 to 69 yr improves the % $\dot{V}O_{2peak}$ reached from 77% to 87% without increasing the risk of type II error to more than 5%.

In the age group 70 yr and older, a large spread of criteria reached is observed. Low criteria, such as RER 1.0 and 1.05 or 85% and 90% $APMHR_{208}$, were not achieved by several participants, but they, nevertheless, exhibited a $\dot{V}O_2$ plateau. Secondary exhaustion criteria do not seem to work well in the age group older than 70 yr as the spread of criteria reached is quite large among the participants fulfilling the criterion standard criteria and reaching a $\dot{V}O_{2pleateau}$. Optimal criteria for this group could not be defined as criteria based on tolerance intervals with a 95% confidence level and a coverage of 0.9 were very low, with RER 0.95, 98% $APMHR_{210}$, or 83% $APMHR_{208}$. These cutoff values would most likely result in large type I errors, as shown by the % $\dot{V}O_{2peak}$ values reached for these criteria presented in Table 2 and the cumulative frequencies in Figure 1 indicate.

BL_{max} and RPE_{max}. Tolerance intervals for BL_{max} and RPE were 6.7, 5.3, 4.0, and 1.8 and 19, 17, 18, and 16 for the age groups 20 to 39 yr, 40 to 59 yr, 60 to 69 yr, and 70 yr or older, respectively. BL_{max} was highly variable between participants. Lactate concentration ranged from 1.7 to 16.2 $mmol \cdot L^{-1}$, making it impossible to set a secondary $\dot{V}O_{2peak}$ criterion even when the study cohort is stratified into the four age categories. Further, a previous study showed that the reliability of BL_{max} is poor and displayed a high degree of variance during the day (34).

The suggested RPE criteria for 20- to 39-yr-old participants, RPE 19, are in line with the results by Knaier et al. (34), who investigated 18- to 35-yr-old, highly trained individuals.

Our findings are particularly relevant when interpreting study results that have repeated measurements. These data demonstrate that, in the context of criteria that is too low, it would be possible to derive a substantial but fallacious increase or decrease in measured $\dot{V}O_{2peak}$ of up to 32%, on average, without there being a real increase in an individual's cardiovascular and muscle capacity to utilize oxygen, thus leading to a higher $\dot{V}O_{2max}$. Exhaustion criteria are not only highly relevant in studies with repeated measurements—this also applies to the establishment or publication of reference values for CPET. In recent years, several normal reference values from different countries and for different exercise modes have been published for CPET (16,35–39). Many of them, however, have applied relatively low criteria (16,36–39), such as RER 1.0 or 85% APMHR. Using these reference values plausibly can overestimate aerobic fitness of the participant, client, or patient; hence, they are potentially misclassified as having normal aerobic fitness.

The new proposed values are higher than the criteria used in several previous publications (7). The studies using higher criteria are mostly performed in younger participants, mostly men, and in moderately to highly trained individuals. We are not aware of a systematic analysis addressing secondary exhaustion criteria for CPET specifically performed on a cycle ergometer in people older than 40 yr.

Multiparameter score. The proposed score provides, for the first time, a meaningful combination of several criteria and can therefore reduce type I errors as compared with using a single criterion with the same small type II error of at most 5%. Frequently used scores selected on the basis of intuition rather than an evidence-based approach result in an unknown number of type I and II errors. An example of such a score is "participants need to fulfill one out of three criteria" or "two out of four criteria" or setting low criteria but stipulating that participants need to "fulfill all of the criteria." To use the provided tolerance interval for the score, the standardized *z* values for RER_{max}, $APMHR_{210}$, and age must be calculated. These *z* variables then must be multiplied with their corresponding loading and the resulting values summed to get the final score. Finally, the score must be compared with our tolerance interval to evaluate exhaustion.

Beyond the advantage of being more applicable and time-efficient in comparison to a $\dot{V}O_2$ plateau determination, the criteria $AMPHR_{210}$, $APMHR_{208}$, and RER have the advantage that they can be continuously monitored during CPET and can facilitate further motivation to the participant during the test. These criteria can also be applied immediately after the test without time-consuming raw data analyses and can be of assistance in limiting an underestimation of an individual's $\dot{V}O_{2max}$.

Although type I and II errors will always occur using the proposed secondary exhaustion criteria, the approach described herein minimizes these errors. As discussed above, the criterion standard determination of $\dot{V}O_{2max}$, the $\dot{V}O_2$ plateau, is not present in half the participants. These secondary criteria are, therefore, needed even though they are not able to distinguish between $\dot{V}O_{2max}$ and $\dot{V}O_{2peak}$—per the definition, this can only be done by the detection of plateau—but rather they minimize the bias of an underestimation of the $\dot{V}O_{2peak}$.

The current results suggest that for subjects not exhibiting a plateau, applying the multiparameter approach or one of the proposed secondary criteria would be useful, particularly for research purposes. For nonresearch purposes, the proposed secondary exhaustion criteria $APMHR_{210}$, $APMHR_{208}$, or RER are recommended. Which of these three criteria is applied needs to be defined before the test. It is important to mention that only one of the three criteria needs to be applied. Among the three proposed criteria, the RER criterion reaches the highest % $\dot{V}O_{2peak}$ on average (Table 3) and has the lowest area under the curve (Fig. 1) and is, therefore, recommended. However, $APMHR_{210}$ or $AMPHR_{208}$ appear to work nearly as well.

Strengths and limitations. This work is the first to our knowledge to determine data based secondary exhaustion criteria for healthy adults across a broad age span. In addition, these data are the first to suggest multicriteria score based on data from a large sample in contrast to scores defined by expert opinion (40). A further strength is the large data set and the equal distribution of the participants across age decades and sex. In addition, all CPET were performed under rigorously standardized conditions using the same equipment.

A limitation of this study is that it is uncertain whether the proposed criteria are transferable when CPET are performed on a treadmill instead of a cycle ergometer. Studies have generally reported lower values for cycling compared with running (33,41). Because HR_{max} and RER_{max} depend on the incremental rate (42), the present findings are only valid for cycling protocols that lead to exhaustion in 6 to 18 min.

Even though data were available for participants older than 70 yr, the criteria for elderly subjects cannot be recommended. The variability of RER, $APMHR_{210}$, and $AMPHR_{208}$ was large enough in this subgroup that the calculated criteria could limit type II errors, but were quite low and, therefore, would likely lead to large type I errors. To determine maximal exhaustion in participants older than 70 yr, the only valid method remains the criterion standard determination of a $\dot{V}O_2$ plateau. A further limitation is the limited transferability of the criteria to clinical populations, such as patients with chronic heart failure or chronic obstructive pulmonary disease. Future research could apply a similar approach to evaluate secondary exhaustion criteria for different clinical populations and for CPET performed on treadmills.

CONCLUSIONS

In the general population, high and age-stratified secondary exhaustion criteria must be chosen to distinguish between a maximal and a submaximal effort. Based on our analyses, we recommend the following cutoffs for the age group 20 to 39 yr: $RER_{max} \geq 1.13$, $APMHR_{210} - \text{age} \geq 96\%$, and $APMHR_{208} \times 0.7 \text{ age} \geq 93\%$; for the age group of 40 to 59 yr: $RER_{max} \geq 1.10$, $APMHR_{210} - \text{age} \geq 99\%$, and $APMHR_{208} \times 0.7 \text{ age} \geq 92\%$; and, for the age group of 60 to 69 yr: $RER_{max} \geq 1.06$, $APMHR_{210} - \text{age} \geq 99\%$, and $APMHR_{208} \times 0.7 \text{ age} \geq 89\%$. Lower cutoff values are likely to produce type I errors. Our study and the above recommendations have the potential to improve standardized application and quality of $\dot{V}O_{2max}$ reporting in research and clinical practice.

We thank all participants who made this study possible by their participation.

R. K., J. W. participated in the conceptualization. J. W., R. K., A.S.-T. participated in the methodology. J. W., M. N., D. I., R. K. participated in the formal analysis. J. W., R. K. participated in the investigation. J. W., R. K., M. N. participated in writing the original draft. D. I., H. H., L. S., A. S.-T. participated in the writing—review and editing. J. W., D. I. participated in the visualization. H. H., A. S.-T. participated in the supervision.

Conflicts of interest and funding: The analysis of this data was funded by the Department of Sport, Exercise and Health of the University of Basel, Switzerland. The authors declare that they have no conflicts of interest.

The results of the present study do not constitute endorsement by ACSM. The results of the present study are presented clearly, honestly, and without fabrication, falsification, or inappropriate data manipulation.

REFERENCES

1. Laukkanen JA, Zaccardi F, Khan H, Kurl S, Jae SY, Rauramaa R. Long-term change in cardiorespiratory fitness and all-cause mortality: a population-based follow-up study. *Mayo Clin Proc*. 2016;91(9):1183–8.
2. Imboden MT, Harber MP, Whaley MH, Finch WH, Bishop DL, Kaminsky LA. Cardiorespiratory fitness and mortality in healthy men and women. *J Am Coll Cardiol*. 2018;72(19):2283–92.

3. Ross R, Blair SN, Arena R, et al. Importance of assessing cardiorespiratory fitness in clinical practice: a case for fitness as a clinical vital sign: a scientific statement from the American Heart Association. *Circulation*. 2016;134(24):e653–e99.
4. Howley ET, Bassett DR Jr, Welch HG. Criteria for maximal oxygen uptake: review and commentary. *Med Sci Sports Exerc*. 1995;27(9):1292–301.
5. Mehra MR, Canter CE, Hannan MM, et al. The 2016 International Society for Heart Lung Transplantation listing criteria for heart transplantation:a 10-year update. *J Heart Lung Transplant*. 2016;35(1):1–23.
6. Arena R, Myers J, Williams MA, et al. Assessment of functional capacity in clinical and research settings: a scientific statement from the American Heart Association Committee on Exercise, Rehabilitation, and Prevention of the Council on Clinical Cardiology and the Council on Cardiovascular Nursing. *Circulation*. 2007;116(3):329–43.
7. Midgley AW, McNaughton LR, Polman R, Marchant D. Criteria for determination of maximal oxygen uptake: a brief critique and recommendations for future research. *Sports Med*. 2007;37(12):1019–28.
8. Edvardsen E, Hem E, Anderssen SA. End criteria for reaching maximal oxygen uptake must be strict and adjusted to sex and age: a cross-sectional study. *PLoS One*. 2014;9(1):e85276.
9. Wood RE, Hills AP, Hunter GR, King NA, Byrne NM. Vo2max in overweight and obese adults: do they meet the threshold criteria? *Med Sci Sports Exerc*. 2010;42(3):470–7.
10. Lucia A, Rabadan M, Hoyos J, et al. Frequency of the VO2max plateau phenomenon in world-class cyclists. *Int J Sports Med*. 2006;27(12):984–92.
11. Barker AR, Williams CA, Jones AM, Armstrong N. Establishing maximal oxygen uptake in young people during a ramp cycle test to exhaustion. *Br J Sports Med*. 2011;45(6):498–503.
12. Murias JM, Pogliaghi S, Paterson DH. Measurement of a true [Formula: see text]O2max during a ramp incremental test is not confirmed by a verification phase. *Front Physiol*. 2018;9:143.
13. Possamai LT, Campos FS, Salvador P, et al. Similar VO2max assessment from a step cycling incremental test and verification tests on the same or different day. *Appl Physiol Nutr Metab*. 2019;1–5.
14. Poole DC, Wilkerson DP, Jones AM. Validity of criteria for establishing maximal O2 uptake during ramp exercise tests. *Eur J Appl Physiol*. 2008;102(4):403–10.
15. Poole DC, Jones AM. Measurement of the maximum oxygen uptake Vo2max: Vo2peak is no longer acceptable. *J Appl Physiol (1985)*. 2017;122(4):997–1002.
16. Kokkinos P, Kaminsky LA, Arena R, Zhang J, Myers J. A new generalized cycle ergometry equation for predicting maximal oxygen uptake: the Fitness Registry and the Importance of Exercise National Database (FRIEND). *Eur J Prev Cardiol*. 2018;25(10):1077–82.
17. Fleg JL, Morrell CH, Bos AG, et al. Accelerated longitudinal decline of aerobic capacity in healthy older adults. *Circulation*. 2005;112(5):674–82.
18. Wagner J, Knaier R, Infanger D, et al. Functional aging in health and heart failure: the COmPLETE Study. *BMC Cardiovasc Disord*. 2019;19(1):180.
19. Agostoni P, Bianchi M, Moraschi A, et al. Work-rate affects cardiopulmonary exercise test results in heart failure. *Eur J Heart Fail*. 2005;7(4):498–504.
20. Midgley AW, Bentley DJ, Luttikholt H, McNaughton LR, Millet GP. Challenging a dogma of exercise physiology: does an incremental exercise test for valid VO 2 max determination really need to last between 8 and 12 minutes? *Sports Med*. 2008;38(6):441–7.
21. American Thoracic S. American College of Chest P. ATS/ACCP Statement on cardiopulmonary exercise testing. *Am J Respir Crit Care Med*. 2003;167(2):211–77.
22. Jones AM, Carter H. Oxygen uptake-work rate relationship during two consecutive ramp exercise tests. *Int J Sports Med*. 2004;25(6):415–20.
23. Niemeyer M, Bergmann TGJ, Beneke R. Oxygen uptake plateau: calculation artifact or physiological reality? *Eur J Appl Physiol*. 2019;120:231–42.
24. Midgley AW, Carroll S, Marchant D, McNaughton LR, Siegler J. Evaluation of true maximal oxygen uptake based on a novel set of standardized criteria. *Appl Physiol Nutr Metab*. 2009;34(2):115–23.
25. Gravelle BM, Murias JM, Spencer MD, Paterson DH, Kowalchuk JM. Adjustments of pulmonary O2 uptake and muscle deoxygenation during ramp incremental exercise and constant-load moderate-intensity exercise in young and older adults. *J Appl Physiol (1985)*. 2012;113(9):1466–75.
26. Boone J, Koppo K, Bouckaert J. The VO2 response to submaximal ramp cycle exercise: Influence of ramp slope and training status. *Respir Physiol Neurobiol*. 2008;161(3):291–7.
27. Roecker K, Striegel H, Dickhuth HH. Heart-rate recommendations: transfer between running and cycling exercise? *Int J Sports Med*. 2003;24(3):173–8.
28. Tanaka H, Monahan KD, Seals DR. Age-predicted maximal heart rate revisited. *J Am Coll Cardiol*. 2001;37(1):153–6.
29. Wongpakaran N, Wongpakaran T, Wedding D, Gwet KL. A comparison of Cohen's Kappa and Gwet's AC1 when calculating inter-rater reliability coefficients: a study conducted with personality disorder samples. *BMC Med Res Methodol*. 2013;13.
30. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33(1):159–74.
31. Myers J, Walsh D, Sullivan M, Froelicher V. Effect of sampling on variability and plateau in oxygen uptake. *J Appl Physiol (1985)*. 1990;68(1):404–10.
32. Beltrami FG, Wong Del P, Noakes TD. High prevalence of false-positive plateau phenomena during VO2max testing in adolescents. *J Sci Med Sport*. 2014;17(5):526–30.
33. Gordon D, Mehter M, Gernigon M, Caddy O, Keiller D, Barnes R. The effects of exercise modality on the incidence of plateau at VO2max. *Clin Physiol Funct Imaging*. 2012;32(5):394–9.
34. Knaier R, Niemeyer M, Wagner J, et al. Which cutoffs for secondary V O2max criteria are robust to diurnal variations? *Med Sci Sports Exerc*. 2019;51(5):1006–13.
35. Kaminsky LA, Imboden MT, Arena R, Myers J. Reference standards for cardiorespiratory fitness measured with cardiopulmonary exercise testing using cycle ergometry: data from the Fitness Registry and the Importance of Exercise National Database (FRIEND) Registry. *Mayo Clin Proc*. 2017;92(2):228–33.
36. Mylius CF, Krijnen WP, van der Schans CP, Takken T. Peak oxygen uptake reference values for cycle ergometry for the healthy Dutch population: data from the LowLands Fitness Registry. *ERJ Open Res*. 2019;5(2):00056–2018.
37. Koch B, Schaper C, Ittermann T, et al. Reference values for cardiopulmonary exercise testing in healthy volunteers: the SHIP study. *Eur Respir J*. 2009;33(2):389–97.
38. Hakola L, Komulainen P, Hassinen M, et al. Cardiorespiratory fitness in aging men and women: the DR's EXTRA study. *Scand J Med Sci Sports*. 2011;21(5):679–87.
39. Genberg M, Andren B, Lind L, Hedenstrom H, Malinovschi A. Commonly used reference values underestimate oxygen uptake in healthy, 50-year-old Swedish women. *Clin Physiol Funct Imaging*. 2018;38(1):25–33.
40. Robergs RA, Dwyer D, Astorino T. Recommendations for improved data processing from expired gas analysis indirect calorimetry. *Sports Med*. 2010;40(2):95–111.
41. Davis JA, Vodak P, Wilmore JH, Vodak J, Kurtz P. Anaerobic threshold and maximal aerobic power for three modes of exercise. *J Appl Physiol*. 1976;41(4):544–50.
42. Adami A, Sivieri A, Moia C, Perini R, Ferretti G. Effects of step duration in incremental ramp protocols on peak power and maximal oxygen consumption. *Eur J Appl Physiol*. 2013;113(10):2647–53.