Expert Reviews

Reference values for cardiopulmonary exercise testing in healthy adults: a systematic review

Expert Rev. Cardiovasc. Ther. 12(12), 1439–1453 (2014)

Davy Paap[1] **and Tim Takken***[1–3]

[1]University Utrecht and University Medical Center Utrecht, Master Clinical Health Sciences, Program in Physical Therapy Science, Utrecht, The Netherlands
[2]Child Development and Exercise Centre, Wilhelmina Children's Hospital, University Medical Center Utrecht, Utrecht, The Netherlands
[3]The Physiology Academy, Alphen aan den Rijn, The Netherlands
**Author for correspondence:*
Tel.: +31 887 554 030
t.takken@umcutrecht.nl

Reference values (RV) for cardiopulmonary exercise testing (CPET) provide the comparative basis for answering important questions concerning the normality of exercise response in patients and significantly impacts the clinical decision-making process. The aim of this study is to systematically review the literature on RV for CPET in healthy adults. A secondary aim is to make appropriate recommendations for the practical use of RV for CPET. Systematic searches of MEDLINE, EMBASE and PEDro databases up to March 2014 were performed. In the last 30 years, 35 studies with CPET RV were published. There is no single set of ideal RV; characteristics of each population are too diverse to pool the data in a single equation. Therefore, each exercise laboratory must select appropriate sets of RV that best reflect the characteristics of the population/patient tested, and equipment and methodology utilized.

Keywords: cardiopulmonary exercise testing • exercise physiology • healthy adults • maximal oxygen consumption • reference values

Physical activity (PA) is a key component of healthy lifestyle and disease prevention. Health professionals should stimulate PA and prescribe exercise for health-related fitness in their patients [1]. Before an exercise regimen is prescribed, it is important to assess the exercise capacity of a patient. Many different exercise tests are available for this purpose [2]; the gold standard for exercise testing remains cardiopulmonary exercise testing (CPET) using respiratory gas exchange analysis during incremental exercise [3].

This form of testing provides a global assessment of the integrative exercise response involving the pulmonary, cardiovascular, hematopoietic, neuropsychological and skeletal muscle systems, which are not adequately reflected through the measurement of individual organ system function [4]. This relative noninvasive, dynamic physiologic overview permits the evaluation of both submaximal and peak exercise response, providing the physician relevant information for clinical decision making [4]. Examples of usefulness of CPET for important clinical decisions are evaluating by lung, heart-lung transplantation patients, preoperative evaluation, exercise intolerance, or patient with cardiovascular diseases [5].

Normal reference values (RV) provide the comparative basis for answering important questions concerning the normality of exercise responses in patients, and can significantly impact the clinical decision-making process [5]. As recommended by the American Thoracic Society/American College of Chest Physicians (ATS/ACCP) guideline, each exercise laboratory must select an appropriate set of RV that best reflects the characteristics of the population tested, and the equipment and methodology utilized [4]. Several papers reporting RV for CPET are available, obtained in different populations.

From a historical perspective, a review of available literature for maximal oxygen uptake (VO_{2max}) on population samples was done by Shephardin in 1966 [6]. Shvartz *et al.* published in 1990 a meta-analysis of the most important CPET parameters VO_{2max}, minute ventilation (V_E), and maximal heart rate (HR_{peak}) [7]. These

10.1586/14779072.2014.985657 ISSN 1477-9072

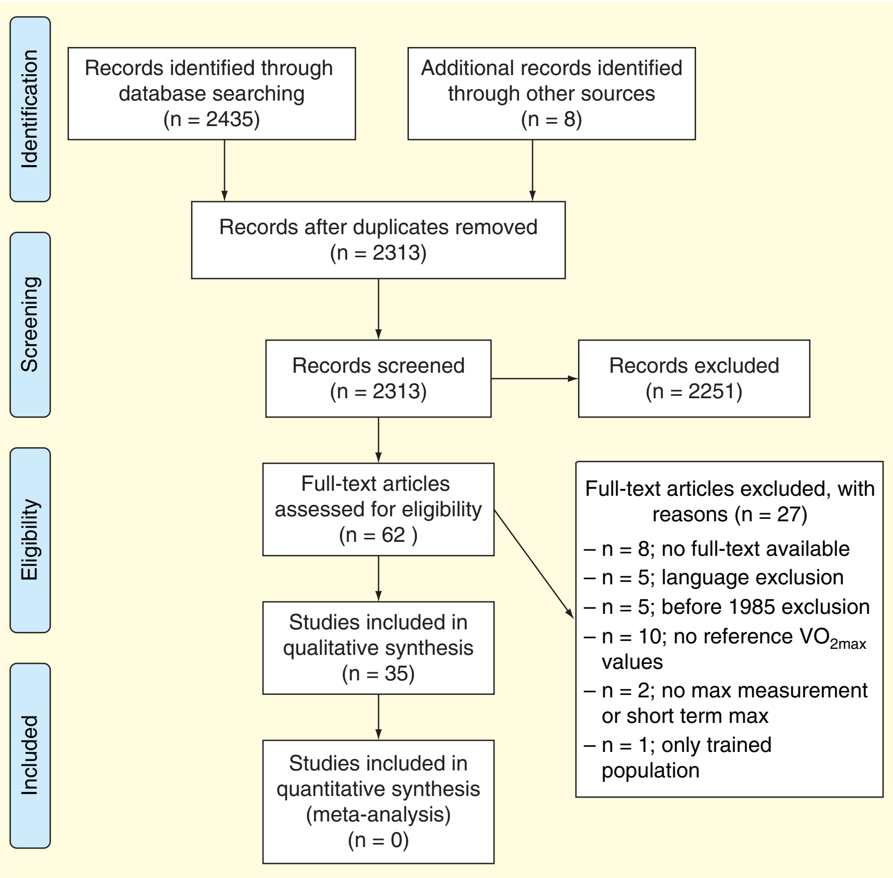

Figure 1. Flow scheme of included studies (Prisma 2009 flow diagram).

two reviews were not systematic and were based on older literature with predominantly small sample sizes.

However, despite increasing use of RV in exercise testing and increasing importance, there is no systematic review of the RV for CPET. A systematic review might aid a clinician in choosing the set of reference value that best reflect the characteristics of the patient tested.

Therefore, the aim of this study is to systematically review the literature on RV for CPET in healthy adults. Furthermore, we aim to give appropriate recommendations for the practical use of RV for CPET.

Methods

This systematic review of existing literature followed the guidelines of the PRISMA statement [8].

Data sources & searches

A search strategy was created by the first author (DP) with the support of a medical librarian and critically reviewed and approved by an experienced exercise physiologist (TT). After approval, published articles in the following databases were searched: MEDLINE, from 1950; PEDro from 1980; and EMBASE, from 1980 to March 2014. We used the systematic search strategy as described in Appendix A. The search strategy did not have any limitations on ethnicity. Relevant reference lists were hand-searched.

Selection of studies

Combining the results of electronic searches, duplicates were removed by the first author (DP). All unique records were screened by two individual reviewers (DP and TT) for potential relevance using the title, abstract or descriptors, or both. The first author (DP) assessed the remaining articles for compliance with the eligibility criteria, based on full text. Reasons for possible exclusion based on full text were noted.

Eligibility criteria

Studies with the objective to evaluate RV of maximal CPET were included. Furthermore, inclusion criteria were: healthy adults from age 18–80 years, studies with CPET using cycle ergometer or treadmill, cross-sectional studies or cohort studies and studies that reported the outcomes of VO_{2max} or maximal workload (WR_{peak}). We excluded studies with participants with a mean age younger than 18 (or 70 and older) years, studies before 1985, studies that were not published in English, studies without the availability of full text and intervention studies.

Data extraction

The first reviewer (DP) extracted data using a standard extraction form. Data extracted from the included articles are shown in Table 1. If data were missing or further information was required, serious attempts were made to contact the first two authors to request information.

Methodological quality

The methodological quality of the selected studies was assessed using a quality list based on the ATS/ACCP guidelines (Supplementary Appendix B [supplementary material can be found online at www.informahealthcare.com/suppl/10.1586/14779072.2014.985657]) [4]. This list is a combination of the study requirements for an optimal set of normal RV as described in the ATS/ACCP guidelines and the code number scheme of shortcomings and limitations. Each criterion was scored as ‘yes’, ‘no’ or ‘don’t know’ with one point for each ‘yes’. A study was considered to be high quality when it scored ≥10 points (≥75% of maximum score of 14), moderated quality when it scored 7–9 points and low quality when it scored ≤ 6 points.

Two persons performed independently a quality assessment of each study. Disagreements about the eligibility of a study or differences between the two sets of information extracted were resolved by consensus or by referring to a third person when disagreement persisted. There was no blinding on authors or journal.

Table 1. Study characteristics.

Study (year)	Sample size (M/F)	Age (year)	Sample characteristics	Country	Smokers included	Treadmill or cycle	Protocol	Primary parameters measured	Methodology	Time averaging (s)	Ref.
Neder *et al.* (2001)	60 M/60 F	20–80	University population, prospective, randomized	Brazil	Yes	CY	Ramp, 10–30 W/min	VO_2, WR, HR, Ve, VCO_2, VAT, $PETO_2$, $PETCO_2$	BxB, gas analyzers	15	[9]
Arstila *et al.* (1990)	423 M/289 F	30–67	Retrospective? population based	Finland	Yes	CY	Workload increase linear by heart rate 5 beat/min every min	WR, HR	Electronic cycle ergometer	30	[10]
Aspenes *et al.* (2011)	2368 M/2263 F	20–90	Prospective, population based	Norway	Yes (take into account in analyze)	TM	Individualized protocol modif. elef, 1–2% each step; velocity 0.5–1 km/h^{-1}	HR, VO_2, Ve	Mixing chamber, gas analyzers ergospirometry	?	[11]
Blackie *et al.* (1989)	47 M/81 F	55–80	Prospective, hospital based, senior centers	UK	Yes	CY	Incremental 16 W/min	VO_2,VCO_2, WR, HR	Turbine, mixing chamber	30	[39]
Blackie *et al.* (1991)	120 M/111 F	20–80	Prospective, community based and lectures	Canada	Yes	CY	Incremental 16 W/min	VO_2, VCO_2, VT, Ve	Mixing chamber, infrared gas analyzers	60	[12]
Bromley *et al.* (2006)	50 M/52 F	19–64	Prospective, population based	UK	Yes	TM	Bruce	VAT, VE, VO_2, VCO_2, WR, $PETO_2$, $PETCO_2$	BXB, gas analyzers	60	[13]
Brudin *et al.* (2013)	1043 M/747 F	7–80	Hospital based, retrospective	Sweden	Yes	CY	Incremental 10–15 W/min	HR, WR	Electronic cycle ergometer	?	[14]
Davis *et al.* (1997)	103 M/101 F	20–70	Prospective, community based	USA	No	CY	Incremental 15 W/min	VO_2, VCO_2, Ve	BxB, gas analyzers	30	[15]
Davis *et al.* (2002)	115 M/115 F	20–70	?	USA	No	CY	Incremental 15 W/min	VO_2, VCO_2, HR, Ve	Mixing chamber, BxB, gas analyzers	30	[16]
Dubowy *et al.* (2008)	647 M/548	4–75	Prospective, population based	Germany	Yes	TM	Modif stepwise. elef 3% each step; velocity 0.5 km/h^{-1}	VO_2, VCO_2, Ve, VAT, $PETO_2$, $PETCO_2$	BxB, gas analyzers	30	[17]
Edvardsen, *et al.* (2013)	349 M/365 F	20–85	Caucasian, prospective, population based, randomized	Norway	Yes	TM	Balke modif. elef 2%/min	VO_2, VCO_2, Ve, HR, venous lactate	Gas analyzers	30	[18]
Fairbarn *et al.* (1994)	111 M/120 F	20–80	Prospective, general population	Canada	Yes	CY	Incremental 16 W/min	VO_2, HR	Turbine, mixing chamber	30	[41]

?: Not stated; ABGs: Arterial blood gases; BxB: Breath by breath; CY: Cycle; F: Female; HR: Heart rate; M: Male; Modif: Modification; P(a-A-) O_2: Alveolar-arterial difference for oxygen pressure; $PETCO_2$: End-tidal PCO_2; $PETO_2$: End-tidal PO_2; TM: Treadmill; VAT: Ventilatory anaerobic threshold; Vd/VT: Ratio physiologic dead space to tidal volume; Ve: Minute ventilation; VCO_2: Carbon dioxide output; VO_2: Oxygen uptake; WR: Work rate; Yr: Year.

Table 1. Study characteristics (cont.).

Study (year)	Sample size (M/F)	Age (year)	Sample characteristics	Country	Smokers included	Treadmill or cycle	Protocol	Primary parameters measured	Methodology	Time averaging (s)	Ref.
Farazdaghi *et al.* (2001)	87 F	20–80	Prospective, population based, randomized	Sweden	Yes	CY	Incremental 5 W/30 s	HR, WR	Electronic cycle ergometer	?	[19]
Habedank *et al.* (1998)	56 M/45 F	16–75	General population, prospective	Germany	Yes	TM	Naughton modif, incremental 2-min stage	VO_2, VCO_2,Ve, $PETO_2$, $PETCO_2$	BxB, gas analyzers	?	[20]
Hakola *et al.* (2011)	672 M/677 F	57–78	Population based, retrospective, randomized	Finland	Yes	CY	Incremental 20 W/min	VO_2, WR	BxB, gas analyzers	20	[21]
Herdy *et al.* (2011)	2388 M/1534 F	15–74	Hospital based, sedentary and active population, retrospective	Brazil	No	TM	Ramp, ?	VO_2, VCO_2, Ve, HR	Mixing chamber, gas analyzers	?	[22]
Hollenberg *et al.* (1998)	846 M/1246 F	55–95	Population-based study, retrospective	USA	Yes	TM	Bruce treadmill exercise protocol	VO_2, VCO_2, Ve, HR	BxB, gas analyzers	?	[23]
Inbar *et al.* (1994)	1424 M	20–70	General sedentary population? retrospective	Israel	Yes	TM	Balke modif, elef 2%/min	HR, VO_2, VCO_2, Ve, VT, $PETO_2$, $PETCO_2$, VAT	BxB, gas analyzers	30	[24]
Itoh *et al.* (2013)	387 M/362 F	20–78	Prospective, population based, randomized	Japan	No	CY/ TM	Cy; ramp protocols 10 W/min/20 W/min/30 W/Min Tr; ramp protocols, incremental in speed and grade	VO_2, VCO_2, Ve, HR, WR, $PETO_2$, $PETCO_2$	BxB, gas analyzers	30	[25]
John *et al.* (2011)	82 M/ 19 F	18–69	Staff and caregivers of patients, prospective	India	No	TM	Balke incremental exercise protocol 1%/min	HR, VO_2, VCO_2, Ve	?	?	[26]
Jones *et al.* (1985)	50 M/50 F	15–71	Prospective, population (local and university) based	Canada	Yes	CY	Incremental 16 W/min	VO_2, VCO_2, VAT, Ve, HR	Mixing chamber, Gas analyzers	15	[27]
Jones *et al.* (1989)	732 M/339 F	20–70	Hospital referrals; retrospective	Canada	?	CY	Incremental 16 W/min	WR	Electronic cycle ergometer	30	[28]
Koch *et al.* (2009)	253 M/281 F	25–80	Prospective, population based	Germany	No	CY	Jones modif. Incremental 16 W/min	VO_2, VCO_2, Ve, $PETO_2$, $PETCO_2$, Vd/VT	BxB, gas analyzers	10	[43]

?: Not stated; ABGs: Arterial blood gases; BxB: Breath by breath; CY: Cycle; F: Female; HR: Heart rate; M: Male; Modif: Modification; P(a-A-) O_2: Alveolar-arterial difference for oxygen pressure; $PETCO_2$: End-tidal PCO_2; $PETO_2$: End-tidal PO_2; TM: Treadmill; VAT: Ventilatory anaerobic threshold; Vd/VT: Ratio physiologic dead space to tidal volume; Ve: Minute ventilation; VCO_2: Carbon dioxide output; VO_2: Oxygen uptake; WR: Work rate; Yr: Year.

Table 1. Study characteristics (cont.).

Study (year)	Sample size (M/F)	Age (year)	Sample characteristics	Country	Smokers included	Treadmill or cycle	Protocol	Primary parameters measured	Methodology	Time averaging (s)	Ref.
Magrani *et al.* (2010)	92 M/45 F	?	Prospective, randomized	Brazil	No	CY	Continuous graded maximal exercise protocol	VO_2, VCO_2, HR	Pneumotach, metabolic	20	[29]
Meyer *et al.* (1994)	69 M	20–59	Clinical staff	Germany	Yes	CY	Incremental 12.5 W/min	VO_2, VCO_2, Ve, VAT, $PETO_2$, $PETCO_2$, Vd/VT	BxB, gas analyzers	?	[30]
Nelson *et al.* (2010)	816 M	30–69	Retrospective? cross sectional and longitudinal, laboratory population, sedentary	Canada	?	TM	Balk modif. Elef. 2%/min ?	VO_2, VCO_2, Ve	Mixing chamber, gas analyzers	20	[38]
Nordenfelt 1985	93 M/95 F	20–79	Prospective, population based, randomized	Sweden	?	CY	Incremental 10 W/min	WR, HR	Electronic cycle ergometer	?	[31]
Ong *et al.* (2002)	48 M/47 F	20–70	Prospective, general population	China	Yes	CY	Incremental 10–30 W/min	VO_2, VCO_2, Ve, VAT	Turbine, gas analyzers	30	[42]
Singh *et al.* (1989)	167 M	13–59	?	Malaysia	?	CY	Incremental 16 W/min	VO_2, VCO_2, Ve, HR	Mixing chamber, gas analyzers	?	[32]
Storer *et al.* (1990)	115 M/116 M	20–70	Prospective, general population, sedentary	USA	Yes	CY	Incremental 15 W/min	VO_2, VCO_2, WR, Ve	Mixing chamber, BxB, turbine	30	[40]
Sun *et al.* (2012)	281 M/136 F	17–78	?	USA/ Spain	?	TM	Ramp protocol, incremental in speed and grade ?	VO_2, VCO_2, Ve, VAT	?	30	[33]
Tammelin *et al.* (2004)	63 M/60 F	31	Prospective, cohort population-based sample	Finland	Yes	CY	Incremental 20 – 25 W/2 min	VO_2	Gas analyzers	?	[34]
Vogel *et al.* (1986)	1514 M/375 F	17–55	Prospective, random US soldiers	USA	?	TM	Discontinuous, 3-min stages	VO_2	Douglas bag, gas analyzers	60	[37]
Wisen *et al.* (2004)	25 F	22–44	Prospective, population-based sample, randomized	Sweden	Yes	CY	Ramp protocol, incremental 5 W/30 s	VO_2, VCO_2, Ve,	BxB, gas analyzers	10	[35]
Wohlfart *et al.* (2003)	81 M	20–80	Prospective, population based, randomized	Sweden	Yes	CY	Incremental 5 W/30 s	HR, WR			[36]

?: Not stated; ABGs: Arterial blood gases; BxB: Breath by breath; CY: Cycle; F: Female; HR: Heart rate; M: Male; Modif: Modification; P(a-A-) O_2: Alveolar-arterial difference for oxygen pressure; $PETCO_2$: End-tidal PCO_2; $PETO_2$: End-tidal PO_2; TM: Treadmill; VAT: Ventilatory anaerobic threshold; Vd/VT: Ratio physiologic dead space to tidal volume; Ve: Minute ventilation; VCO_2: Carbon dioxide output; VO_2: Oxygen uptake; WR: Work rate; Yr: Year.

Table 2. Methodological quality.

Study (year)	1	2	3	4	5	6	7	8	9	10	11	12	13	14	Total score	Ref.
Neder *et al.* (2001)	1	1	0	0	1	0	1	1	0	1	1	0	1	1	9	[9]
Arstila *et al.* (1990)	1	1	0	1	1	0	0	0	0	1	0	1	0	1	7	[10]
Aspenes *et al.* (2011)	1	1	0	1	1	1	0	1	0	1	1	0	0	0	8	[11]
Blackie *et al.* (1989)	0	0	1	0	1	0	0	1	0	1	1	1	1	1	8	[39]
Blackie *et al.* (1991)	1	0	0	0	1	0	0	1	0	1	1	1	0	1	7	[12]
Bromley *et al.* (2006)	1	0	0	0	1	0	0	1	0	1	1	1	0	0	6	[13]
Brudin *et al.* (2013)	0	0	0	0	1	0	0	0	0	1	0	0	0	1	3	[14]
Davis *et al.* (1997)	1	1	0	1	1	0	0	1	0	1	1	1	1	1	10	[15]
Davis *et al.* (2002)	0	1	0	1	1	0	0	0	0	1	1	1	0	1	7	[16]
Dubowy *et al.* (2008)	1	0	0	0	1	0	0	1	0	1	1	1	0	0	6	[17]
Edvardsen *et al.* (2013)	1	1	1	0	1	0	1	1	0	1	1	1	0	1	10	[18]
Fairbarn *et al.* (1994)	1	0	1	0	0	0	0	1	0	1	1	1	1	1	8	[41]
Farazdaghi *et al.* (2001)	1	0	0	0	1	0	1	1	0	1	0	0	0	1	6	[19]
Habedank *et al.* (1998)	1	0	0	0	0	0	0	1	0	1	1	0	0	0	4	[20]
Hakola *et al.* (2011)	1	1	0	0	1	0	1	0	0	1	1	1	0	1	8	[21]
Herdy *et al.* (2011)	0	1	0	1	1	0	0	0	0	1	1	0	0	0	5	[22]
Hollenberg *et al.* (1998)	1	1	0	0	1	0	0	0	0	1	1	0	0	1	6	[23]
Inbar *et al.* (1994)	1	1	1	0	1	0	0	0	0	1	1	1	0	1	8	[24]
Itoh *et al.* (2013)	1	1	1	1	1	0	1	1	0	1	1	1	0	1	11	[25]
John *et al.* (2011)	0	1	1	1	0	0	0	1	0	1	0	0	0	1	6	[26]
Jones *et al.* (1985)	1	0	0	0	1	0	0	1	0	1	1	1	0	1	7	[27]
Jones *et al.* (1989)	0	0	0	0	1	0	0	0	0	1	1	1	1	1	6	[28]
Koch *et al.* (2009)	1	0	0	1	1	0	0	1	0	1	1	1	0	1	8	[43]
Magrani *et al.* (2010)	1	0	0	1	1	0	1	1	0	1	0	0	1	1	8	[29]
Meyer *et al.* (1994)	0	1	0	0	0	0	0	0	0	1	1	0	0	0	3	[30]
Nelson *et al.* (2010)	0	1	0	0	1	0	0	0	0	0	1	0	1	1	5	[38]
Nordenfelt *et al.* (1985)	1	0	0	0	1	0	1	1	0	1	0	0	0	0	5	[31]
Ong *et al.* (2002)	1	1	0	1	1	0	0	1	0	1	0	1	1	0	8	[42]
Singh *et al.* (1989)	1	0	0	0	1	0	0	0	0	1	1	0	0	0	4	[32]
Storer *et al.* (1990)	1	1	0	0	1	1	0	1	0	1	1	1	1	1	10	[40]
Sun *et al.* (2012)	1	0	0	0	0	0	0	0	0	1	0	1	1	1	5	[33]
Tammelin *et al.* (2004)	1	1	0	0	0	0	0	1	0	1	0	0	1	1	6	[34]
Vogel *et al.* (1986)	1	1	0	0	1	0	1	1	0	1	1	1	0	1	9	[37]
Wisen *et al.* (2004)	1	1	0	0	1	0	1	1	0	1	1	0	1	1	9	[35]
Wohlfart *et al.* (2003)	1	1	0	0	1	0	1	1	0	1	0	0	0	1	7	[36]

Table 3. Reference values for cardiopulmonary exercise testing.

Study (year)	Variables	Equations[†]	R^2[#]	SEE[#]	Ref.
$VO_{2\ max}$ ml·min^{-1}					
Blackie *et al.* (1989)	$VO_{2\ max}$ l·min^{-1}, male	3.015 + 0.0142 (height) – 0.0494 (age) + 0.00257 (weight)	0.40	0.34	[39]
Blackie *et al.* (1989)	$VO_{2\ max}$ l·min^{-1}, female	0.651 + 0.0142 (height) – 0.0115 (age) + 0.00974 (weight)	0.27	0.31	[39]
Koch *et al.* (2009)[§]	$VO_{2\ max}$ ml·min^{-1}, both	6107.67 – 502.67 (age) – 8.3333 (age^2) – 1124.33 (sex) + 177.6667 (age·sex)	?	?	[43]
Inbar *et al.* (1994)	$VO_{2\ max}$ ml·min^{-1}, male	3499.5 – 25.55 (age)	0.46	477.7	[24]
Edvardsen *et al.* (2013)	$VO_{2\ max}$ l·min^{-1}, male	4.97 – 0.033 (age)	0.65	?	[18]
Edvardsen *et al.* (2013)	$VO_{2\ max}$ l·min^{-1}, female	3.31 – 0.022 (age)	0.64	?	[18]
Fairbarn *et al.* (1994)	$VO_{2\ max}$ l·min^{-1}, male	-0.332 – 0.031 (age) + 0.023 (height) + 0.0117 (weight)	0.70	?	[41]
Fairbarn *et al.* (1994)	$VO_{2\ max}$ l·min^{-1}, female	0.207 – 0.027 (age) + 0.0158 (height) + 0.00899 (weight)	0.43	?	[41]
Jones *et al.* (1985)	$VO_{2\ max}$ l·min^{-1}, male	-3.76 + 0.034 (height) – 0.028 (age) + 0.022 (weight)	0.80	0.483	[27]
Jones *et al.* (1985)	$VO_{2\ max}$ l·min^{-1}, female	-2.26 + 0.025 (height) – 0.018 (age) + 0.010 (weight)	0.66	0.388	[27]
Nelson *et al.* (2010)	$VO_{2\ max}$ l·min^{-1}, male	4.83 – 0.032·(age)	0.16	?	[38]
Magrani *et al.* (2010)	$VO_{2\ max}$ l·min^{-1}, male	0.518 + (0.01016 WR_{max}) + (0.01482. BMI) – (0.0292 age)	0.71	0.40	[29]
Ong *et al.* (2002)	$VO_{2\ max}$ ml·min^{-1}, both	7.6929 – 0.0060 (age) – 0.3522 (sex) + 0.0009 (height) + 0.0052 (weight)	0.66	?	[42]
Singh *et al.* (1989)	$VO_{2\ max}$ l·min^{-1}, male	1.99+ 0.035 (weight) – 0.04 (age)	0.55	?	[32]
Storer *et al.* (1990)	$VO_{2\ max}$ ml·min^{-1}, male	519,3 + 10.51 (WR_{max}) + 6.35 (weight) - 10.49 (age)	0.94	212	[40]
Storer *et al.* (1990)	$VO_{2\ max}$ ml·min^{-1}, female	136,7 + 9.39 (WR_{max}) + 7.7 (weight) - 5.88 (age)	0.93	147	[40]
$VO_{2\ max}$ ml·min^{-1}·kg^{-1}					
Edvardsen *et al.* (2013)	$VO_{2\ max}$ ml·min^{-1}·kg^{-1}, male	60.9 – 0.43 (age)	0.61	?	[18]
Edvardsen *et al.* (2013)	$VO_{2\ max}$ ml·min^{-1}·kg^{-1}, female	48.2 – 0.32 (age)	0.61	?	[18]
Habedank *et al.* (1998)	$VO_{2\ max}$ ml·min^{-1}·kg^{-1}, male	51.5 – 0.36 (age)	0.59	?	[20]
Habedank *et al.* (1998)	$VO_{2\ max}$ ml·min^{-1}·kg^{-1}, female	44.6 – 0.34 (age)	0.67	?	[20]

[†] Sex, male, 0; female, 1; age, years; height (ht), centimeters; Weight (kg); HR: Heart rate after step test, beat.min^{-1}; PA: Frequency of brisk physical activity, times per week; BMI: Body mass index; R^2: Coefficient of determination; SEE: Standard error of the estimate

[‡] Farazdaghi: Height (m); [§] Koch: body mass index (BMI) was code as 0 for BMI ≤25 kg.m^{-2} and 1 for BMI >25 kgm^{-2}. There were five age groups: 25–34, 35–44, 45–54,55–64 and ≥64 yrs of age. The content of each column was multiplied with the code characteristics.

Table 3. Reference values for cardiopulmonary exercise testing (cont.).

Study (year)	Variables	Equations[†]	R^2[#]	SEE[#]	Ref.
$VO_{2\ max}$ ml·min^{-1}·kg^{-1} (cont.)					
Hollenberg *et al.* (1998)	$VO_{2\ max}$ ml·min^{-1}·kg^{-1}, male	54.19 – 34 (age) + 14.54 (height) – 0.17 (height · age)	0.31	?	[23]
Hollenberg *et al.* (1998)	$VO_{2\ max}$ ml·min^{-1}·kg^{-1}, female	58.68 – 43 (age) + 0.97 (height)	0.29	?	[23]
Itoh *et al.* (2013) (cycle data)	$VO_{2\ max}$ ml·min^{-1}·kg^{-1}, both	42.05 – 0.268 (age) – 7.22 (sex) – 0.0811 (sex age)	?	?	[25]
Itoh *et al.* (2013) (treadmill data)	$VO_{2\ max}$ ml·min^{-1}·kg^{-1}, both	61.07 – 0.510 (age) – 20.4 (sex) – 0.301 (sex age)	?	?	[25]
John *et al.* (2011)	$VO_{2\ max}$ ml·min^{-1}·kg^{-1}, male	-531.795 + (13.26. age) + (10.542. height) + (15.503. weight)	0.51	?	[26]
John *et al.* (2011)	$VO_{2\ max}$ ml·min^{-1}·kg^{-1}, female	-1144.126 + (-7.911. age) + (10.542. height) + (12.436. weight)	0.66	?	[26]
Koch *et al.* (2009)[§]	$VO_{2\ max}$ ml·min^{-1}·kg^{-1}, both	61.3721 – 1.9479 (age) – 0.3053 (age^2) – 9.1229 (sex) + 3.8892 (BMI) – 1.9492 (age. BMI) – 6.7455 (sex. BMI) + 0.0716 (age. sex) + 1.6900 (age. sex. BMI)	?	?	[43]
Nelson *et al.* (2010)	$VO_{2\ max}$ ml·min^{-1}·kg^{-1}, male	54.3 – 0.361·(age)	0.13	?	[38]
Singh *et al.* (1989)	$VO_{2\ max}$ ml·min^{-1}·kg^{-1}, male	67.7 – 0.77 (age)	0.55	?	[32]
Tammelin *et al.* (2004)	VO_{2max} ml·min^{-1}·kg^{-1}, male	76.02 – 0.11 (HR) – 0.79 (BMI) – 1.35 (PA)	0.51	4.49	[34]
Tammelin *et al.* (2004)	VO_{2max} ml·min^{-1}·kg^{-1}, female	61.07 – 0.11 (HR) – 0.56 (BMI)- 1.34 (PA)	0.49	4.42	[34]
Work rate, kpm/min					
Arstila *et al.* (1990)	Work rate, W min^{-1} (W_{max6}), male	-164.7 + 2.222 (height) – 0.2422 (weight) – 1.435 (age)	?	?	[10]
Arstila *et al.* (1990)	Work rate, W·min^{-1} (W_{max6}), female	123.5 – 0.1288 (height) – 0.8984 (weight) – 1.070 (age)	?	?	[10]
Blackie *et al.* (1989)	Work rate, W_{max}), kpm/min male	1.704 + 6.1 (height) – 26.1 (age) + 0.04 (weight)	0.43	167	[39]
Blackie *et al.* (1989)	Work rate, W_{max}), kpm/min female	52 + 7.4 (height) – 13 (age) + 3.78 (weight)	0.50	110	[39]
Brudin *et al.* (2013)	Work rate, kpm·min^{-1}, male	(-0.96 [ln (age)]^2 + 6.34 ln (age) – 7.73) (1.04 height – 85)	0.96	?	[14]
Brudin *et al.* (2013)	Work rate, kpm·min^{-1}, female	(-0.50 [ln (age)]^2 + 3.17 ln (age) – 3.25) (1.02 height – 68)	0.96	?	[14]
Farazdaghi *et al.* (2001)**	Work rate, kpm·min^{-1}, female	(137. 7 height – 23. 1)/(1+ exp (0.064. (age – 75.9)))	0.66	?	[19]
Itoh *et al.* (2013) (cycle data)	Work rate, kpm·min^{-1}, both	3.55 – 0.0200 (age) – 0.281 (sex (female)) + 0.00327 (sex age) – 0.465 (protocol (10W/min) ?	?	?	[25]

[†] Sex, male, 0; female, 1; age, years; height (ht), centimeters; Weight (kg); HR: Heart rate after step test, beat.min^{-1}; PA: Frequency of brisk physical activity, times per week; BMI: Body mass index; R^2: Coefficient of determination; SEE: Standard error of the estimate

[‡] Farazdaghi: Height (m); [§] Koch: body mass index (BMI) was code as 0 for BMI ≤25 kg.m^{-2} and 1 for BMI >25 kgm^{-2}. There were five age groups: 25–34, 35–44, 45–54,55–64 and ≥64 yrs of age. The content of each column was multiplied with the code characteristics.

Table 3. Reference values for cardiopulmonary exercise testing (cont.).

Study (year)	Variables	Equations[†]	R^{2}[#]	SEE[#]	Ref.
Work rate, kpm/min (cont.)					
Joh *et al.* (2011)	Work rate, kpm·min^{-1}, male	5.05 + (-1.998 age) + (1.056 height) + (1.344 weight)	0.32	?	[26]
John *et al.* (2011	Work rate, kpm·min^{-1}, female	-581.974 + (-1.962 age) + (4.778 height) + (0.492 weight)	0.30	?	[26]
Jones *et al.* (1985)	Work rate, kpm·min^{-1}, male	-2759 + 25.3 (height) – 9.06 (age)	0.72	245	[27]
Jones *et al.* (1985)	Work rate, kpm·min^{-1}, female	-756 + 9.5 (height) – 9.21 (age) + 6.1 (weight)	0.67	177	[27]
Jones *et al.* (1989)	Work rate, kpm·min^{-1}, male	1506. height$^{2.70}$·age$^{-0.46}$	0.78	?	[28]
Jones *et al.* (1989)	Work rate, kpm·min^{-1}, female	969. height$^{2.80}$·age$^{-.43}$	0.77	?	[28]
Ong *et al.* (2002)	Work rate, kpm·min^{-1}, both	4.1394 – 0.0103 (age) – 0.3131 (sex) + 0.0076 (height) + 0.0058 (weight)	0.75	?	[42]
HR, beats/min					
Edvardsen *et al.* (2013)	HR, beats min^{-1}, male	220 – 0.88 (age)	0.7	?	[18]
Edvardsen *et al.* (2013)	HR, beats min^{-1}, female	208 – 0.66 (age)	0.7	?	[18]
Farazdaghi *et al.* (2001)[‡]	HR, beats min^{-1}, female	190.2/ (1+exp (0.0453. (age-107.5)))	?	?	[19]
Hollenberg *et al.* (1998)	HR, beats min^{-1}, male	220.62 – 1.10 (age) + 4.93 (height)	0.23	?	[23]
Hollenberg *et al.* (1998)	HR, beats min^{-1}, female	207.28 – 0.94 (age) + 4.53 (height)	0.17	?	[23]
Inbar *et al.* (1994)	HR, beats min^{-1}, male	205.0 – 0.605 (age)	0.60	6.4	[24]
Itoh *et al.* (2013) (cycle data)	HR, beats min^{-1}, both	191.7 – 0.743 (age) – 3.88 (sex) + 0.0669 (sex · age)	?	?	[25]
Itoh *et al.* (2013) (treadmill data)	HR, beats min^{-1}, both	202.8 – 0.763 (age) – 11.10 (sex) + 0.209 (sex · age)	?	?	[25]
John *et al.* (2011)	HR, beats min^{-1}, male and female	194.473 – 0.865 (age)	0.12	?	[26]
Jones *et al.* (1985)	HR, beats min^{-1}, male	206 – 0.8 (age)	0.72	11.6	[27]
Jones *et al.* (1985)	HR, beats min^{-1}, female	198 – 0.63 (age)	0.73	8.9	[27]
Nelson *et al.* (2010)	HR, beats min^{-1}, male	220.0 – 0.970 (age)	0.264	?	[38]
Ong *et al.* (2002)	HR, beats min^{-1}, both	5.2183 – 4.7100 10^{-5} (age) – 0.0108 (sex)	0.5012	?	[42]

[†] Sex, male, 0; female, 1; age, years; height (ht), centimeters; Weight (kg); HR: Heart rate after step test, beat.min^{-1}; PA: Frequency of brisk physical activity, times per week; BMI: Body mass index; R^{2}: Coefficient of determination; SEE: Standard error of the estimate

[‡] Farazdaghi: Height (m); [§] Koch: body mass index (BMI) was code as 0 for BMI ≤25 kg.m^{-2} and 1 for BMI >25 kgm^{-2}. There were five age groups: 25–34, 35–44, 45–54,55–64 and ≥64 yrs of age. The content of each column was multiplied with the code characteristics.

Table 4. The maximal oxygen uptake (VO_{2max} $L\cdot min^{-1}$) of different studies.

Study (year)	Males						Females						Ref.
	Age (years)						Age (years)						
	20–29	*30–39*	*40–49*	*50–59*	*60–69*	*70–79*	*20–29*	*30–39*	*40–49*	*50–59*	*60–69*	*70–79*	
Neder *et al.* (2001)	2.621	2.621	2.085	2.085	1.585	1.585	1.679	1.679	1.319	1.319	1.052	1.052	[9]
Aspenes *et al.* (2011)	4.3	4.22	4.01	3.62	3.23	2.71	2.77	2.74	2.63	2.35	2.15	1.79	[11]
Davis *et al.* (1997)	3.25	3.01	2.68	2.5	1.97		1.94	1.76	1.62	1.39	1.28		[15]
Edvardsen *et al.* (2013)	3.91	3.84	3.56	3.14	2.74	2.45	2.66	2.54	2.33	2.14	1.94	1.54	[18]
Fairbarn *et al.* (1994)	3.58	3.42	3.33	3.03	2.44	1.88	2.67	2.58	2.20	1.77	1.58	1.35	[41]
Inbar *et al.* (1994)	2.88		2.38		1.86								[24]
Meyer *et al.* (1994)	3.393	3.061	2.817	2.589									[30]
Nelson *et al.* (2010)		3.6	3.4	3.1	2.8								[38]
Singh *et al.* (1989)	3	2.6	2.2	1.7									[32]

Results

Selected studies

We identified 2313 potential studies. After initial screening, 62 were regarded potentially eligible. After reading the full text, 35 studies were found eligible for inclusion [9–43]. A flowchart displaying exact details of the selection process, including the reasons of exclusion, is presented in Figure 1.

Study characteristics

Table 1 shows overall study characteristics and reported adherence to ATS/ACCP guideline [4]. The 35 studies assessed 25,826 adults in total. A total of 15,550 participants were male and 10,276 female. The age of included subjects were between 20 and 70 years. The recruitment period ranged from 1985 to 2013. CPET was performed in 23 studies using a cycle ergometer, 13 studies using a treadmill and 1 study used both. There was a wide variety in CPET protocols, 10 studies used 15 or 16 W/min incremental protocol. Other studies used different incremental protocols or a modified version from an existing protocol. The studies were performed in 14 different countries, most represented countries were the US (n = 6), Sweden (n = 5), and Canada (n = 5). Sample size ranged from 25 to 4631.

Methodological quality assessment

The quality of the included studies varied, and none of the studies fulfilled all 14 quality criteria. A 'quality score' between 10 and 14 was seen in 4 studies, 16 studies received a score of 7–9 and 15 studies received a score of ≤6. Frequently observed weaknesses were a lack of power analysis, quality assurance of equipment and methodologies and reference equation validation. Table 2 provides a detailed overview of the quality of the reviewed studies.

Meta-analysis

Each of the included studies has various numbers of shortcomings and limitations that are noted in Table 2. Meta-analysis of the data is not meaningful, because of heterogeneity of methods and subjects (including sampling bias, uneven quality of primary data and inadequate statistical treatment of the data).

Results of individual studies

Table 3 shows the RV for VO_{2max}, WR_{peak} and HR_{max}. Studies differed in reporting their RV. Studies that did report the RV using regression equations are included in Table 3. Another method to report RV was by presenting mean values in a table. An overview of different VO_{2max} values is presented in Table 4, WR_{peak} is presented in Table 5, and HR_{max} is presented in Table 6. The mean values in Tables 4 to Table 6 are separately reported for male and female in six different age groups.

VO_{2max}

Three different outcomes of VO_{2max} were found, VO_{2max}, $ml\cdot min^{-1}$; VO_{2max}, $l\cdot min^{-1}$; and VO_{2max} $ml\cdot min^{-1}$ kg^{-1}. Four

Table 5. The maximal workload (WR_{peak}, Watts) of different studies.

Study (year)	Males						Females						Ref.
	Age (years)						Age (years)						
	20–29	*30–39*	*40–49*	*50–59*	*60–69*	*70–79*	*20–29*	*30–39*	*40–49*	*50–59*	*60–69*	*70–79*	
Arstila *et al.* (1990)		223	203	195	161			144	134	128	112		[10]
Brudin *et al.* (2013)	273.95	268.35	256.4	229.8	201.8	162.1	175.5	168.6	160.65	144.5	127.4	104.2	[14]
Fairbarn *et al.* (1994)	255	241	232	208	175	141	188	177	151	132	109	85	[41]
Farazdaghi *et al.* (2001)							200	190	181	168	137	100	[19]
Meyer *et al.* (1994)	225	222	215	188									[30]
Nordenfelt *et al.* (1985)	223	211	199	172	139	106	135	133	131	120	101	82	[31]
Singh *et al.* (1989)	195.3	185.5	158.6	129.2									[32]
Wohlfart *et al.* (2003)							303	288	240	257	192	186	[36]

studies used VO_{2max} $ml{\cdot}min^{-1}$ as outcome for the CPET equations, two equations for a male [24,40], one for female [40] and two for both genders [42,43]. There were seven studies that used VO_{2max}, $l{\cdot}min^{-1}$, seven different equations for male and two females [18,27,29,32,38,39,41,44]. The equations coefficient of determination (R^2) ranged between 0.16 and 0.80. Standard error of the estimate (SEE) varied from 0.31 to 0.483. Total of nine studies presented their VO_{2max} RV outcome in VO_{2max} $ml{\cdot}min^{-1}{\cdot}kg^{-1}$ [18,20,23,25,26,32,34,38,43]. Seven studies reported an equation for male, five female and three both. R^2 ranged between the lowest score of 0.13 to highest score of 0.67. Only one study reported the SEE; this was 4.49 for males and 4.42 for females [34]. VO_{2max} values in Table 4 were consistently higher in male than in female across age groups. VO_{2max} decreased with age. Younger subjects possessed a higher VO_{2max} than older subjects, and males higher than females.

Workload

Nine studies reported equations for WR_{peak} [10,14,19,25–28,39,42]. These studies reported six different equations for male, seven for female and two for both genders. R^2 ranged between the lowest score 0.30 to the highest score 0.96. Two studies reported the SEE ranging between 110 and 245 [27,39]. Males systematically scored higher than female in mean WR_{peak} and younger subjects scored higher WR_{peak} values than older subjects (Table 5).

Heart rate

Nine studies provided equations for HR_{max} [18,19,23–27,38,42]. These studies reported five different equations for male, four for female and four for both genders. The HR_{max} equations R^2 ranged between 0.12 and 0.73. Two studies reported the SEE ranging between 6.4 and 11.6 [24,27]. HR_{max} decreased with age (Table 6). Younger subjects possessed a higher HR_{max} than older subjects. There were some small differences between males and females.

Discussion

The aim of our study was to review the existing RV for CPET in healthy adults. We found 35 articles with RV for CPET. In the studies included, the Caucasians population was by far most frequently studied. CPET was assessed in 14 different countries. Despite all recent research on RV, the applicable populations are small. The appropriate knowledge of normal range in VO_{2max}, WR_{peak}, HR_{max} and other CPET variables during the CPET is necessary to interpret the results of a CPET. Females were relatively understudied compared to males. As many as 34% fewer female subjects were measured in the included studies.

Cycle ergometry was the most commonly employed CPET method. Furthermore, there was a wide variety in CPET protocols, equipment, study population, methodology and measurements reported.

A total of 16 studies reported VO_{2max} equations (Table 3). There are similarities and dissimilarities in the equations. However, they are difficult to compare, because studies have differences in equations terms, protocols, sample size and VO_{2max} might be determined differently. From the studies included, VO_{2max} was dependent on age, sex and anthropometric properties and could be affected by training status. The studies used different methods to normalize VO_{2max} by index of body size. However, there was

Table 6. The maximal heart rate (HR_{max}, beats·min^{-1}) of different studies.

Study (year)	Males						Females						Ref.
	Age (years)						Age (years)						
	20–29	*30–39*	*40–49*	*50–59*	*60–69*	*70–79*	*20–29*	*30–39*	*40–49*	*50–59*	*60–69*	*70–79*	
Arstila *et al.* (1990)		177	166	160	149			167	165	159	144		[10]
Brudin *et al.* (2013)	184.75	180.95	172.95	164.5	156	145.1	182.1	173.7	169.95	163.45	154.65	143.58	[14]
Edvardsen *et al.* (2013)	193.7	189.4	162.3	170.2	163	151.7	189.5	184.7	179.6	172.8	165.9	156.8	[18]
Fairbarn *et al.* (1994)	185	181	176	170	152	153	186	180	173	163	156	142	[19]
Farazdaghi *et al.* (2001)							186	184	179	174	167	155	
Hollenberg *et al.* (1998)				161	155.5	144.5				159	154	145.5	[23]
Inbar *et al.* (1994)	185		172		161								[24]
Meyer *et al.* (1994)	177	176	168	157									[30]
Nelson *et al.* (2010)		182.9	176.2	168.9	158.7								[38]
Nordenfelt *et al.* (1985)	177	178	174	159	142	145	176	177	168	163	153	131	[31]
Singh *et al.* (1989)	182.6	180.2	177.6	178									[32]
Wolhfart *et al.* (2003)							191	182	181	170	154	151	[36]

no uniform consensus in the literature on the best method for adjusting these indices. The most used normalization in the selected studies was bodyweight in kilograms. However, it may not be the appropriate method of reference for comparing or 'normalizing' the metabolic rate across subjects of different size, for instance, in underweight and overweight/obese subjects [45]. More studies are needed to gain consensus on this topic.

Meta-analysis of the data was not meaningful, because of heterogeneity of methods and subjects. Previous research pooling the large number of reports on predicted values for VO_{2max} demonstrated the limitations of such an approach [46]. Moreover, when in further research more homogeneity in methods and subjects is performed, meta-analysis is a valuable option.

Strengths & limitations

The quality of the included studies varied and none of the studies fulfilled all the 14 quality criteria. Most of the studies had a poor or moderate 'quality score'. Frequently observed weaknesses were a lack of power analysis, quality assurance of equipment and methodologies and reference equation validation. Almost all studies lack of reporting a sample size calculation and quality assurance, maybe these criteria are too strict. However in our opinion this was of vital importance for the quality of the studies. Furthermore, these criteria points are consistent with the ATS/ACCP guidelines [4].

It is notable that most of the studies did have a poor description of the study population. For example, the study of Neder *et al.* (2001) did not report which study population they measured; based on the study description, the population could be Brazilian or British [9]. Only four studies have an exclusion of different racial groups. Ideally, studies should report separately on different ethnic groups.

The findings of this review are difficult to compare with the past two reviews done in 1966 and 1990 on this subject [6,7], because studies before 1985 were excluded in this study. RV obtained before 1985 might not be applicable to contemporary subjects because of population differences (e.g., physical (in) activity levels, body composition) and the changes in methodology of the CPET (e.g., equipment, protocols).

The differences in reported RV emphasize the fact that each country or region should have its own RV and that RV may change over time and should be regularly updated. This is in accordance with the recommendations of the ATS/ACCP guidelines, each exercise laboratory must select an appropriate set of RV that best reflects the characteristics of the population tested, and equipment and methodology utilized [4]. It is also recommended that tests should include 10 healthy males and 10 healthy females of similar age, anthropometric characteristics, level of PA and be relative to the studied patient and that the results should be compared empirically with different sets of RV. The RV set that better characterizes the sample of healthy volunteers tested as normal should be selected. This review might help the clinician with reference values they should chose. For this purpose, we devised a flow chart (Figure 2) to facilitate this process. Furthermore, the identified sets with RV can be used by CPET equipment manufacturers to program these data into their software.

Recommendation for future research

A small variety of population norms is a clear lack. Caucasian, Japanese and Scandinavian populations were most frequently included, whereby the Caucasian white men were by far the most measured. Data from other populations from Asia, Middle-East, Africa and South America are needed. Furthermore, RV may change over time and should be regularly updated/validated. Therefore, standardization of normal RV processes/practice for CPET is necessary to facilitate interpretation and optimize clinical application [47,48]. RV beyond a healthy population, such as disease/disability specific RV for CPET, are of interest for clinical populations. The development of RV for different clinical populations could be used for better understanding of effect of intervention or better understanding of dosage.

Furthermore, additional studies reporting on other (sub)maximal CPET parameters, such as ventilatory anaerobic threshold, VE/VCO$_2$ slope and oxygen uptake efficiency slope, are warranted.

Conclusion

Thirty-five studies of RV for CPET in healthy adults were identified that were published in the last three decades. In the most studies, there are significant differences in the population characteristics, sample size, equipment, methodology and measurements reported. It was concluded that none of these studies provided an optimal set of CPET RV. Therefore, each exercise laboratory must select appropriate set of RV that best reflects the characteristics of the population tested, equipment, and methodology utilized.

Expert commentary

RV for CPET parameters are important for its interpretation but are still a challenging area. Consensus on data harmonization regarding reporting and determination of CPET parameters is required as well as consensus on the methodology to generate RV. In addition, more emphasis must be placed on quality control of CPET data using biological calibration or a metabolic simulator. More emphasis should be placed on the reporting of specific software and hardware settings of the equipment used as well.

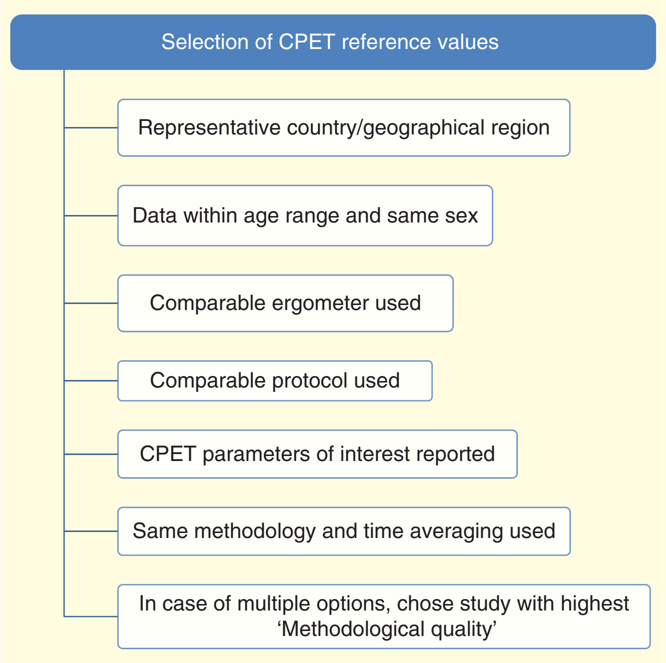

Figure 2. Flow chart for the selection of cardiopulmonary exercise testing reference values.

The current 'reference against the normal range' approach in CPET interpretation does not provide a ranking of an individual patient relative to age, sex, and disease/disability-matched peers. The development of disease/disability specific RV for CPET, as has been done for patients with congenital heart disease, will help to overcome this limitation [49].

Researchers with unidentified datasets of healthy subjects are encouraged to submit their manuscripts and datasets to the authors to facilitate a future update of the current review.

Five-year view

Within 5 years, consensus has been reached over the standardization of normal RV processes as well as the practice for analysis of CPET data. Consensus has been reached over, for instance, time averaging, definition of variables measured, protocols used as well as the reporting of CPET data in scientific publications.

Furthermore, manufacturer-independent open source software such as XINT [50] is increasingly used to standardize the analysis and interpretation of CPET data. Currently identified RV are implemented by manufacturers of CPET systems in their software packages.

Disease/disability specific RV for CPET are being developed and published for many different conditions.

Financial & competing interests disclosure

The authors have no relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript. This includes employment, consultancies, honoraria, stock ownership or options, expert testimony, grants or patents received or pending, or royalties.

No writing assistance was utilized in the production of this manuscript.

Key issues

- There is no single set of ideal reference values (RV); the population characteristics of each population are too diverse to pool the data in a single equation.
- Each exercise laboratory must select appropriate set of RV that best reflect the characteristics of the population/patient tested, and equipment and methodology utilized.
- Normal RV provide the comparative basis for answering important questions concerning the normality of exercise responses in patients and can significantly impact the clinical decision-making process.
- Maximal oxygen uptake is dependent on age, sex and anthropometric properties, and can be affected by training status.
- Peak workload decreased with age, and males systematically scored higher than females.
- Peak heart rate decreased with age, with some small differences between males and females.
- There is a small variety in population in which norms are established; Caucasian, Japanese and Scandinavian populations were most frequently studies, whereby Caucasian white men were by far the most measured. Data from other populations from Asia, Middle-East, Africa and South America are needed.
- RV may change over time and should be regularly updated/validated.
- Standardization of the methodology to generate RV, reporting of cardiopulmonary exercise testing parameters, reporting on specific software and hardware settings of the equipment and data harmonization are necessary to facilitate interpretation and optimize clinical application of cardiopulmonary exercise testing.

References

Papers of special note have been highlighted as:
• of interest
•• of considerable interest

1. Guide to physical therapist practice. second edition. American physical therapy association. Phys Ther 2001;81(1):9-746
2. Noonan V, Dean E. Submaximal exercise testing: clinical application and interpretation. Phys Ther 2000;80(8):782-807
3. Mezzani A, Agostoni P, Cohen-Solal A, et al. Standards for the use of cardiopulmonary exercise testing for the functional evaluation of cardiac patients: a report from the Exercise Physiology Section of the European Association for Cardiovascular Prevention and Rehabilitation. Eur J Cardiovasc Prev Rehabil 2009;16(3):249-67

   •• **The European Association for Cardiovascular Prevention and Rehabilitation cardiopulmonary exercise testing (CPET) guideline.**
4. ATS/ACCP Statement on cardiopulmonary exercise testing. Am J Respir Crit Care Med 2003;167(2):211-77

   •• **The American Thoracic Society/American College of Chest Physicians (ATS/ACCP) guideline.**
5. Wasserman K, Hansen JE, Sue DY, et al. Principles of exercise testing and interpretation. J Cardiopulm Rehabil Prev 1987;7(4):189

   •• **Principles of Exercise Testing and Interpretation. An evaluation of the patient with exercise limitation due to pulmonary symptomatology.**
6. Shephard RJ. World standards of cardiorespiratory performance. Arch Environ Health 1966;13(5):664-72

   • **First review of literature of reference values (RV) CPET.**
7. Shvartz E, Reibold RC. Aerobic fitness norms for males and females aged 6 to 75 years: a review. Aviat Space Environ Med 1990;61(1):3-11

   • **Meta- analysis of the most important RV CPET parameters.**
8. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Int J Surg (London England) 2010;8(5):336-41
9. Neder J, Nery LE, Peres C, Whipp BJ. Reference values for dynamic responses to incremental cycle ergometry in males and females aged 20 to 80. Am J Respir Crit Care Med 2001;164(8 I):1481-6
10. Arstila M, Impivaara O, Maki J. New ergometric reference values for clinical exercise tests. Scand J Clin Lab Invest 1990;50(7):747-55
11. Aspenes ST, Nilsen TI, Skaug EA, et al. Peak oxygen uptake and cardiovascular risk factors in 4631 healthy women and men. Med Sci Sports Exerc 2011;43(8):1465-73
12. Blackie SP, Fairbarn MS, McElvaney NG, et al. Normal values and ranges for ventilation and breathing pattern at maximal exercise. Chest 1991;100(1):136-42
13. Bromley PD, Hodges LD, Brodie DA. Physiological range of peak cardiac power output in healthy adults. Clin Physiol Funct Imaging 2006;26(4):240-6
14. Brudin L, Jorfeldt L, Pahlm O. Comparison of two commonly used reference materials for exercise bicycle tests with a Swedish clinical database of patients with normal outcome. Clin Physiol Funct Imaging 2014;34(4):297-307
15. Davis JA, Storer TW, Caiozzo VJ. Prediction of normal values for lactate threshold estimated by gas exchange in men and women. Eur J Appl Physiol Occup 1997;76(2):157-64
16. Davis JA, Storer TW, Caiozzo VJ, Pham PH. Lower reference limit for maximal oxygen uptake in men and women. Clin Physiol Funct Imaging 2002;22(5):332-8
17. Dubowy KO, Baden W, Bernitzki S, Peters B. A practical and transferable new protocol for treadmill testing of children and adults. Cardiol Young 2008;18(6):615-23
18. Edvardsen E, Scient C, Hansen BH, et al. Reference values for cardiorespiratory response and fitness on the treadmill in a 20- to 85-year-old population. Chest 2013;144(1):241-8
19. Farazdaghi GR, Wohlfart B. Reference values for the physical work capacity on a bicycle ergometer for women between 20 and 80 years of age. Clin Physiol (Oxford, England) 2001;21(6):682-7

20. Habedank D, Reindl I, Vietzke G, et al. Ventilatory efficiency and exercise tolerance in 101 healthy volunteers. Eur J Appl Physiol Occup Physiol 1998;77(5):421-6
21. 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
22. Herdy AH, Uhlendorf D. Reference values for cardiopulmonary exercise testing for sedentary and active men and women. Arq Bras Cardiol 2011;96(1):54-9
23. Hollenberg M, Ngo LH, Turner D, Tager IB. Treadmill exercise testing in an epidemiologic study of elderly subjects. J Gerontol A Biol Sci Med Sci 1998;53(4): B259-67
24. Inbar O, Oren A, Scheinowitz M, et al. Normal cardiopulmonary responses during incremental exercise in 20- to 70-yr-old men. Med Sci Sports Exerc 1994;26(5): 538-46
25. Itoh H, Ajisaka R, Koike A, et al. Heart rate and blood pressure response to ramp exercise and exercise capacity in relation to age, gender, and mode of exercise in a healthy population. J Cardiol 2013;61(1): 71-8
    - **RV study of CPET in Japan, with highest methological score in these review.**
26. John N, Thangakunam B, Devasahayam AJ, et al. Maximal oxygen uptake is lower for a healthy Indian population compared to white populations. J Cardiopulm Rehabil Prev 2011;31(5):322-7
27. Jones NL, Makrides L, Hitchcock C, et al. Normal standards for an incremental progressive cycle ergometer test. Am Rev Respir Dis 1985;131(5):700-8
28. Jones NL, Summers E, Killian KJ. Influence of age and stature on exercise capacity during incremental cycle ergometry in men and women. Am Rev Respir Dis 1989; 140(5):1373-80
29. Magrani P, Pompeu FA. [Equations for predicting aerobic power (VO(2)) of young Brazilian adults]. Arq Bras Cardiol 2010; 94(6):763-70
30. Meyer K, Hajric R, Samek L, et al. Cardiopulmonary exercise capacity in healthy normals of different age. Cardiology 1994;85(5):341-51
31. Nordenfelt I, Adolfsson L, Nilsson JE, Olsson S. Reference values for exercise tests with continuous increase in load. Clin Physiol 1985;5(2):161-72
32. Singh R, Singh HJ, Sirisinghe RG. Cardiopulmonary fitness in a sample of Malaysian population. Japanese J Physiol 1989;39(4):475-85
33. Sun XG, Hansen JE, Stringer WW. Oxygen uptake efficiency plateau: physiology and reference values. Eur J Appl Physiol 2012; 112(3):919-28
34. Tammelin T, Nayha S, Rintamaki H. Cardiorespiratory fitness of males and females of northern Finland birth cohort of 1966 at age 31. Int J Sports Med 2004; 25(7):547-52
35. Wisen AG, Wohlfart B. Aerobic and functional capacity in a group of healthy women: reference values and repeatability. Clin Physiol Functional Imaging 2004; 24(6):341-51
36. Wohlfart B, Farazdaghi GR. Reference values for the physical work capacity on a bicycle ergometer for men - A comparison with a previous study on women. Clin Physiol Functional Imaging 2003;23(3): 166-70
37. Vogel JA, Patton JF, Mello RP, Daniels WL. An analysis of aerobic capacity in a large United States population. J Applied Physiol (Bethesda, Md. 1985) 1986;60(2):494-500
38. Nelson MD, Petersen SR, Dlin RA. Effects of age and counseling on the cardiorespiratory response to graded exercise. Med Sci Sports Exercise 2010;42(2):255-64
39. Blackie SP, Fairbarn MS, McElvaney GN, et al. Prediction of maximal oxygen uptake and power during cycle ergometry in subjects older than 55 years of age. Am Rev Resp Dis 1989;139(6):1424-9
40. Storer TW, Davis JA, Caiozzo VJ. Accurate prediction of VO2max in cycle ergometry. Med Sci Sports Exercise 1990;22(5):704-12
41. Fairbarn MS, Blackie SP, McElvaney NG, et al. Prediction of heart rate and oxygen uptake during incremental and maximal exercise in healthy adults. Chest J 1994; 105(5):1365-9
42. Ong KC, Loo CM, Ong YY, et al. Predictive values for cardiopulmonary exercise testing in sedentary Chinese adults. Respirol (Carlton, Vic.) 2002;7(3):225-31
43. Koch B, Schaper C, Ittermann T, et al. Reference values for cardiopulmonary exercise testing in healthy volunteers: the SHIP study. Eur Res J 2009;33(2):389-97
44. Akinola AB, Land JM, Mathias CJ, et al. Contribution of nitric oxide to exercise-induced hypotension in human sympathetic denervation. Clin Autonomic Res 1999;9(5):263-9
45. Astrand P, Rodahl K. Textbook of work physiology. Physiological Bases of Exercise (3rd ed.). New York: McGraw-Hill Book Company; 1986. 412-85
46. Bailar JC 3rd. The promise and problems of meta-analysis. N Engl J Med 1997; 337(8):559-61
47. Weisman I, Zeballos R. Cardiopulmonary exercise testing: the need for standardization. Pulm Perspect 1992;9:5-8
48. Roca J, Whipp B, Agusti A, et al. Clinical exercise testing with reference to lung diseases: indications, standardization and interpretation strategies. Eur Respir J 1997; 10(11):2662-89
49. Kempny A, Dimopoulos K, Uebing A, et al. Reference values for exercise limitations among adults with congenital heart disease. Relation to activities of daily life–single centre experience and review of published data. Eur Heart J 2012;33(11): 1386-96
50. XINT. Available from: www.xint.org