AUTHOR QUERIES

DATE 3/24/2018
JOB NAME CPT

ARTICLE CPT-2017-010
QUERIES FOR AUTHORS Tolstrup Larsen et al

THIS QUERY FORM MUST BE RETURNED WITH ALL PROOFS FOR CORRECTIONS

Please confirm the given names (pink) and surnames (blue) of authors have been identified correctly.

AU1) Please check and confirm the suggested running head.
AU2) Please provide professional degrees (e.g., PhD, MD) for the authors "Rod S. Taylor, Patrick Doherty, Henning Langberg."
AU3) Please check the edits made in the affiliation "1."
AU4) Please provide the Department/Unit (if any) in affiliations "1, 5, 6."
AU5) Please provide City name in affiliation "7."
AU6) Please confirm the conflict of interest statement.
AU7) Please provide Department/Unit, University/Institution for the corresponding author.
AU8) Please check the sentence "Although AC offers an important alternative form of exercise testing…" for clarity.
AU9) Please note that as per style, the abstract sub headings should be "Purpose, Methods, Results, and Conclusions." Please check and change if necessary.
AU10) Please note that keywords have been taken from the provided pdf. Please check and correct if necessary.
AU11) The abbreviation "LC" has been defined as "leg cycling" as well as "leg cycle." Please check if the representation should be made consistent in the article.
AU12) The abbreviation "AC" has been defined as "arm cycling" as well as "arm cycle." Please check if the representation should be made consistent in the article.
AU13) Please spell out abbreviation "RCT" in the text and Table.
AU14) Please check the sentence "To fully use the values in a meta-analysis…" for clarity.
AU15) Please note that the spelling of the author has been given as "Astrand" in the reference section. Hence, the spelling "Aastrand" in the file has been replaced throughout by "Astrand" for consistency. Please check and confirm.
AU16) Please provide the volume number and page range in reference "15."
AU17) Please provide publisher location for references "17–19, 25."
AU18) Please provide publisher name and location for references "21, 23, 27."
AU19) Please provide the year of publication for the reference "25."
AU20) Please check the edits made in the reference "30."
AU21) Please check the layout of "Tables 1–3."
AU22) Please note that Table numbers "Table 2.1 and 2.2" have been changed to "Table 2 and Table 3," respectively. Please check.
AU23) Please note that Figure numbers "Fig 3.1, 3.2, 4.1, and 4.2" has been changed to "Fig. 3, 4, 5, and 6" respectively. Please check.
AU24) Please define the abbreviations "CAD," "CHD," "CT," "LVRS," "HF," and "PAD" in the footnotes of tables.

Comparison of Physical Fitness and Cardiopulmonary Exercise Test Performance Using Arm Versus Leg Cycling in Patients With Cardiovascular or Pulmonary Disease—A Systematic Review and Meta-analysis

Rasmus Tolstrup Larsen, PT, MSc;[1,2] *Lars Hermann Tang, PT, MSc, PhD;*[1,3] *Camilla Keller, PT;*[4] *Jan Christensen, PT, MSc;*[2,4] *Rod S. Taylor,*[1,5,6] *Patrick Doherty,*[7] *Henning Langberg, ~~Dr Med;~~*[2] *Ann-Dorthe Zwisler, MD*[1]

[1]REHPA, Danish Knowledge Centre for Rehabilitation and Palliative Care, Odense University Hospital, University of Southern Denmark, Denmark
[2]CopenRehab, Section of Social Medicine, Department of Public Health, University of Copenhagen, Copenhagen, Denmark
[3]Department of Rehabilitation and Nutrition, Faculty of Health and Technology, Metropolitan University College, Copenhagen, Denmark
[4]Department of Occupational- and Physiotherapy Therapy, Copenhagen University Hospital, Copenhagen, Denmark
[5]Institute of Health Research, University of Exeter Medical School, Exeter, United Kingdom
[6]National Institute of Public Health, University of Southern Denmark, Copenhagen, Denmark
[7]Department of Health Sciences, University of York, United Kingdom

~~Introduction:~~ Alternative modes of cardiopulmonary exercise testing are needed and ~~arm cycling~~ (AC) is a promising alternative to the gold standard of ~~leg cycling~~ (LC). The aim of this study was to undertake a systematic review comparing maximal oxygen consumption (VO_{2max}) obtained from AC and LC in patient populations with cardiovascular disease (CVD) and pulmonary disease (PD). **Methods:** A systematic review was undertaken with literature searches on December 5, 2016. Studies were included if they directly compared aerobic capacity values obtained from AC and LC in patients with CVD or PD. Results across studies were pooled using random effects meta-analysis and univariate meta-regression were used to assess potential associations between variables. **Results:** A total of 14 studies in 411 patients were included. On average, VO_{2max} of LC exceeded AC mean difference by 3.48 $mL \cdot kg^{-1} \cdot min^{-1}$, (95% confidence interval [CI]: 1.94, 5.03) and a mean AC/LC ratio of 0.83, (95% CI: 0.77, 0.90). VO_{2max} differences between AC and LC were similar in patients with CVD compared with PD but were found to be higher in older individuals and those with higher VO_{2max}. **Conclusions:** Although AC offers an important alternative form of exercise testing for patients with CVD or PD, clinicians must take into account that VO_{2max} values obtained from AC are consistently lower than those ~~seem to~~ LC. However, the results of this article offer an evidence-based estimation for the proportional differences between AC values and LC values for groups with CVD or groups with PD. **(Cardiopulm Phys Ther J. 2018;00: 1–12)** *Key Words: cardiopulmonary exercise testing, arm cycle, leg cycle, cardiovascular disease, chronic pulmonary disease, physical fitness*

INTRODUCTION

Cardiopulmonary exercise testing (CPET) is the recommended noninvasive assessment of the physical fitness both in healthy populations as well as in the context of rehabilitation in various patient groups, including patients with cardiovascular disease (CVD) and pulmonary disease (PD).[1-5] When CPET is used in rehabilitation, it typically has 3 purposes: an objective measure of patients' physical fitness, a way of prescribing exercise intensity, and a measure of intervention-specific change over time.[1-3]

Cardiopulmonary exercise testing is traditionally performed using lower limb exercise and often on a leg cycle (LC).[5-8] However, patients with CVD and chronic obstructive pulmonary disease (COPD) may be unable to perform CPET on an LC due to lower limb comorbidities such as peripheral vascular disease, neurological disease among patients,[9] or loss of muscle mass in the lower limbs.[10] Alternative modes of delivery such as CPET performed using an arm cycle (AC) are therefore needed.

A recently published systematic review and meta-analysis in healthy adults reported a mean difference in oxygen uptake of 12.5 $mL \cdot kg^{-1} \cdot min^{-1}$ favoring LC compared with AC, and a linear relationship with a ratio of 0.7 between the tests.[11] This systematic review was limited to studies on healthy adults and it is uncertain whether the results are generalizable to patient populations.[11] Although there may be equations to estimate LC CPET values from AC obtained CPET values, these equations are expected to have a large amount of error. However, evidence within patient populations confirms the expected lower VO_{2max} measurements during AC-testing compared with LC-testing (eg, patients with vascular surgery, COPD, orthotropic cardiac transplants, and other CVD conditions), which can be explained by the use of a smaller amount of muscle mass when performing AC testing.[9,12-14] In order for clinicians to use AC values to prescribe the correct exercise training intensity for patients with CVD or PD, the extent of the difference between obtained maximum values from AC and obtained maximum values from LC needs to be clarified.

The objective of this article is to undertake a systematic review and a meta-analysis of studies directly comparing aerobic capacity obtained from AC and LC in CVD and PD patient populations and subsequently to establish a ratio between AC and LC for use in everyday clinical practice.

Correspondence: Rasmus Tolstrup Larsen, PT, MSc, Henrik Pontoppidans Vej 6, 1st floor, DK-2200 Copenhagen N, Denmark (rala@sund.ku.dk).

The authors declare no conflicts of interest.

Supplemental digital content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal's Web site (www.cptj.com).

DOI: 10.1097/CPT.0000000000000081

METHOD

A protocol for this systematic review was published in the International Prospective Register of Systematic Reviews (PROSPERO—CRD42016048767),[15] and the reporting of the study was done according to the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) statement.[16]

Criteria for Considering Studies for This Review

Types of Studies. Randomized controlled trials, observational cohort studies, cross-sectional studies, and case-control studies were considered eligible for inclusion.

Types of Cardiopulmonary Exercise Testing and Outcome Measures. Studies directly comparing maximum or peak oxygen uptake (VO_2) values on AC and LC, preferably as milliliter oxygen per kilogram per minute ($mL \cdot min^{-1} \cdot kg^{-1}$) or otherwise as liters per minute (L/min), were considered eligible for inclusion. The VO_{2max} values had to be obtained from a nonassisted test (no external help, eg, functional electrical stimulation or therapist-assisted CPET), and patients had to perform both AC and LC testing in a within-comparison design.

Types of Patients. The following patient groups were included: Patients with CVD according to the World Health Organization (WHO) definition, ie, coronary heart disease, cerebrovascular disease, peripheral arterial disease, rheumatic heart disease, congenital heart disease, deep vein thrombosis, and pulmonary embolism.[17] Patients with PD according to the WHO definition, ie, asthma, bronchiectasis, chronic obstructive lung disease including COPD, bronchitis, and emphysema, chronic rhinosinusitis, hypersensitivity pneumonitis, lung cancer, and neoplasms of respiratory and intrathoracic organs, lung fibrosis, chronic pleural diseases, pneumoconiosis, pulmonary eosinophilia, pulmonary heart disease and diseases of pulmonary circulation including pulmonary embolism, pulmonary hypertension, cor pulmonale, rhinitis, sarcoidosis, and sleep apnea syndrome.[18]

Data Sources and Searches

Electronic searches of the databases CENTRAL, MEDLINE, EMBASE, and CINAHL were undertaken on December 5, 2016. Identified studies from preliminary searches were screened to identify additional search terms.

The search strategy consisted of a combination of relevant keywords and MeSH/Thesaurus terms for: (1) LC test, (2) AC test, and (3) physical fitness. To avoid excluding any relevant subgroups, evaluation of eligibility of the studies in terms of the patient population was done by reviewing the full-text publication. The authors of

unobtainable studies or studies with missing data were contacted.

Two authors (R.T.L. and C.K.) independently screened titles and abstracts and assessed eligible articles in full-text. Any inconsistencies between authors were solved by discussion and, if relevant, a third author (L.T.) was consulted.

Data Extraction

The following information was extracted from included studies: sample size, study design, patient population (CVD or PD), sex distribution, mean age, mean height, mean weight, mean body mass index (BMI), highest obtained VO_2 value from AC, highest obtained VO_2 value from LC, mean difference between the highest obtained VO_2 values, correlation between the highest obtained VO_2 values from AC and LC, mean peak respiratory exchange ratio (RER) from AC and from LC, protocol type (ramp/incremental), starting Watt level on AC and LC, increments on AC and LC, order of testing, and hours between testing. The preferred outcome was body mass relative VO_{2max} reported as $mL{\cdot}kg^{-1}{\cdot}min^{-1}$. However, older studies tend to report the absolute VO_{2max} instead of the body mass relative VO_{2max}.[11] Because of this and if possible, both types of outcome were extracted or calculated and used in the analyses. Two authors (R.T.L. and C.K.) independently extracted the above listed data from all included studies. Any inconsistencies between authors was discussed and solved with consultation of a third author (L.T.).

Risk of Bias Assessment

The Quality Assessment Tool for Observational Cohort and Cross-Sectional Studies[19] was used as a template for assessing the methodological quality of all included studies. Quality assessment items 5 and 9 to 12 (groups recruited from the same population, concurrent controls, exposure assessment priori outcome assessment, exposure measures and assessment, and blinding of exposure assessor) were not applicable for the research question in this review and the items did not contribute to the quality rating. Each item was assessed for "low risk," "unclear risk," or "high risk" of bias. Two authors (R.T.L. and C.K.) independently undertook the quality assessment. Any inconsistencies between authors were solved by discussion and, if relevant, a third author (L.T.) was consulted. The risk of bias assessment was performed rigorously according to the PROSPERO protocol.[15]

Data Analysis

Two studies included more than one group in the CPET (men/women or different groups in ~~an RCT~~) and hence did also report values for the specific groups.[12,20] Because of this, the analyses were performed with the results of the specific groups. The analyses on respectively $mL{\cdot}kg^{-1}{\cdot}min^{-1}$ and L/min were evaluated by 2 random-effects meta-analyses pooling the mean differences between the obtained AC values and the obtained LC values. Second, 2 random-effects meta-analyses pooling the ratios between the obtained AC values and the obtained LC values for respectively $mL{\cdot}kg^{-1}{\cdot}min^{-1}$ and for L/min were undertaken. The analyses were stratified on patient population (CVD or PD).

To perform the analyses of the mean difference of values reported as $mL{\cdot}kg^{-1}{\cdot}min^{-1}$ or L/min, the mean difference was calculated for each study. Positive mean differences favor the LC as having the largest values. The SD of this difference was calculated and adjusted for the within-subject correlation using the method of chapter 16.4.6.1 in the Cochrane Handbook.[21] As no studies reported correlation values, we chose to impute a correlation value of 0.5 as it represents a moderate positive relationship between the test values, which was described in the protocol.[15]

To obtain the values needed for the meta-analysis, the following calculation was performed for each individual study. With AC value be denoted as VO_{2AC}, LC value as VO_{2LC}, the correlation value as r, and the SDs as SD_{diff}, SD_{AC}, and SD_{LC}.

$$\text{Mean difference} = VO_{2LC} - VO_{2AC}$$

$$SD_{diff} = \sqrt{SD_{AC}^2 + SD_{LC}^2 - (2 \times r \times SD_{AC} \times SD_{LC})}$$

We calculated the ratio and used the values in a meta-analysis to express the association between the 2 types of CPET. The ratio between the maximal obtained VO_2 from AC and from LC was calculated for each study as a ratio of the mean values.[22] Ratios below 1.0 favor the LC as having the largest values. To fully use the values in a meta-analysis, a logarithmic transformation was needed ~~use~~ and the re-transformation was done in Stata using the *eform* command to get the pooled ratio and 95% confidence intervals (CIs). For a study reporting the values, let the ratio be denoted as VO_{2ratio}.

$$\ln(VO_{2ratio}) = \ln\left(\frac{VO_{2AC}}{VO_{2LC}}\right)$$

$$SE[\ln(VO_{2ratio})] = \sqrt{\frac{1}{n} \times \left(\frac{SD_{AC}}{VO_{2AC}}\right)^2 + \frac{1}{n} \times \left(\frac{SD_{LC}}{VO_{2LC}}\right)^2}$$

The levels of statistical heterogeneity were assessed using the I^2 score from each analysis. I^2 values from 0% to 25% were interpreted as the meta-analysis having a low level of heterogeneity, values from 26% to 50% as a moderate level, values from 51% to 75% as a high level, and from 76% to 100% as a considerable level.[23,24]

To avoid excluding studies from the meta-analyses, missing SDs were imputed from the median covariate-specific SD, respectively from CVD studies and PD studies, according to the Cochrane Handbook.[25]

Subgroup analyses were performed by stratification on patient population (CVD or PD). Sensitivity analyses on L/min outcome were performed. Sensitivity analyses on small study bias were performed using the Egger test and, if significant, a metatrim analysis was performed to evaluate small study bias from imprecise studies. Meta-analyses stratified on the risk of bias assessment were used to evaluate if any heterogeneity in the analyses of the mean difference and the ratio was associated with methodological quality (low risk, unclear risk, or high risk).

Univariate meta-regressions were performed on the following continuous outcome measures: mean age, mean BMI (calculated for studies not reporting BMI), sex distribution (percentage of males), and mean difference in peak RER values. Size of aerobic capacity (based on the Astrand classification—"low," "fair," "average," "good," or "high")[26] was analyzed as a categorical outcome. [AQ:15]

All analyses were performed using Stata 14.0 software (StataCorp, 2013. Stata Statistical Software: Release 14.9; College Station, TX: StataCorp LP). Flowchart and risk of bias is presented using Review Manager 5.3 (Cochrane collaboration) software. A P value $\leq .05$ was considered as statistically significant.

RESULTS

Study Inclusion

The electronic searches identified 4154 records. After removing 707 duplicates, 3447 studies remained. Screening title and abstract resulted in exclusion of 3222 studies as they did not meet the CPET inclusion criteria, and thus, 225 studies were considered eligible for review in full text. In the full-text review, 211 studies were excluded in total; 118 were excluded due to lack of patients with CVD/PD, 25 for not reporting VO_2 values, 46 for lack of AC testing, 7 for lack of LC testing, 1 for lack of within-comparison design, 1 for lack of data, and 12 studies were unobtainable in full-text versions or only available in versions unable to translate (Persian, Turkish, and Japanese). Thus, 14 studies (published between 1971 and 2009) were included in the review. The study selection process is summarized in Figure 1. [F1] Characteristics of the 14 included studies (17 groups, 411 participants) are listed in the Appendix, **Supplemental Digital Content 1** (see ~~Table 3~~, http://links.lww.com/CPTJ/A2). A summary of the included studies is listed in Table 1. [T1]

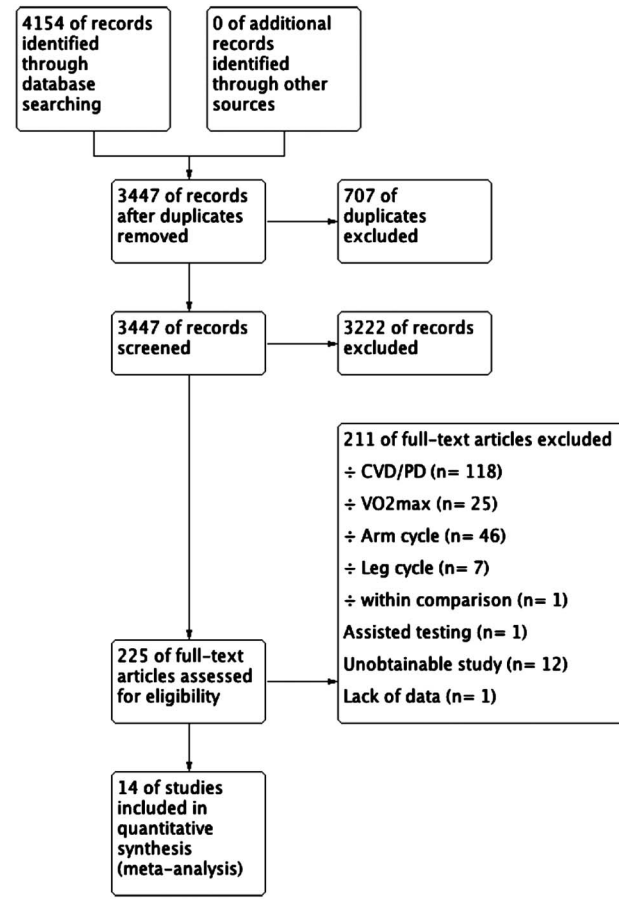

Fig. 1. Study flow diagram.

Risk of Bias in Included Studies

Risk of bias in included studies is illustrated in detail in Figure 2. [F2]

Mean Difference in Physical Fitness

The random-effects meta-analysis for the mean difference in $mL \cdot kg^{-1} \cdot min^{-1}$ is shown in Figure 3. [F3] A total of 14 groups from 11 studies (359 participants) reported data for the mean difference measured in $mL \cdot kg^{-1} \cdot min^{-1}$. The overall mean difference was 3.54 $mL \cdot kg^{-1} \cdot min^{-1}$, (95% CI: 2.08, 5.01), $I^2 = 91.5\%$, $P < .001$, favoring LC. For patients with CVD, a pooled mean difference of 4.30 $mL \cdot kg^{-1} \cdot min^{-1}$, (95% CI: 1.96, 6.65), $I^2 = 92.6\%$, $P < .001$, favoring LC was found. For patients with PD, a pooled mean difference of 2.25 $mL \cdot kg^{-1} \cdot min^{-1}$, (95% CI: 0.57, 3.94), $I = 86.7\%$, $P < .001$, favoring LC was found. According to the I^2 values, the results are likely to be affected by considerable heterogeneity.

Sensitivity Analysis on Mean Difference Measured in L/min

Random-effects meta-analysis for the mean difference in L/min is shown in Figure 4. [F4] All 17 groups from the 14 studies (411 participants) reported data for the mean difference measured in L/min. The overall mean difference was 0.25 L/min, (95% CI: 0.13, 0.23), $I^2 = 92.6\%$, $P < .001$, favoring LC. For patients with CVD, a pooled mean difference of 0.32 L/min, (95% CI: 0.17, 0.47), $I^2 = 94.8\%$, $P < .001$, favoring LC was found. For patients with PD, a pooled mean difference of 0.18 L/min, (95% CI: 0.13, 0.23), $I^2 = 83.7\%$, $P < .001$, favoring LC was found. According to the I^2 values, the results are likely to be affected by considerable heterogeneity.

TABLE 1

AQ:21 Study Characteristics of the 14 Included Studies

Population	Cardiovascular Disease N = 10, n = 207	Pulmonary Disease N = 7, n = 204	Combined N = 17, n = 411
Continent of publication, %			
North America	62.5	50	57.2
Europe	37.5	33.33	35.7
Australia	0	16.67	7.1
Study design, %			
RCT	12.5	0	7.1
Non-RCT	12.5	16.67	14.3
Case-control	25	33.33	28.6
Cross-sectional	50	50	50
Participant characteristics			
Gender, %			
Male only	60	57.1	47.1
Female only	0	14.3	5.8
Mixed	40	28.6	47.1
Participant characteristics, median (IQR)			
Mean age, years	59 (52–66)	62 (61–66.5)	61 (52.25–66.25)
Mean BMI, kg/m^2	26.6 (25–28.4)	26 (23.75–26.9)	26.35 (24.75–26.35)
Aerobic capacity, %			
Low	75	100	85.7
Average	25	0	14.3
Test characteristics, %			
Order on AC/LC test			
AC first	0	0	0
LC first	0	33.37	7.2
Random order	87.5	16.67	71.4
Not reported	12.5	50	21.4
Ramp	0	0	0
Incremental	37.5	33.33	35.7
Different protocol for AC/LC	0	16.67	7.2
Not reported	62.5	50	57.1
Test characteristics, median (IQR)			
Time between tests (hours)	24 (12.5–36)	2 (2–24)	24 (2–24)
AC start level (W)	12.5 (9–15.67)	0 (0–5)	9.5 (4.5–15)
LC start levels (W)	15.7 (9–27.26)	0 (0–0)	9 (0–24.46)
AC increase/min (W)	7 (5–7)	10 (5–15)	7 (5–7)
LC increase/min (W)	14 (8.17–14)	15 (10–30)	14 (10–16.34)

Aerobic capacity based on the definition of Astrand.

AC, arm cycle; BMI, body mass index; IQR, interquartile range; LC, leg cycle; $mL \cdot kg^{-1} \cdot min^{-1}$, milliliters per kilogram bodyweight per minute; N, number of groups; n, number of patients.

Sensitivity Analysis for Small Study Bias

The Egger test showed a significant result ($P = .030$) for $mL \cdot kg^{-1} \cdot min^{-1}$ and a nonsignificant result ($P = .15$) for L/min, suggesting that small study bias is present in the random-effects meta-analysis of the mean difference in $mL \cdot kg^{-1} \cdot min^{-1}$. This means that the smallest study reports the largest mean difference. The following metatrim analysis adjusted the overall mean difference to 1.95 $mL \cdot kg^{-1} \cdot min^{-1}$, (95% CI: 0.22, 3.68).

Analyses on the Ratios

The random-effects meta-analysis for the ratio is shown in Figure 5. A total of 14 groups from 11 studies [F5] (349 participants) reported results in $mL \cdot kg^{-1} \cdot min^{-1}$. The overall ratio was 0.83, (95% CI: 0.77, 0.89), $I^2 = 0\%$, $P = .909$, favoring LC. For patients with CVD, a pooled of ratio of 0.83, (95% CI: 0.75, 0.91), $I^2 = 0\%$, $P = .718$, favoring LC was found. For patients with PD, a pooled

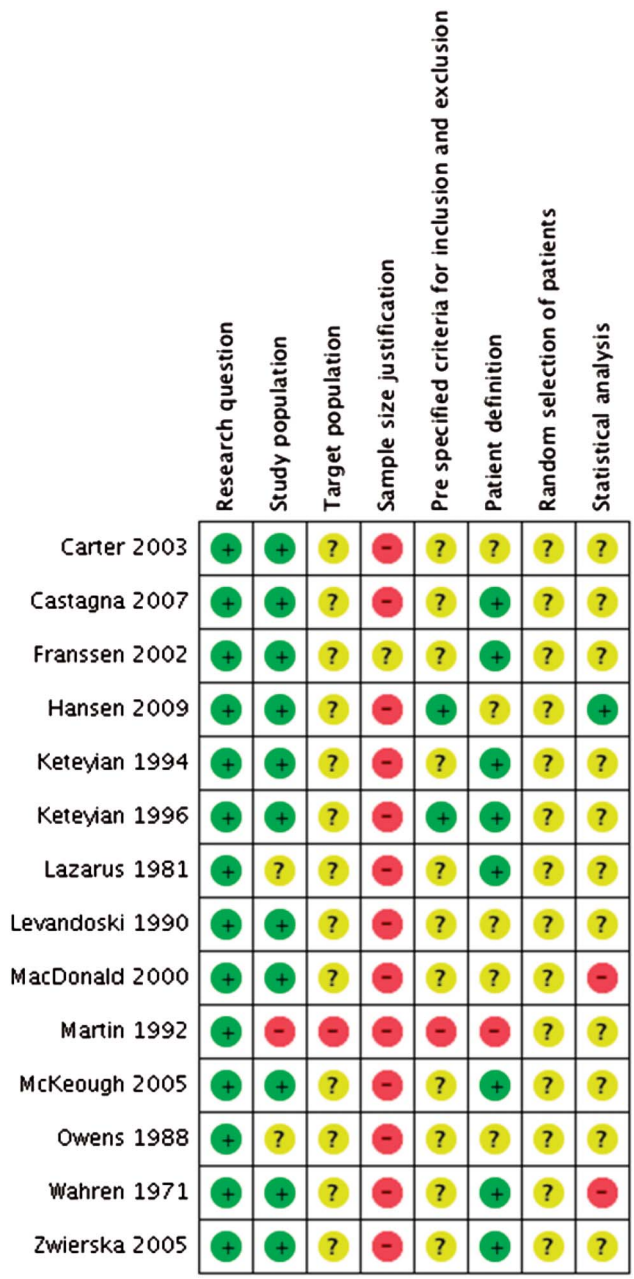

Fig. 2. Risk of bias summary: review authors' judgments about each risk of bias item for each included study.

ratio of 0.83, (95% CI: 0.74, 0.92), $I^2 = 0\%$, $P = .827$, favoring LC was found. According to the I^2 values, the results are not likely to be affected by heterogeneity.

Sensitivity Analysis on Results in L/min

[F6] The random-effects meta-analysis for the ratio is shown in Figure 6. All 17 groups from the 14 studies (411 participants) reported results in L/min. The overall ratio was 0.83, (95% CI: 0.78, 0.89), $I^2 = 0\%$, $P = .937$. For patients with CVD, a pooled ratio of 0.84, (95% CI: 0.77, 0.92), $I^2 = 0\%$, $P = .654$, favoring LC was found. For patients with PD a pooled ratio of 0.83, (95% CI: 0.75, 0.92), $I^2 = 0\%$, $P = .959$, favoring LC was found. According to the I^2 values, the results are not likely to be affected by heterogeneity.

Sensitivity Analysis for Small Study Bias

The Egger test showed no significant result ($P = .943$ and $P = .966$), respectively, for the ratios of the results in $mL \cdot kg^{-1} \cdot min^{-1}$ and L/min in the random-effects meta-analyses of the ratios.

Stratified Analysis on Risk of Bias

Figures 5.1–5.8, 6.1–6.8, 7.1–7.2, and 8.1–8.8, found in the appendix (see **Supplemental Digital Content 1**, http://links.lww.com/CPTJ/A2), show stratified analysis on the risk of bias. Mean difference results were significantly different on the risk of bias evaluation on study population in $mL \cdot kg^{-1} \cdot min^{-1}$ (see **Fig. 5.2**, **Supplemental Digital Content 1**, http://links.lww.com/CPTJ/A2), target population in $mL \cdot kg^{-1} \cdot min^{-1}$ (see **Fig. 5.3**, **Supplemental Digital Content 1**, http://links.lww.com/CPTJ/A2), prespecified criteria for inclusion and exclusion in $mL \cdot kg^{-1} \cdot min^{-1}$ (see **Fig. 5.5**, **Supplemental Digital Content 1**, http://links.lww.com/CPTJ/A2), and patient definition in $mL \cdot kg^{-1} \cdot min^{-1}$ (see **Fig. 5.6**, **Supplemental Digital Content 1**, http://links.lww.com/CPTJ/A2). The results were also significantly different on study population in L/min (see **Fig. 6.2**, **Supplemental Digital Content 1**, http://links.lww.com/CPTJ/A2), target population in L/min (see **Fig. 6.3**, **Supplemental Digital Content 1**, http://links.lww.com/CPTJ/A2), sample size justification in L/min (see **Fig. 6.4**, **Supplemental Digital Content 1**, http://links.lww.com/CPTJ/A2), prespecified criteria for inclusion and exclusion in L/min (see **Fig. 6.5**, **Supplemental Digital Content 1**, http://links.lww.com/CPTJ/A2), and patient definition in L/min (see **Fig. 6.6**, **Supplemental Digital Content 1**, http://links.lww.com/CPTJ/A2), all with a higher mean difference among the studies assessed with high risk of bias in the above-mentioned figures. No significant difference on the ratio was found on the risk of bias in both outcomes.

Univariate Meta-regressions on the Mean Difference

Results from univariate meta-regressions are shown in Table 2. Not all included studies reported all values for [T2] the groups. Hence, between 5 and 14 groups were included in the analysis on the $mL \cdot kg^{-1} \cdot min^{-1}$ outcome. Significant associations between mean difference in VO_2 in $mL \cdot kg^{-1} \cdot min^{-1}$ and mean age (coefficient -0.27, $P < .001$), Astrand classification (coefficient 5.58, $P = .027$), and mean aerobic capacity obtained from LC (coefficient 0.39, $P < .001$) were found. Between 7 and 17 groups were included in the analysis on the V02 L/min outcome. Significant associations between mean difference in L/min and mean age (coefficient -0.02, $P < .001$), Astrand classification (coefficient 0.43, $P = .018$), and mean

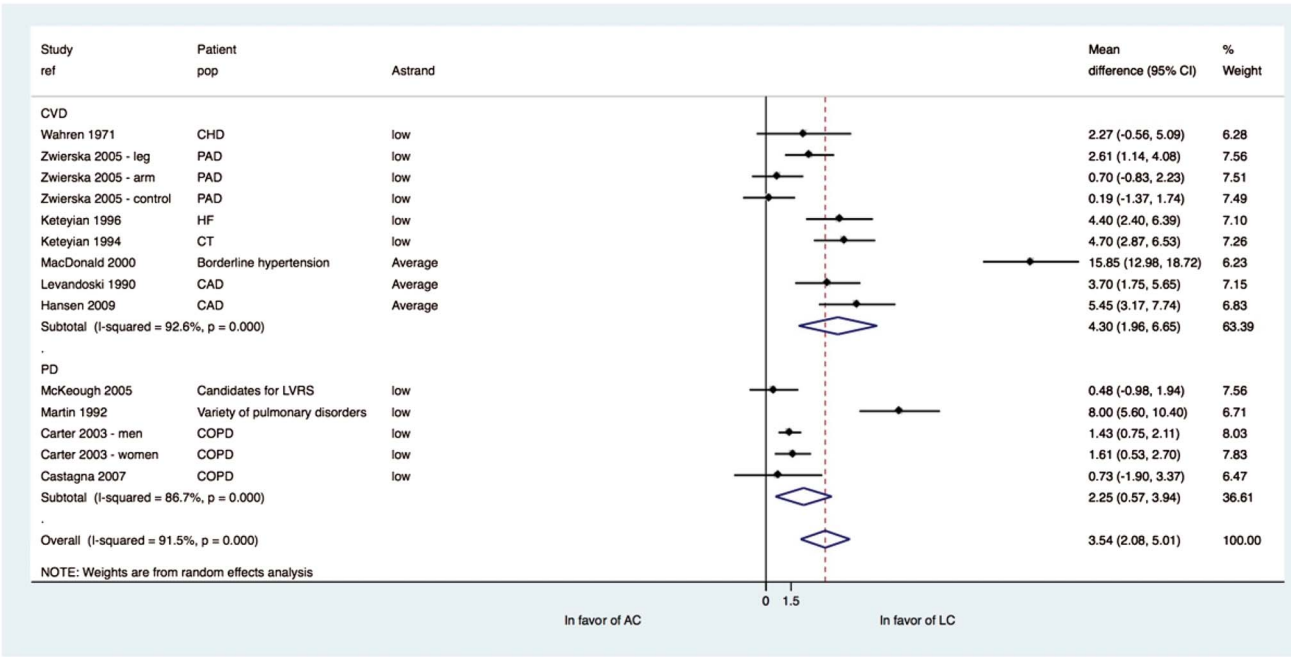

Fig. 3. Random-effects meta-analysis on the mean difference between the arm cycle and the leg cycle reported as $mL \cdot kg^{-1} \cdot min^{-1}$, stratified on cardiovascular disease (CVD)/pulmonary disease (PD). CI, confidence interval; COPD, chronic obstructive pulmonary disease.

aerobic capacity obtained from LC (coefficient 0.38, $P < .001$) were found. Adjusted R^2 values were above 80% in the univariate meta-regressions between mean difference ($mL \cdot kg^{-1} \cdot min^{-1}$ and L/min) and mean age and mean maximum obtained VO_2/mean aerobic capacity obtained from LC. Hence, most of the heterogeneity found in the meta-analysis on mean difference in $mL \cdot kg^{-1} \cdot min^{-1}$ (Fig. 3) and the meta-analysis on mean difference in L/min (Fig. 4) is explained by mean age and aerobic capacity.

Univariate Meta-regressions on the Ratios

Results from univariate meta-regressions are shown in Table 3. Not all included studies reported values for all variables for the included groups. Hence, between 5 and [T3]

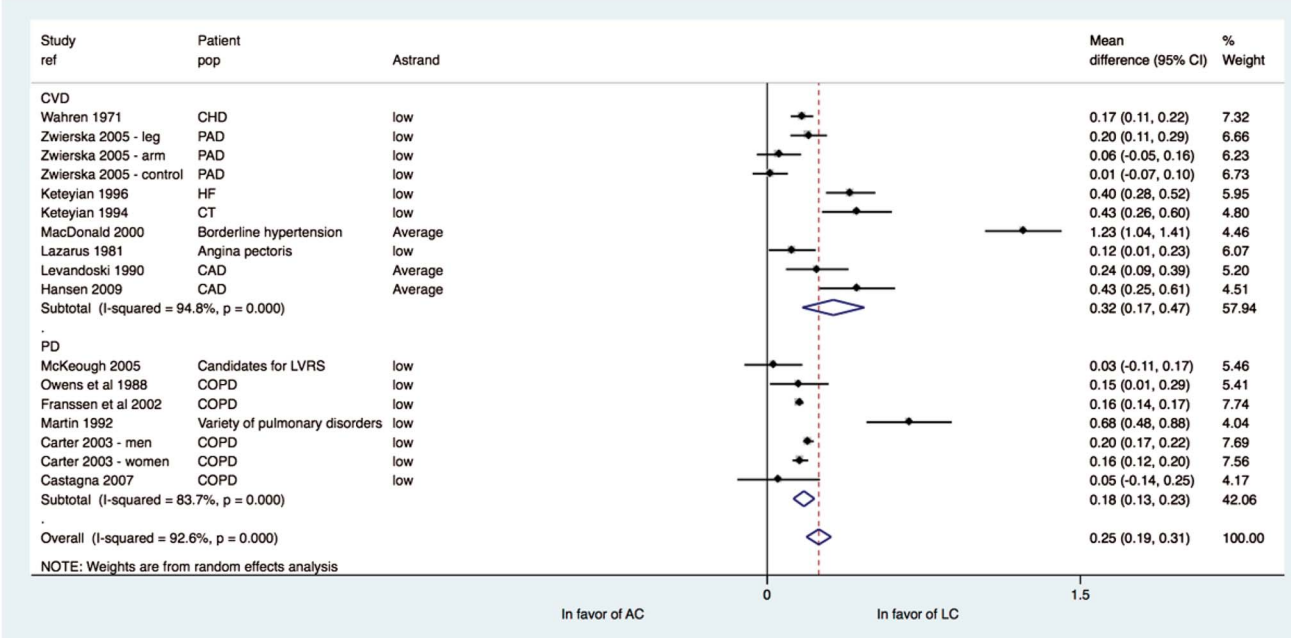

Fig. 4. Random-effects meta-analysis on the mean difference between the arm cycle and the leg cycle reported as L/min, stratified on cardiovascular disease (CVD)/pulmonary disease (PD). CI, confidence interval; COPD, chronic obstructive pulmonary disease.

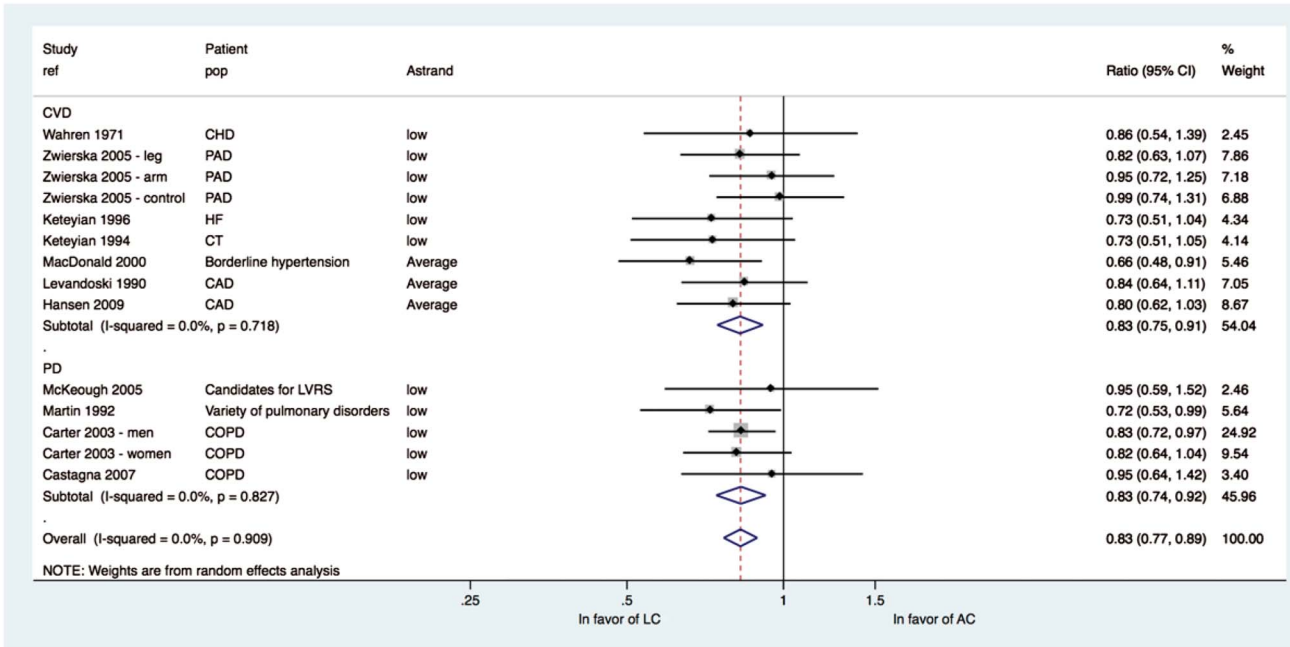

Fig. 5. Random-effects meta-analysis on the ratio between the arm cycle and the leg cycle reported as $mL \cdot kg^{-1} \cdot min^{-1}$, stratified on cardiovascular disease (CVD)/pulmonary disease (PD). CI, confidence interval; COPD, chronic obstructive pulmonary disease.

14 groups were included in the analysis on the ratio in $mL \cdot kg^{-1} \cdot min^{-1}$. No significant associations between the ratio in $mL \cdot kg^{-1} \cdot min^{-1}$ and independent variables were found. Between 7 and 17 groups were included in the analysis on ratio of the L/min outcome. Significant associations between the ratio in L/min and mean age (coefficient 0.01, $P = .047$) were found. No R^2 values were present as no heterogeneity was found in the meta-analyses on the ratio from Figures 5 and 6.

DISCUSSION

This systematic review included results from 17 groups and 411 patients. The overall mean difference

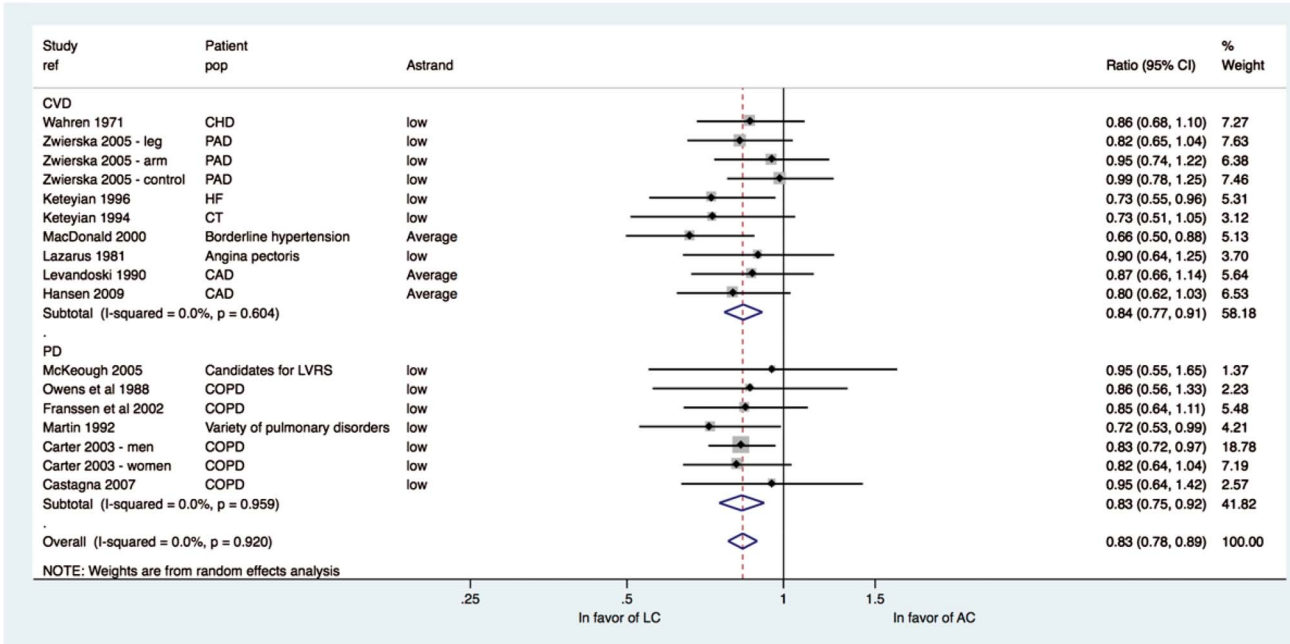

Fig. 6. Random-effects meta-analysis on the ratio between the arm cycle and the leg cycle reported as L/min, stratified on cardiovascular disease (CVD)/pulmonary disease (PD). CI, confidence interval; COPD, chronic obstructive pulmonary disease.

TABLE 2

AQ:22 Univariate Meta-Regressions on Mean Difference Between Arm Cycle Values and Leg Cycle Values

Independent Variable, $mL \cdot kg^{-1} \cdot min^{-1}$	N	Mean Coefficient (95% CI)	*P*	Adjusted R^2, %
Mean age	14	−0.27 (−0.35, −0.19)	<.001	91.60
Mean BMI	11	−0.72 (−2.72, 1.27)	.433	−3.52
Sex distribution	14	0.029 (−0.059, 0.12)	.484	−4.28
Diff RER	5	−76.10 (−413.40, 261.20)	.558	−20.46
Astrand	14	5.68 (0.84, 10.54)	.025	33.71
Mean aerobic capacity obtained from LC	14	0.39 (0.30, 0.48)	<.001	97.66
Independent Variable, L/min	**N**	**Mean Coefficient (95% CI)**	***P***	**Adjusted R^2, %**
Mean age	17	−0.021 (−0.027, −0.015)	<.001	83.96
Mean BMI	12	−0.038 (−0.17, 0.094)	.529	−6.14
Sex distribution	17	0.0019 (−0.004, 0.008)	.521	−4.32
Diff RER	7	−1.33 (−20.50, 17.8)	.865	−25.20
Astrand	17	0.43 (0.085, 0.78)	.018	31.93
Mean VO_{2max} obtained from LC	17	0.38 (0.30, 0.46)	<.001	94.39

Astrand: Astrand classification of aerobic capacity as categorical outcome (low/average).

BMI, body mass index; CI, confidence interval; diff RER, difference in respiratory exchange ratio (positive values favors leg cycle); $mL \cdot kg^{-1} \cdot min^{-1}$, milliliters per kilogram bodyweight per minute; N, number of groups included in the analysis; LC, leg cycle.

between AC and LC values was found to be 3.48 $mL \cdot kg^{-1} \cdot min^{-1}$, (95% CI: 1.94, 5.03) and the overall ratio was found to be 0.83 (95% CI: 0.77, 0.90). The mean difference measured in $mL \cdot kg^{-1} \cdot min^{-1}$ was significantly associated with age and aerobic capacity, meaning that studies with older patients, and lower aerobic capacity, reported a smaller mean difference. The ratio of the results in L/min was significantly associated with age, meaning that studies with older patients reported a larger ratio between the test values. The meta-analyses on the mean difference in $mL \cdot kg^{-1} \cdot min^{-1}$ (Fig. 3) and L/min (Fig. 4) indicate that the mean differences were similar when comparing patients with CVD to patients with PD. The main finding of the meta-analyses on the ratios of the results in $mL \cdot kg^{-1} \cdot min^{-1}$

TABLE 3

Univariate Meta-Regressions on the Ratio Between Arm Cycle Values and Leg Cycle Values

Independent Variable, $mL \cdot kg^{-1} \cdot min^{-1}$	N	Mean Coefficient (95% CI)	*P*	Adjusted R^2, %
Mean age	14	0.01 (−0.0004, 0.013)	.064	0
Mean BMI	11	0.027 (−0.048, 0.103)	.437	0
Sex distribution	14	−0.0004 (−0.0032, 0.0025)	.790	0
Diff RER	5	0.94 (−17.72, 19.6)	.883	0
Astrand	14	−0.081 (−0.28, 0.121)	.398	0
Mean aerobic capacity obtained from LC	14	−0.01 (−0.0176, 0.0018)	.101	0
Independent Variable, L/min	**N**	**Mean Coefficient (95% CI)**	***P***	**Adjusted R^2, %**
Mean age	17	0.01 (0.0001, 0.012)	.047	0
Mean BMI	12	0.022 (−0.034, 0.078)	.402	0
Sex distribution	17	−0.0002 (−0.003, 0.0025)	.881	0
Diff RER	7	−0.122 (−13.65, 13.89)	.983	0
Astrand	17	−0.089 (−0.27, 0.096)	.321	0
Mean VO_{2max} obtained from LC	17	−0.1 (−0.211, 0.0031)	.056	0

Astrand: Astrand classification of aerobic capacity as categorical outcome (low/average).

BMI, body mass index; CI, confidence interval; diff RER, difference in respiratory exchange ratio (positive values favors leg cycle); LC, leg cycle; $mL \cdot kg^{-1} \cdot min^{-1}$, milliliters per kilogram bodyweight per minute; N, number of groups included in the analysis.

(Fig. 5) and in L/min (Fig. 6) indicates that the ratio of 0.83 is similar when comparing patients with CVD to patients with PD.

The sensitivity analysis and investigation for small study bias on mean difference in $mL \cdot kg^{-1} \cdot min^{-1}$ (Fig. 3) indicate that the smallest study reports the largest mean difference. The metatrim analysis adjusted the mean difference for small study bias and imputed 3 fictive studies with mean differences favoring the AC. Although part of an appropriate analysis rationale, this imputation seems irrelevant as neither of our results nor previous findings in healthy adults[11] did indicate that groups of participants would obtain larger mean aerobic capacity from AC. Hereby, the most important information from this sensitivity analysis is that the smallest study is influencing the pooled mean difference toward a larger estimate.

The I^2 values from the analyses on the mean difference indicate considerable heterogeneity between the study results, whereas the I^2 values from the analyses on the ratios indicate no heterogeneity. The considerable heterogeneity found can be explained by the narrow CIs seen in Figures 3 and 4 and the nonexisting heterogeneity can be explained by the overlapping CIs seen in Figures 5 and 6.[27] The considerable heterogeneity found in the analyses on the mean difference is evaluated in the univariate meta-regressions from Table 2. The results show that the mean difference decreases with higher age and increases with higher aerobic capacity. The correlation is most likely explained in the association between age and aerobic capacity first described by Astrand;[26] thus, older patients will have lower aerobic capacity and lower mean difference between AC and LC. In general, the considerable heterogeneity and diversity in the results of the meta-analyses of the mean difference are explained mostly on the level of aerobic capacity (R^2 value above 90%). The univariate meta-regressions with negative adjusted R^2 values, presented in Table 2, describe the mean difference poorly and thus, BMI, sex distribution, and difference in RER values cannot be not be used to explain the heterogeneity in the results. A significant positive association between the ratio of the results in L/min and mean age were found in Table 3 ($P = .047$); hence, older patients will have AC values closer to LC values, compared with younger patients. The results from Table 3 are affected by the wide CIs found in the analyses on the ratios. As no heterogeneity was found in Figures 5 and 6, no R^2 could be calculated in Table 3. However, the explanation of the association between the ratio and age is probably also found in the previously described correlation between aerobic capacity and age.[26]

The included study of MacDonald et al[28] included borderline hypertensive patients and reported a higher physical fitness than the other studies. It could be argued that the included patients are only borderline CVD patients.[28] Four sensitivity analyses without the study were performed and in general the overall mean difference decreased and the overall ratio decreased. The results of the mean difference without MacDonald et al were 2.65 $mL \cdot kg^{-1} \cdot min^{-1}$ (95% CI: 1.62, 3.68) and 0.19 L/min, (95% CI: 0.15, 0.23), which were smaller than the mean differences from the main analyses. The ratio of the results measured in $mL \cdot kg^{-1} \cdot min^{-1}$ without MacDonald et al were 0.84 (95% CI: 0.78, 0.90) and 0.85 of the results in L/min, (95% CI: 0.79, 0.90), which were larger than the ones from the main analyses on the ratio. However, this does not change our conclusion, but highlights that groups of patients with higher physical fitness will have a greater difference between values from AC and LC. On the other side, the CIs on the ratio reported from each individual study in Figures 5 and 6 show that some individual patients with either PD or CVD will achieve a higher VO_2 value on AC, compared with the corresponding value on LC. The values reported in this article are on group level and it is expected that some individual patients will differ from the mean results.

Our results have some limitations as the main type of included studies was observational (prospective and retrospective) (according to Table 1). Such study designs lack any form of preplanned control to account for variables that may affect the results. None of the included articles described a random selection of patients and the possibility of selection bias is present. The risk of bias assessment also highlights the poor description of the target population, justification of sample size, and description of the statistical methods in the included studies. Especially, the study published by Martin et al[14] was assessed with high or unclear risk of bias in all of the items except the research question. According to the meta-analysis on the mean difference, the study by Martin et al[14] also reports the second highest mean difference in all of the included articles. However, except the above-mentioned study, we were not able to conclude whether risk of bias systematically affected our results toward an overestimation or underestimation of the mean difference or ratio. Another limitation is low generalizability to female patients because most of the included articles were performed on males or on mixed populations. The study by Carter et al[12] was the only study with a group of only female patients. However, the univariate meta-regressions on sex distribution do not indicate any affect by age on the mean difference nor on the ratio. Seven studies (with 7 groups) reported RER values.[14,29-34] Age-specific cuts for RER values are often used to determine whether the VO_2 outcome can be categorized as a maximum value.[35] Two of the 7 groups did not meet the age-specific criteria for maximal testing[32,33] and 4 of the 7 groups did meet the age-specific criteria.[14,29-31,34] However, it did not seem to affect the results of the mean difference nor the ratio whether the RER criteria was met or not, but it should be stated that the VO_2 values from the studies of Owens et al and Franssen et al cannot be categorized as maximum values.[32,33,35]

CLINICAL IMPLICATIONS

Implication of AC in testing of patients with CVD is high and it is already commonly used.[1] Arm cycling is often used in patients with coronary artery disease due to exercise-induced claudication in LC testing.[1] Our results show that patients with CVD as well as patients with PD

will obtain a more similar measure of physical fitness on the 2 exercise tests compared with healthy adults.[11] It is unknown whether the larger ratio between AC and LC among patients compared with healthy adults[11] is caused by patients being unable to reach the maximal aerobic capacity on LC, due to atrophy of the legs after inactivity or central limitations, or that the exercise-induced pain and discomfort is less on AC compared with LC. The results presented in this article support the implication of using AC in patients with CVD. Although the primary method of using CPET in patients with COPD is on the treadmill or with LC,[2] our results support using AC as an alternative for CPET. Whether these results can be used on other patient populations is unknown. However, it could be hypothesized that patients with central limitations such as CVD and PD will appear with the same results. New research should focus on mean difference and ratio and how the relationships between AC and LC values are affected by disease severity and also what exercise testing modality individual patients prefer.

The pooled ratio of 0.83 is larger than the ratio of 0.7 previously reported for healthy adults,[11] suggesting that differences between AC and LC are smaller in patient populations. The ratio between AC and LC is only significantly associated with age and thus, the ratio can be used in a clinical setting throughout an exercise program where patients are expected to enhance aerobic capacity. Furthermore, the pooled mean difference of 3.5 $mL \cdot kg^{-1} \cdot min^{-1}$ is smaller than the mean difference on 12.5 $mL \cdot kg^{-1} \cdot min^{-1}$ previously reported for healthy adults.[11] In conclusion, the results in this article show that patients with a low physical fitness will obtain more similar values on AC versus LC compared with patients with a higher physical fitness or healthy adults.[11] When the physical fitness and aerobic capacity is low among patients, the mean difference between test values from the AC and from the LC will also be low and the ratio between the tests will be large.

REFERENCES

1. Fletcher GF, Ades PA, Kligfield P, et al. Exercise standards for testing and training: A scientific statement from the American Heart Association. *Circulation.* 2013;128(8):873-934.
2. Gloeckl R, Marinov B, Pitta F. Practical recommendations for exercise training in patients with COPD. *Eur Respir Rev.* 2013;22(128): 178-186.
3. Haykowsky MJ, Daniel KM, Bhella PS, Sarma S, Kitzman DW. Heart failure: Exercise-based cardiac Rehabilitation: Who, when, and how intense? *Can J Cardiol.* 2016;32(10S2):S382-S387.
4. Keteyian SJ, Piña IL, Hibner BA, Fleg JL. Clinical role of exercise training in the management of patients with chronic heart failure. *J Cardiopulm Rehabil Prev.* 2010;30(2):67-76.
5. Mezzani A, Hamm LF, Jones AM, et al. Aerobic exercise intensity assessment and prescription in cardiac rehabilitation: A joint position statement of the European Association for Cardiovascular Prevention and Rehabilitation, the American Association of Cardiovascular and Pulmonary Rehabilitation and the Canadian Association of Cardiac Rehabilitation. *Eur J Prev Cardiol.* 2013;20(3):442-467.
6. Puente-Maestu L, Palange P, Casaburi R, et al. Use of exercise testing in the evaluation of interventional efficacy: An official ERS statement. *Eur Respir J.* 2016;47(2):429-460.
7. Albouaini K, Egred M, Alahmar A, Wright DJ. Cardiopulmonary exercise testing and its application. *Postgrad Med J.* 2007;83(985):675-682.
8. Paap D, Takken T. Reference values for cardiopulmonary exercise testing in healthy adults: A systematic review. *Expert Rev Cardiovasc Ther.* 2014;12(12):1439-1453.
9. Loughney L, West M, Pintus S, et al. Comparison of oxygen uptake during arm or leg cardiopulmonary exercise testing in vascular surgery patients and control subjects. *Br J Anaesth.* 2014;112(1):57-65.
10. Barreiro E, Gea J. Respiratory and limb muscle dysfunction in COPD. *COPD.* 2015;12(4):413-426.
11. Larsen RT, Christensen J, Tang LH, et al. A systematic review and meta-analysis comparing cardiopulmonary exercise test values obtained from the arm cycle and the leg cycle respectively in healthy adults. *Int J Sports Phys Ther.* 2016;11(7):1006-1039.
12. Carter R, Holiday DB, Stocks J, Tiep B. Peak physiologic responses to arm and leg ergometry in male and female patients with airflow obstruction. *Chest.* 2003;124(2):511-518.
13. Keteyian SJ, Marks CR, Levine AB, Kataoka T, Fedel F, Levine TB. Cardiovascular responses to submaximal arm and leg exercise in cardiac transplant patients. *Med Sci Sports Exerc.* 1994;26(4):420-424.
14. Martin TW, Zeballos RJ, Weisman IM. Use of arm crank exercise in the detection of abnormal pulmonary gas exchange in patients at low altitude. *Chest.* 1992;102(1):169-175.
15. Larsen RT, Keller C, Tang LH, et al. Protocol: Difference in aerobic capacity between a cardiopulmonary exercise test performed by arm cycling or leg cycling in patients with cardiovascular disease or chronic respiratory disease: A systematic review and meta-analysis. *PROSPERO.* 2016. AQ:16
16. Moher D, Liberati A, Tetzlaff J, Altman DG; The PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Med.* 2009;6(7):e1000097.
17. WHO. *WHO\|Cardiovascular Diseases (CVDs).* WHO; 2016. AQ:17
18. WHO. *WHO\|Chronic Respiratory Diseases (CRDs).* WHO; 2016.
19. NHLBI. *Quality Assessment of Case-control Studies.* NHLBI, NIH; 2011.
20. Zwierska I, Walker RD, Choksy SA, Male JS, Pockley AG, Saxton JM. Upper- vs lower-limb aerobic exercise rehabilitation in patients with symptomatic peripheral arterial disease: A randomized controlled trial. *J Vasc Surg.* 2005;42(6):1122-1130.
21. Higgins J, Green S. *16.4.6.1 Mean Differences. Version 5.1.0. Cochrane Collaboration.* 2011. AQ:18
22. Friedrich JO, Adhikari NKJ, Beyene J. Ratio of means for analyzing continuous outcomes in meta-analysis performed as well as mean difference methods. *J Clin Epidemiol.* 2011;64(5):556-564.
23. Higgins J, Green S. *9.5.2 Identifying and Measuring Heterogeneity. Version 5.1.0. Cochrane Collaboration.* 2011.
24. Higgins JPT, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ.* 2003;327(7414):557-560.
25. Higgins J, Green AR. *16.1.3.1 Imputing Standard Deviations. Version 5.1.0.* Cochrane Collaboration. AQ:19
26. Astrand I. Aerobic work capacity in men and women with special reference to age. *Acta Physiol Scand Suppl.* 1960;49(169):1-92.
27. Higgins J, Green S. *9.5 Heterogeneity—Cochrane Handbook for Systematic Reviews of Interventions.* 2011.
28. MacDonald JR, MacDougall JD, Hogben CD. The effects of exercising muscle mass on post exercise hypotension. *J Hum Hypertens.* 2000;14 (5):317-320.
29. Keteyian SJ, Marks CR, Brawner CA, Levine AB, Kataoka T, Levine TB. Responses to arm exercise in patients with compensated heart failure. *J Cardpulm Rehabil.* 1996;16(6):366-371.
30. Levandoski S, Sheldahl L, Wilke N, Tristani F, Hoffmann M. Cardiorespiratory responses of coronary artery disease Patients to arm and leg cycle ergometry. *J Cardiopul Rehabil.* 1990;10:39-44. AQ:20
31. Hansen D, Dendale P, Berger J. Type of exercise alters VE/VCO_2 slope in CAD patients. *Int J Sports Med.* 2009;30(4):302-306.
32. Owens GR, Thompson FE, Sciurba FC, Robertson R, Metz KF, Volmer RR. Comparison of arm and leg ergometry in patients with

moderate chronic obstructive lung disease. *Thorax*. 1988;43(11): 911-915.
33. Franssen FME, Wouters EFM, Baarends EM, Akkermans MA, Schols AMWJ. Arm mechanical efficiency and arm exercise capacity are relatively preserved in chronic obstructive pulmonary disease. *Med Sci Sports Exerc*. 2002;34(10):1570-1576.
34. Castagna O, Boussuges A, Vallier JM, Prefaut C, Brisswalter J. Is impairment similar between arm and leg cranking exercise in COPD patients? *Respir Med*. 2007;101(3):547-553.
35. 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.