




Reliability of maximum oxygen uptake in cardiopulmonary exercise testing with continuous laryngoscopy

Mette Engan ^{1,2,6}, Ida Jansrud Hammer^{1,6}, Marianne Bekken², Thomas Halvorsen ^{1,2,3}, Zoe Louise Fretheim-Kelly^{2,4}, Maria Vollsæter^{1,2}, Lars Peder Vatshelle Bovim⁵, Ola Drange Røksund^{1,5} and Hege Clemm^{1,2}

Affiliations: ¹Dept of Pediatric and Adolescent Medicine, Haukeland University Hospital, Bergen, Norway. ²Dept of Clinical Science, University of Bergen, Bergen, Norway. ³Dept of Sports Medicine, Norwegian School of Sport Sciences, Oslo, Norway. ⁴Faculty of Veterinary Medicine, Norwegian University of Life Sciences, Oslo, Norway. ⁵Faculty of Health and Social Sciences, Western Norway University of Applied Sciences, Bergen, Norway. ⁶These authors contributed equally.

Correspondence: Mette Engan, Dept of Pediatric and Adolescent Medicine, Haukeland University Hospital, Jonas Lies vei 65, N-5021 Bergen, Norway. E-mail: metteengan@hotmail.com

ABSTRACT

Aims: A cardiopulmonary exercise test (CPET) is the gold standard to evaluate symptom-limiting exercise intolerance, while continuous laryngoscopy performed during exercise (CLE) is required to diagnose exercise-induced laryngeal obstruction. Combining CPET with CLE would save time and resources; however, the CPET data may be distorted by the extra equipment. We therefore aimed to study whether CPET with CLE influences peak oxygen uptake ($V'_{O_2,peak}$) and other gas exchange parameters when compared to a regular CPET.

Methods: Forty healthy athletes without exercise-related breathing problems, 15–35 years of age, performed CPET to peak exercise with and without an added CLE set-up, in randomised order 2–4 days apart, applying an identical computerised treadmill protocol.

Results: At peak exercise, the mean difference (95% confidence interval) between CPET with and without extra CLE set-up for $V'_{O_2,peak}$, respiratory exchange ratio (RER), minute ventilation (V'_E) and heart rate (HR) was 0.2 (−0.4 to 0.8) mL·kg^{−1}·min^{−1}, 0.01 (−0.007 to 0.027) units, 2.6 (−1.3 to 6.5) L·min^{−1} and 1.4 (−0.8 to 3.5) beats·min^{−1}, respectively. Agreement (95% limits of agreement) for $V'_{O_2,peak}$, RER and V'_E was 0.2 (±3.7) mL·kg^{−1}·min^{−1}, 0.01 (±0.10) units and 2.6 (±24.0) L·min^{−1}, respectively. No systematic or proportional bias was found except for the completed distance, which was 49 m (95% CI 16 to 82 m) longer during CPET.

Conclusion: Parameters of gas exchange, including $V'_{O_2,peak}$ and RER, obtained from a maximal CPET performed with the extra CLE set-up can be used interchangeably with data obtained from standard CPET, thus preventing unnecessary additional testing.



@ERSpublications

Cardiopulmonary exercise testing (CPET) with concurrent continuous laryngoscopy provides reliable measures for maximal oxygen consumption and other CPET data <https://bit.ly/36wVJhb>

Cite this article as: Engan M, Hammer IJ, Bekken M, *et al.* Reliability of maximum oxygen uptake in cardiopulmonary exercise testing with continuous laryngoscopy. *ERJ Open Res* 2021; 7: 00825-2020 [<https://doi.org/10.1183/23120541.00825-2020>].



Introduction

Exercise-related breathing problems are common, with causes varying from poor physical condition, overweight or de-training, to disease conditions such as respiratory or cardiac disorders [1]. Asthma and exercise-induced bronchoconstriction are important causes of exercise-related breathing problems; however, the significance of exercise-induced laryngeal obstruction (EILO) is increasingly acknowledged [2]. With a prevalence of 5–7% in unselected adolescent populations [3, 4], and even higher in groups where exercise is particularly important [5, 6], EILO is now recognised as an important cause of exertional breathlessness [3, 7, 8]. We know that misdiagnosing exercise-related breathing problems might lead to unfortunate consequences and delay institution of adequate treatment [9–12]. We also know that diagnostic evaluations based solely on patients' symptom descriptions are inaccurate, and guidelines therefore prescribe objective evaluation [13]. Despite good intentions, diagnostic errors are made, often because objective tests are not performed [5, 14–17].

A cardiopulmonary exercise test (CPET) is the gold standard for evaluating symptom-limiting exercise intolerance in patients with suspected respiratory or cardiac disease [18]. Continuous laryngoscopy exercise (CLE) testing is the gold standard for diagnosing EILO and requires flexible video laryngoscopy performed during a maximal CPET [8, 19]. Combining CPET and CLE on a treadmill was first performed by our research group, published in 2006 [20]. This combined set-up has benefits in clinical and scientific situations, enabling comprehensive physiological evaluations of gas exchange, and cardiovascular and respiratory capacities and limitations, in addition to concomitant visual assessment of the larynx [1].

Measurement of maximal oxygen consumption (V'_{O_2}) based on a standard CPET is highly reproducible [21]. The combined set-up of CPET and CLE represents more stress for the patient, adds extra weight, and can introduce air leaks, all of which might influence performance and gas exchange. MIRZA *et al.* [22] found no systematic difference between important parameters of gas exchange obtained from bicycle CPETs performed with and without continuous laryngoscopy. We aimed to expand on that knowledge, and conducted a full method comparison study using treadmill exercise, evaluating reliability and the upper and lower limits of agreement between peak oxygen consumption (V'_{O_2} peak) and other parameters of gas exchange obtained from a CPET with and without added CLE equipment.

Material and methods

Subjects and study design

In this method comparison study, a convenience sample of healthy 15–35-year-old athletes participating in endurance training for a minimum of four times per week, were recruited to perform CPET to peak exercise, with and without added CLE set-up, in a randomised order. The two tests were performed 2–4 days apart applying the same computerised treadmill protocol. Except for the modifications required for the CLE test, the equipment and the pre-test preparations were identical. The test personnel consisted of three people, providing instructions and encouragement in similar ways.

At inclusion, all participants completed a questionnaire on past medical history and habits of physical activity and exercise. Subjects who reported breathing difficulties at rest or during exercise, or who had any medical issues (except well-controlled asthma) were not enrolled. Information on the use of nicotine, food intake, hours of nightly sleep, time of previous workout, or any symptoms of illness were obtained for the 24 h preceding both tests.

Pulmonary function

Spirometry was performed on both test days, using a Vyntus PNEUMO spirometer (Vyaire Medical GmbH, Leibnizstrasse, Hoechberg, Germany) according to guidelines [23]. Forced vital capacity (FVC) and forced expiratory volume in 1 s (FEV_1) were recorded and reported as raw data and as z-scores calculated by the Global Lung Function Initiative online calculator [24].

CPET

An incremental treadmill (Woodway PPS 55 Med, Weil am Rhein, Germany) test was applied, using a pre-set modified Bruce protocol [25]. Speed and elevation were gradually increased every 60 s from an initial slow walking phase. A facemask (Hans Rudolph Inc., Kansas City, MO, USA) connected the patient to a Vyntus CPX unit powered by SentrySuite software (Vyaire Medical GmbH, Leibnizstrasse, Hoechberg, Germany). After baseline variables were established, subjects ran to exhaustion. Parameters of gas exchange and airflow were measured breath-by-breath and averaged over 30 s.

The test was considered successful when the subjects indicated exhaustion, preferably supported by a plateau in V'_{O_2} and/or the heart rate (HR) response. Duration and completed treadmill distance were recorded. V'_{O_2} , carbon dioxide output (V'_{CO_2}), tidal volume (V_T), respiratory rate (RR) and HR were measured directly, while minute ventilation (V'_E) was calculated from V_T and RR.

CLE test

The CPET part of the CLE test was performed as described above. After application of local anaesthesia (Lidocaine), a flexible fiberoptic laryngoscope (Olympus ENF-V2, Tokyo, Japan) with diameter 3.4 mm was introduced into the pharyngeal space *via* a tight-fit opening in a slightly modified Hans Rudolph facemask and through one nostril (figure 1). The mask was tested for air leaks and the laryngoscope was secured with tape and positioned for a good view of the laryngeal entrance, including supraglottic structures and the vocal folds. The body of the laryngoscope was secured to the head of the test person by custom-made headgear. During the treadmill test, the added CLE equipment allowed for videorecording of the laryngeal inlet (figure 1).

Analysis

Data were reported as means with 95% confidence intervals, ranges or SD for continuous data or counts with percentages for categorical data. Cardiopulmonary exercise data from tests without the CLE set-up were labelled “CPET”, whereas data from tests with the added CLE set-up were labelled “CLE”. A paired t-test was used to compare mean differences.

An intraclass correlation coefficient (ICC) using a two-way mixed effect model based on single ratings was calculated to assess the absolute agreement of the repeated measurements at peak exercise; $V'_{O_2\text{peak}}$, $V'_{CO_2\text{peak}}$, respiratory exchange ratio (RER), RR, HR, V_T and V'_E . Values >0.6 were considered to indicate good reliability, and values >0.9 to indicate excellent reliability [26]. To assess the reproducibility of the data, the within-subject coefficients of variation (CoV) for the parameters were calculated assuming that the SD was reasonably constant across the concentration intervals.

Agreement between the ergospirometry data produced by the CPET and CLE method was assessed by calculating the SD of the mean difference between the measurements obtained with the two techniques. The SD was used to calculate 95% limits of agreement (LoA) between the two techniques, representing ± 1.96 SD of the differences [27]. Results were expressed as values and percentages of the average of the paired differences and visualised in Bland–Altman plots for $V'_{O_2\text{peak}}$, RER and V'_E . We considered $3.5 \text{ mL}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$ to be the maximum acceptable difference within pairs of peak V'_{O_2} measurements, equivalent to one metabolic equivalent of task. One-sample t-test *versus* zero was used to examine for a systematic bias between the two methods. The differences between the two methods were regressed on their average to test for proportional bias (*i.e.* if the differences were influenced by the numerical size of the measurements). Normal distribution of the differences was verified by a Kolmogorov–Smirnov test.

The criterion for statistical significance was set at $p < 0.05$. Statistical calculations were performed using the statistical software SPSS version 25 (IBM SPSS Statistics, Armonk, NY, USA) and MedCalc version 19.5.3 (MedCalc Software Ltd., Ostend, Belgium).

We used MedCalc statistical software to determine an appropriate sample size for agreement of peak V'_{O_2} (α values set to 0.05 and power to 90%). A minimum of 31 pairs were required, providing an expected mean difference of $0.4 \text{ mL}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$, an expected SD for the mean difference of $1.0 \text{ mL}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$, and 95% LoA set to $3.5 \text{ mL}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$.

Ethics

The Regional Committee for Medical and Health Research Ethics of the Western Norway Health Authority approved of the study (REC 2014/601). Written informed consent was obtained from all participants. For participants under 16 years, parents also signed the consent.

Results

Of 47 participants, 40 (21 females) successfully completed both the CPET and the CLE test. Their characteristics are given in table 1 and the CPET results are summarised in table 2. Four of the seven incomplete datasets represent participants who were declined further participation because of an asthma exacerbation or for EILO symptoms discovered at the first test, one had difficulties running on the treadmill, one could not find time for the second test, and one participant actively declined to perform the repeat test.

Ergospirometry data obtained with or without the CLE set-up did not differ. The distance completed on the treadmill was 49 m (95% CI 16 to 82 m) longer during CPET than during the CLE test (systematic bias), but there was no proportional bias. For all other parameters, there was no systematic or proportional bias (*i.e.* the data produced by the two methods did not systematically differ from each other), nor were there differences between the measurements influenced by the size of their average.

Table 3 summarises the reliability and reproducibility of the parameters and figure 2 illustrates the agreement for $V'_{O_2\text{peak}}$, V'_E and RER in Bland–Altman plots. For $V'_{O_2\text{peak}}$ the mean difference (95%



FIGURE 1 Illustration of the set-up for cardiopulmonary exercise test (CPET) with and without continuous video laryngoscopy. The upper left and right images demonstrate the facemask used for ordinary CPET. The middle left and right images demonstrate the modified facemask with a flexible transnasal laryngoscope positioned through a tight-fit opening. A custom-made headgear secures the body of the laryngoscope. The attached transnasal flexible laryngoscope enables video recording of the laryngeal inlet during the maximal exercise treadmill test (bottom image).

TABLE 1 Characteristics of the 40 subjects comparing peak oxygen consumption measured with or without continuous laryngoscopy during cardiopulmonary exercise test on a treadmill

Female sex n (%)	21 (53)
Age years mean (range)	24.8 (15.0 to 35.0)
BMI kg·m ⁻²	22.3±2.3
Smoking n (%)	0 (0)
Exercise	
4–6 days per week n (%)	19 (47.5)
Daily n (%)	21 (52.5)
Asthma n (%)	1 (3)
FVC CPET L	5.12±1.26
FVC CLE L	5.15±1.25
FVC z-score CPET	0.25±0.81
FVC z-score CLE	0.30±0.79
FEV₁ CPET L	4.31±1.32
FEV₁ CLE L	4.20±0.92
FEV₁ z-score CPET	0.05±0.83
FEV₁ z-score CLE	0.07±0.79

Data are presented as mean±SD, unless otherwise stated. BMI: body mass index; FVC: forced vital capacity; CPET: cardiopulmonary exercise test; CLE: continuous laryngoscopy during exercise; FEV₁: forced expiratory volume in 1 s.

LoA) was 0.2 (±3.7) mL·kg⁻¹·min⁻¹, and the ICC was excellent with a low CoV (table 3 and figure 2). For V_E at peak exercise the mean difference (95% LoA) was 2.6 (±24) L·min⁻¹, the ICC was excellent, and the CoV was low. For RER, the agreement was good and the ICC was fair with a low CoV (table 3 and figure 2).

Discussion

This is the first study to establish the LoA between parameters of gas exchange obtained during maximal cardiopulmonary treadmill exercise with and without concurrent laryngoscopy, performed in physically active and healthy subjects. The study indicates that the CLE equipment does not disturb the CPET measurements. CPET and CLE tests are time- and resource-consuming for patients and healthcare providers. Combining the two saves time and ensures that advanced patient evaluation can be performed at lower personal and financial costs.

Our findings are consistent with the only similar study comparing CPET with and without CLE; however, the other study by MIRZA *et al.* [22] was performed in sedentary or moderately physically active individuals with EILO, using a cycle ergometer. They found a mean (95% CI) difference in V_{O₂peak} of 22 (–125 to 81) mL·min⁻¹, compared to our results of 11 (–33 to 56) mL·min⁻¹. MIRZA *et al.* [22] did not report the degree of expected agreement between the gas exchange parameters obtained with the two methods.

TABLE 2 Ergospirometry data for the 40 participants obtained from cardiopulmonary exercise tests performed without and with added CLE test set-up

Parameter	CPET without CLE		CPET with CLE		Difference	
	Mean	SD	Mean	SD	Mean	(95% CI)
V _{O₂peak} mL·min ⁻¹	3818	873.8	3806	883.4	11.4	(–33.5, 56.2)
V _{O₂peak} mL·kg ⁻¹ ·min ⁻¹	54.9	6.3	54.7	6.6	0.22	(–0.40, 0.83)
V _{CO₂peak} mL·min ⁻¹	4618	1016	4565	1006	53.1	(–25.9, 132.0)
V _T L	2.61	0.65	2.61	0.62	0.005	(–0.08, 0.071)
V _E L·min ⁻¹	142.3	33.1	139.7	34.2	2.58	(–1.33, 6.49)
HR beats·min ⁻¹	186	8.7	185	9.0	1.4	(–0.8, 3.5)
RR breaths·min ⁻¹	55	9.3	54	10.0	1.2	(–1.5, 4.0)
RER units	1.21	0.05	1.20	0.05	0.01	(–0.007, 0.027)
Distance m	1199	219	1150	210	49	(15.7, 82.0)

CLE: continuous laryngoscopy during exercise; CPET: cardiopulmonary exercise test; V_{O₂peak}: oxygen consumption at peak exercise; V_{CO₂peak}: carbon dioxide production at peak exercise; V_T: tidal volume; V_E: minute ventilation; HR: heart rate; RR: respiratory rate; RER: respiratory exchange rate.

TABLE 3 Reliability and reproducibility of the ergospirometry data obtained for the 40 participants performing cardiopulmonary exercise test with and without CLE test set-up

Parameter	Agreement	Lower LoA	Upper LoA	ICC	[95% CI]	WS-sd	CoV %
$V_{O_2\text{peak}}$ mL·min ⁻¹	11.4	-263	286	0.988	(0.977, 0.993)	98	2.6
$V_{O_2\text{peak}}$ mL·kg ⁻¹ ·min ⁻¹	0.2	-3.5	4.0	0.957	(0.920, 0.977)	1.3	2.5
$V_{CO_2\text{peak}}$ mL·min ⁻¹	53.1	-431	537	0.970	(0.943, 0.984)	176	3.8
V_T L	0.0	-0.47	0.46	0.933	(0.876, 0.964)	0.16	6.3
V_E L·min ⁻¹	2.6	-21.4	26.5	0.933	(0.877, 0.964)	8.7	6.2
HR beats·min ⁻¹	1.4	-11.6	14.3	0.720	(0.532, 0.841)	4.7	2.5
RR breaths·min ⁻¹	-2.5	-19.9	15.0	0.614	(0.375, 0.777)	6.0	11.1
RER units	0.01	-0.09	0.11	0.502	(0.233, 0.700)	0.04	3.1

CLE: continuous laryngoscopy during exercise; LoA: limit of agreement; ICC: intraclass correlation coefficient; WS-sd: within-subject sd; CoV: coefficient of variation; $V_{O_2\text{peak}}$: oxygen consumption at peak exercise; $V_{CO_2\text{peak}}$: carbon dioxide production at peak exercise; V_T : tidal volume; V_E : minute ventilation; HR: heart rate; RR: respiratory rate; RER: respiratory exchange rate.

We opted to apply a treadmill protocol, aiming to achieve the highest possible effort of the subjects. Treadmill exercise has proven feasible in subjects not familiar with exercise testing in general [28] and it is our experience that children and adolescents more often exercise to exhaustion on a treadmill than on a cycle ergometer. Moreover, running involves larger muscle masses and leads to a greater stress on the organ systems mediating the exercise response, generally leading to 5–10% higher maximal oxygen consumption [29, 30]. This might be of importance in athletes, and in some patients where abnormalities may occur only at high metabolic demands, which could be the case for patients with symptoms resembling EILO [29].

$V_{O_2\text{peak}}$ is the most important CPET parameter for determining cardiorespiratory fitness and it is commonly used as a primary end-point for studies to determine the effect of therapeutic interventions or training programmes [31]. DECATO *et al.* [32] provide an overview of studies reporting repeatability of CPET variables both for normal subjects and chronically ill patients. For $V_{O_2\text{peak}}$, the CoV ranges from 3.0% to 13% [33, 34] and large studies using highly consistent exercise test methodology found CoVs of 4.9 to 5.1% [32, 35]. Heterogeneities between studies represent technical errors, technicians with varying experiences, and different exercise protocols in addition to biological variation [36]. In our study there was no systematic bias favouring any of the two methods, and the CoV of $V_{O_2\text{peak}}$ measurements was only 2.5% (*i.e.* a value below what is reported for repeat CPETs). This steadiness may be due to our test

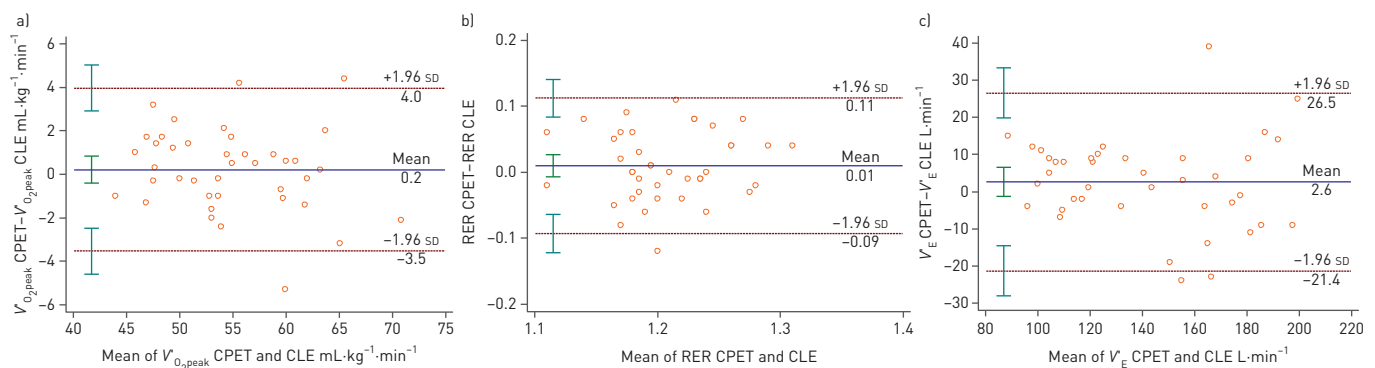


FIGURE 2 Agreement between peak oxygen consumption ($V_{O_2\text{peak}}$), respiratory exchange ratio (RER) and minute ventilation (V_E) obtained from cardiopulmonary exercise testing (CPET) and continuous laryngoscopy performed during exercise (CLE) test. The horizontal lines depict the mean difference between the variables obtained with the two methods, whereas ± 1.96 SD of these differences represent 95% limits of agreement between the two methods. The 95% confidence intervals for the mean, the upper limit of agreement and the lower limit of agreement are indicated by vertical lines. The figures given below in the brackets are the mean difference and the upper and lower limits of agreement, expressed as percentages of the mean of the CPET and CLE value. a) Agreement between $V_{O_2\text{peak}}$, expressed as mL·kg⁻¹·min⁻¹, obtained from CPET and CLE tests. The mean difference was 0.2 mL·kg⁻¹·min⁻¹ (0.4%), the upper limit of agreement was 4.0 mL·kg⁻¹·min⁻¹ (7.0%), and the lower limit of agreement was -3.5 mL·kg⁻¹·min⁻¹ (-6.2%). b) Agreement between RER obtained from CPET and CLE test at peak exercise. The mean difference was 0.01 units (0.7%), the upper limit of agreement was 0.11 units (9.2%), and the lower limit of agreement was -0.09 units (-7.8%). c) Agreement between V_E obtained from CPET and CLE test at peak exercise. The mean difference was 2.58 L·min⁻¹ (2.2%), the upper limit of agreement was 26.5 L·min⁻¹ (18.3%), and the lower limit of agreement was -21.4 L·min⁻¹ (-13.9%).

population, consisting of athletes who conceivably are more used to repeatedly performing to their peak exercise capacity than average sedentary individuals.

The CoV of the parameters V'_{E} , RER, and HR in the present study were generally consistent with or lower than previous reports from repeatability studies of CPET without CLE [32]. The ICCs for V'_{E} and HR were excellent and good, respectively. The ICC for RER was not as good as the other parameters, which may be explained by the diurnal variation of this parameter as the participants were not scheduled for testing at the same hours of the day [32]. The mean completed distance was 49 m longer for the test without the CLE set-up, yet achieving the same $V'_{O_2,peak}$. This result could be explained by the added weight and higher work of breathing with one nostril occlude by the laryngoscope, although this assumption is uncertain, and we cannot provide a good explanation. Regarding peak V'_{E} , LoA was relatively wide and in the area of 15% of its mean value, which is in line with studies testing repeatability of CPET [32]. Peak V'_{E} is included in the calculation of a patient's breathing reserve, which is used to identify ventilatory limitation to exercise. Based on the result from this study, caution should be exercised if a ventilatory limitation to exercise is established solely on the interpretation of a reduced breathing reserve.

CPET is considered the gold standard for evaluating symptom-limited exercise tolerance in patients with pulmonary and cardiac disease [18]. Although response patterns obtained from CPET are not necessarily disease-specific, they nevertheless contribute to the diagnostic process of narrowing down the spectrum of differential diagnoses. Normal CPET response patterns are usually taken as evidence against significant involvement of these systems in exercise limitation. Combining CPET with the CLE test is valuable in several ways. The most obvious advantage is that exercise-induced changes of the laryngeal configuration can be evaluated against corresponding changes of gas exchange parameters. This can be of value in the understanding of individual patients' limitations, as well as facilitate a better understanding of the phenomenon of EILO in research, such as providing outcome measures in intervention studies that aim to improve EILO. Moreover, CPET data will identify if patients reach their maximum level of effort and can serve to identify abnormalities in the oxygen transport pathways, irrespectively of obstructions in the larynx. Lastly, if no signs of abnormalities are revealed, the CPET data inform that the patient's dyspnoea is not likely to be associated with an abnormal physiological limitation, but instead consistent with the actual physical fitness of the tested individual.

Finding a normal $V'_{O_2,peak}$ may be helpful in providing reassurance to patients and prevent unnecessary subsequent testing and evaluations. However, it is important to bear in mind that CPET results should be interpreted based on expected values for the population under investigation [29, 37]. Therefore, it is important to ensure that the gas exchange parameters obtained from a CLE test are correct also in physically active individuals, as shown by this study.

Strength and limitations

The major strengths of this study were the randomised crossover design and the relatively high number of participants performing endurance training at a high level. Measuring $V'_{O_2,peak}$ may be influenced by a learning effect, and performing the tests in randomised order controlled for this possible bias [29].

All our participants were highly physically active performing endurance training at least four times per week with no reports of breathing problems. Their familiarity with high-level training and performance likely reduced the variation between the two tests. Moreover, other studies testing reproducibility of $V'_{O_2,peak}$ have used shorter test protocols and longer "between-test" periods of up to 30 days compared to our maximum of 2 weeks [32].

We did not systematically collect information on previous experiences with exercise testing, flexible laryngoscopy, or CLE testing prior to inclusion. This is a limitation to the study as familiarity with the test situation can influence the results. In the records from our test laboratory, we found that prior to the study, six subjects had performed a CPET test and none had performed a CLE test. The distance to the next cardiopulmonary laboratory capable of performing a CLE test is 600 km. Thus, we find it unlikely that any of these healthy participants had performed a CLE test prior to this study.

Testing only healthy athletic subjects without respiratory complaints, challenges the validity for patients who undergo cardiopulmonary exercise testing for reasons other than EILO. On the other hand, this limitation represents a strength in the context of testing patients with EILO, as EILO is common in athletes and physically active individuals [5]. The test personnel were not technically prevented from access to previous results, and access to such knowledge could in theory have influenced their level of encouragement during the second test. However, the test personnel were skilled and experienced researchers who did in fact not access the previous test results. We did not collect information on how the added laryngoscope influenced perceived dyspnoea or discomfort during test, and we did not use a Borg score or other dyspnoea or discomfort scores. Ultimately, dyspnoea or breathing difficulties during exercise

are the main reason for performing a CPET with CLE, and future studies should explore the added discomfort imposed by the extra necessary gear.

The 95% LoA between V'_{O_2} peak obtained from CPET with and without CLE was $0.2 \pm 3.7 \text{ mL}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$, which was slightly higher than the $3.5 \text{ mL}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$ that we had considered clinically relevant. This was due to the SD of the mean difference being larger than expected, which could have been moderated by including more participants. However, inclusion was complicated by active young people with tight schedules, a busy exercise test laboratory, and a test protocol that prescribed replicate testing within 2 weeks.

In conclusion, V'_{O_2} peak and most of the other ergospirometry data obtained from a CLE test can be used interchangeably with data obtained from a standard CPET in athletes and highly physically trained individuals. Thus, we recommend performing CPET with CLE set-up, if available, when assessing patients with unexplained exertional breathing problems.

Author contributions: M. Engan organised data, carried out the analyses, drafted the initial manuscript and revised the manuscript. I.J. Hammer coordinated and performed data collection and organised data, carried out the initial analyses, drafted the initial manuscript and revised the manuscript. M. Bekken organised data, carried out the initial analyses, and revised the manuscript. Z.L. Fretheim-Kelly revised the analyses and revised the manuscript. T. Halvorsen and M. Vollsæter conceptualised and designed the study, designed the data collection instruments, reviewed, and revised the manuscript. L.P.V. Bovim performed data collection, reviewed, and revised the manuscript. O.D. Røksund conceptualised and designed the study, designed the data collection instruments, performed data collection, reviewed, and revised the manuscript. H. Clemm conceptualised and designed the study, designed the data collection instruments, coordinated and supervised data collection, organised data, carried out the initial analyses, drafted the initial manuscript and revised the manuscript. All authors approved the final manuscript as submitted and agree to be accountable for all aspects of the work.

Data availability: In accordance with the approval granted for this study by The Regional Committee on Medical Research Ethics and The Norwegian Data Inspectorate, the data files are stored securely and in accordance with the Norwegian Law of Privacy Protection. A subset of the data file with anonymised data can be made available to interested researchers upon reasonable request to H. Clemm, providing that Norwegian privacy legislation and the General Data Protection Regulation are respected, and that permission is granted from The Norwegian Data Inspectorate and the data protection officer at Haukeland University Hospital.

Conflict of interest: None declared.

Support statement: This study was supported by the Western Norway Health Authority. Funding information for this article has been deposited with the Crossref Funder Registry.

References

- 1 Abu-Hasan M, Tannous B, Weinberger M. Exercise-induced dyspnea in children and adolescents: if not asthma then what? *Ann Allergy Asthma Immunol* 2005; 94: 366–371.
- 2 Christopher KL, Wood RP 2nd, Eckert RC, et al. Vocal-cord dysfunction presenting as asthma. *N Engl J Med* 1983; 308: 1566–1570.
- 3 Johansson H, Norlander K, Berglund L, et al. Prevalence of exercise-induced bronchoconstriction and exercise-induced laryngeal obstruction in a general adolescent population. *Thorax* 2015; 70: 57–63.
- 4 Christensen PM, Thomsen SF, Rasmussen N, et al. Exercise-induced laryngeal obstructions: prevalence and symptoms in the general public. *Eur Arch Otorhinolaryngol* 2011; 268: 1313–1319.
- 5 Nielsen EW, Hull JH, Backer V. High prevalence of exercise-induced laryngeal obstruction in athletes. *Med Sci Sports Exerc* 2013; 45: 2030–2035.
- 6 Morris MJ, Deal LE, Bean DR, et al. Vocal cord dysfunction in patients with exertional dyspnea. *Chest* 1999; 116: 1676–1682.
- 7 Hull JH. Not all wheeze is asthma: time for patients to exercise their rights. *Thorax* 2015; 70: 7–8.
- 8 Halvorsen T, Walsted ES, Bucca C, et al. Inducible laryngeal obstruction: an official joint European Respiratory Society and European Laryngological Society statement. *Eur Respir J* 2017; 50: 1602221.
- 9 Buchvald F, Phillipsen LD, Hjuler T, et al. Exercise-induced inspiratory symptoms in school children. *Pediatr Pulmonol* 2016; 51: 1200–1205.
- 10 Hull JH, Walsted ES, Pavitt MJ, et al. High prevalence of laryngeal obstruction during exercise in severe asthma. *Am J Respir Crit Care Med* 2019; 199: 538–542.
- 11 Røksund OD, Maat RC, Heimdal JH, et al. Exercise induced dyspnea in the young. Larynx as the bottleneck of the airways. *Respir Med* 2009; 103: 1911–1918.
- 12 Hull JH, Godbout K, Boulet LP. Exercise-associated dyspnea and stridor: thinking beyond asthma. *J Allergy Clin Immunol Pract* 2020; 8: 2202–2208.
- 13 Parsons JP, Hallstrand TS, Mastrorarde JG, et al. American Thoracic Society Subcommittee on Exercise-induced B. An official American Thoracic Society clinical practice guideline: exercise-induced bronchoconstriction. *Am J Respir Crit Care Med* 2013; 187: 1016–1027.
- 14 Ansley L, Kippelen P, Dickinson J, et al. Misdiagnosis of exercise-induced bronchoconstriction in professional soccer players. *Allergy* 2012; 67: 390–395.
- 15 Hull JH, Hull PJ, Parsons JP, et al. Approach to the diagnosis and management of suspected exercise-induced bronchoconstriction by primary care physicians. *BMC Pulm Med* 2009; 9: 29.

- 16 Bush A, Fleming L. Is asthma overdiagnosed? *Arch Dis Child* 2016; 101: 688–689.
- 17 Looijmans-van den Akker I, van Luijn K, Verheij T. Overdiagnosis of asthma in children in primary care: a retrospective analysis. *Br J Gen Pract* 2016; 66: e152–e157.
- 18 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: 429–460.
- 19 Christensen PM, Heimdahl JH, Christopher KL, *et al.* ERS/ELS/ACCP 2013 international consensus conference nomenclature on inducible laryngeal obstructions. *Eur Respir Rev* 2015; 24: 445–450.
- 20 Heimdahl JH, Roksund OD, Halvorsen T, *et al.* Continuous laryngoscopy exercise test: a method for visualizing laryngeal dysfunction during exercise. *Laryngoscope* 2006; 116: 52–57.
- 21 Force ERST, Palange P, Ward SA, *et al.* Recommendations on the use of exercise testing in clinical practice. *Eur Respir J* 2007; 29: 185–209.
- 22 Mirza KK, Walsted ES, Backer V. Ergospirometry with concurrent fibre optic laryngoscopy: a randomised crossover study. *Eur Clin Respir J* 2017; 4: 1399033.
- 23 Miller MR, Hankinson J, Brusasco V, *et al.* Standardisation of spirometry. *Eur Respir J* 2005; 26: 319–338.
- 24 Quanjer PH, Stanojevic S, Cole TJ, *et al.* Multi-ethnic reference values for spirometry for the 3–95-yr age range: the global lung function 2012 equations. *Eur Respir J* 2012; 40: 1324–1343.
- 25 Cumming GR, Everatt D, Hastman L. Bruce treadmill test in children: normal values in a clinic population. *Am J Cardiol* 1978; 41: 69–75.
- 26 Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med* 2016; 15: 155–163.
- 27 Altman DG, Bland JM. Measurement in medicine: the analysis of method comparison studies. *J R Stat Soc Ser D* 1983; 32: 307–317.
- 28 Clemm H, Røksund O, Thorsen E, *et al.* Aerobic capacity and exercise performance in young people born extremely preterm. *Paediatrics* 2012; 129: e97–e105.
- 29 American Thoracic Society, American College of Chest Physicians. ATS/ACCP statement on cardiopulmonary exercise testing. *Am J Respir Crit Care Med* 2003; 167: 211–277.
- 30 Hermansen L, Saltin B. Oxygen uptake during maximal treadmill and bicycle exercise. *J Appl Physiol* 1969; 26: 31–37.
- 31 Gibbons RJ, Balady GJ, Bricker JT, *et al.* ACC/AHA 2002 guideline update for exercise testing: summary article. A report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines (Committee to Update the 1997 Exercise Testing Guidelines). *J Am Coll Cardiol* 2002; 40: 1531–1540.
- 32 Decato TW, Bradley SM, Wilson EL, *et al.* Repeatability and meaningful change of CPET parameters in healthy subjects. *Med Sci Sports Exerc* 2018; 50: 589–595.
- 33 Cox NJ, Hendriks JC, Binkhorst RA, *et al.* Reproducibility of incremental maximal cycle ergometer tests in patients with mild to moderate obstructive lung diseases. *Lung* 1989; 167: 129–133.
- 34 Bingisser R, Kaplan V, Scherer T, *et al.* Effect of training on repeatability of cardiopulmonary exercise performance in normal men and women. *Med Sci Sports Exerc* 1997; 29: 1499–1504.
- 35 Skinner JS, Wilmore KM, Jaskolska A, *et al.* Reproducibility of maximal exercise test data in the HERITAGE family study. *Med Sci Sports Exerc* 1999; 31: 1623–1628.
- 36 Katch VL, Sady SS, Freedson P. Biological variability in maximum aerobic power. *Med Sci Sports Exerc* 1982; 14: 21–25.
- 37 Sydo N, Csulak E, Major D, *et al.* P3105 cardiopulmonary exercise testing normal values for young athletes and non-athletes. *Eur Heart J* 2019; 40: Suppl. 1, ehz745.0180.