Lung (2019) 197:483–492
https://doi.org/10.1007/s00408-019-00233-4

ASTHMA

A Meta-analysis of Diagnostic Test Agreement Between Eucapnic Voluntary Hyperventilation and Cardiopulmonary Exercise Tests for Exercise-Induced Bronchoconstriction

Imran H. Iftikhar[1] · Meredith Greer[1] · Ahmadu Jaiteh[1]

Received: 2 January 2019 / Accepted: 30 April 2019 / Published online: 10 May 2019

Abstract

Introduction Exercise-induced bronchoconstriction (EIB) is very common in athletes. Cardiopulmonary exercise tests (CPET) have traditionally been used for the diagnosis of EIB. However, alternative indirect bronchoprovocation tests have recently been used as surrogate tests. One of these is the eucapnic voluntary hyperventilation (EVH). This meta-analysis studied the agreement between the two tests.

Methods An extensive search in PubMed and Medline was conducted for studies where participants underwent both CPET and EVH with measurement of forced expiratory volume in 1-second (FEV_1). After extracting data using two-by-two contingency tables, pooled positive and negative agreements were first calculated between the two tests, with EVH benchmarked against CPET, and then, pooled positive and negative agreements were calculated with CPET benchmarked against EVH.

Results The pooled positive and negative agreements between EVH and CPET (with CPET as the reference) were 0.62 [(95% confidence interval 0.54–0.70), I^2 77%] and 0.61 [(0.56–0.65)), I^2 81%]. The pooled positive and negative agreements between CPET and EVH (with EVH as the reference) were 0.36 [(0.30–0.42), I^2 93%] and 0.82 [(0.77–0.86), I^2 78%]. The average of positive test results with EVH across all studies was greater than that of CPETs (58.84% vs. 39.51%).

Conclusions Results of this meta-analysis show poor positive agreement between the two tests but high negative agreement (specifically using EVH as reference), suggesting that either test can be used for correctly identifying those without EIB. Results also suggest that the chances of a test resulting positive are higher with EVH than with CPET.

Keywords Exercise-induced bronchoconstriction · Cardiopulmonary exercise test · Eucapnic voluntary hyperventilation

Introduction

Exercise-induced bronchoconstriction (EIB) is the phenomenon of transient reversible narrowing of the airways that can occur during exercise [1]. This condition is highly prevalent in athletes [2]. The prevalence of EIB is estimated to be 6–20% in the general pediatric population and 50–90% in children and adolescents with asthma in whom this condition can be associated with decreased physical activity, impaired psychomotor development and quality of life [3, 4]. Respiratory complaints, such as dyspnea, tightness of the chest, and coughing and wheezing, are frequently reported by people with asthma after vigorous exercise but are neither sensitive nor specific for the diagnosis of EIB [3, 4]. It is generally known that there is a poor relationship between the presence of 'asthma-type' symptoms and objective evidence of EIB in athletes [5–7] as resting spirometric values are poorly predictive of EIB in athletes [8]. Therefore, for the diagnosis of EIB, it is important to perform objective testing to confirm any reversible change in airway function.

Cardiopulmonary exercise tests (CPET) that involve running on treadmills or pedaling on a stationary bike are exercise challenges that have been used for EIB diagnosis. Although these tests are well standardized and widely used for EIB diagnosis in adults and children older than 8 years, other challenge methods have been suggested as

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00408-019-00233-4) contains supplementary material, which is available to authorized users.

✉ Imran H. Iftikhar
imran.hasan.iftikhar@emory.edu

[1] Division of Pulmonary, Allergy, Critical Care & Sleep Medicine, Emory University School of Medicine, 613 Michael St, NE, Atlanta, GA, USA

surrogates [9]. One of these alternatives to CPETs that has gained attention recently is the eucapnic voluntary hyperventilation test (EVH). The EVH methodology was established in 1984 to test army recruits for EIB [10] and since then has been used in the clinical setting for the diagnosis of EIB in athletes. This test requires an athlete to complete a period of voluntary hyperpnea breathing in a dry gas mixture (usually 21% oxygen, 5% carbon dioxide, and balance nitrogen), that results in desiccation of the airways, mimicking the osmotic priming stimulus to EIB [11]. Despite the long history and widespread use of the EVH test in clinical practice, not enough data exists in the literature to support its position as the 'gold standard' in the diagnosis of EIB. This meta-analysis sought to determine the agreement between CPET and EVH in diagnosing EIB.

Methods

Search Strategy and Study Eligibility Criteria

A systematic search of studies in PubMed and Medline was performed in April 2018 and updated in March 2019. Inclusion criteria were: (a) studies reporting information on the sensitivity and specificity of EVH and CPET where every study participant had both EVH and CPET tests (b) and studies including participants aged 8 years and above. Excluded from selection were review articles, non-peer-reviewed papers, and conference proceedings. The search strategy is displayed in Fig. 1. After the initial screening of the titles from the search database, full text articles of short-listed abstracts were independently assessed by all authors (I.H.I, M.G, and A.J) for inclusion in this meta-analysis. Disagreement on any study selection was resolved by consensus. Quality of included studies was assessed with the

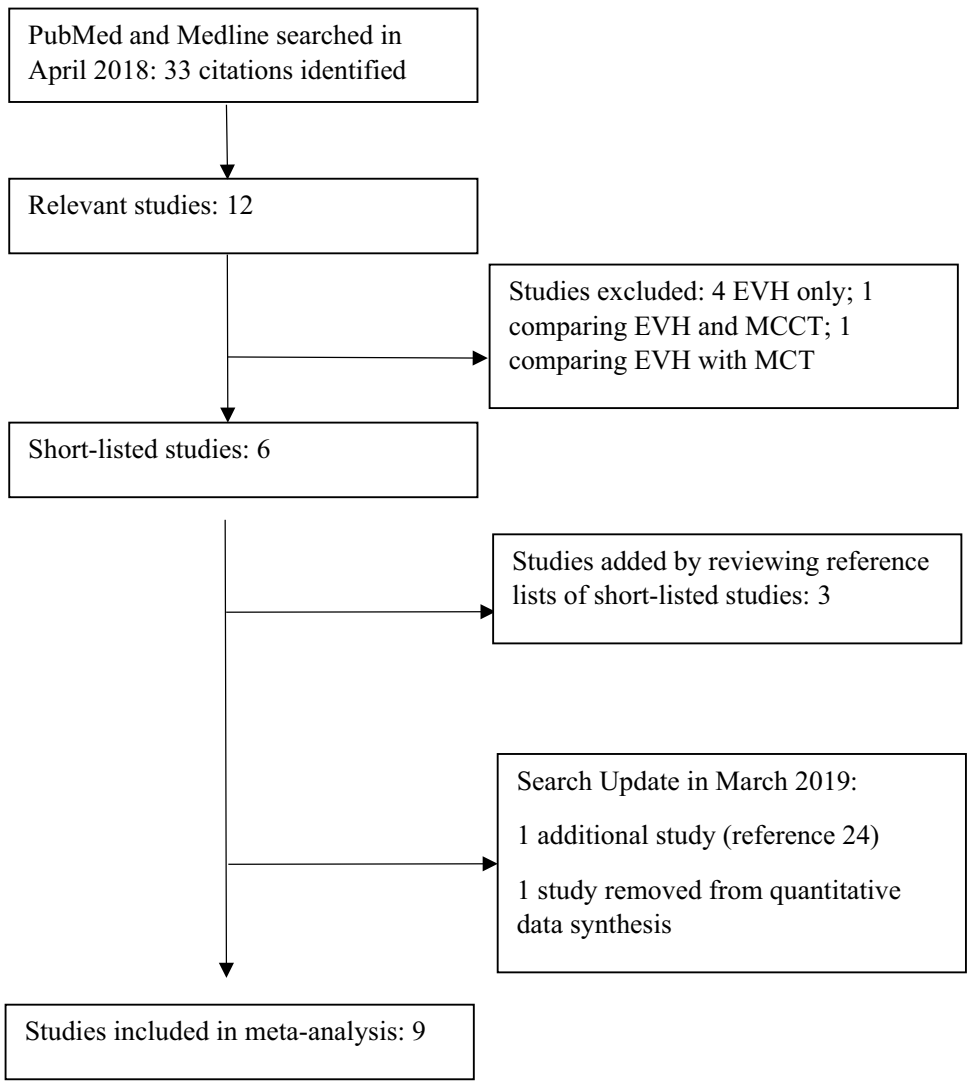

Fig. 1 Study search and selection process. EVH indicates eucapnic voluntary hyperventilation test; MCCT indicates methacholine challenge test; MCT indicates mannitol challenge test

Quality Assessment of Diagnostic Accuracy Studies tool [12] consisting of four key domains: patient selection, index test, reference standard, and the flow and timing (see data supplement). The authors adhered to the PRISMA (preferred reporting items for systematic reviews and meta-analyses) guidelines [13] for all stages of the design, implementation, and reporting of this meta-analysis.

Data Extraction and Synthesis

At least 10% decline in FEV_1 as (with either EVH or CPET) was considered as a 'positive' test result for bronchial hyperreactivity. Test results for EVH and CPET were extracted for each study into two-by-two contingency tables following the format in Table 1. As neither test was considered as the 'reference standard' (to compare one with the other), pooled positive agreement (a/a + c), and pooled negative agreement (d/b + d) between the two tests, were calculated, with EVH benchmarked against CPET, and then, pooled positive and negative agreements were calculated, with CPET benchmarked against EVH, using methods outlined in the Statistical Guidance on Reporting Results from Studies Evaluating Diagnostic Tests by the food and drug administration (FDA) [14]. The total number of study participants in each cell (a, b, c, d) was converted into percentages of the total population in each study. This is equivalent to measuring the sensitivity and specificity of EVH using CPET as the reference standard [15]. Positive and negative agreements were meta-analyzed using random effects meta-analysis of proportions in Meta-DiSc version 1.4, and the results were presented in a tabular format. Heterogeneity was assessed with I^2 index.

Results

The search identified 33 records, of which 12 were taken forward to full text assessment. After short-listing studies and hand-searching their reference lists, which contributed 3 additional studies, and updating the search (updated March, 2019), a total of 9 studies [10, 16–23] were included in this meta-analysis. These studies contributed enough data to calculate positive and negative agreement and be included in the meta-analysis. Table 2 outlines the baseline demographics of the study population as well as the details of the EVH and CPET tests. Most of the study participants were males, aged 8 years and above. One study [18] did not report any positive results with CPET, and two others [10, 22] did not report any negative results with CPET. Hence, in the overall pooled analyses, only six studies [16, 17, 19–21, 23] contributed to the pooled estimates. The pooled positive agreement between EVH and CPET (EVH benchmarked against CPET, latter as reference) was 0.62 [(95% confidence interval 0.54–0.70), I^2 77%], as shown in Fig. 2. The pooled negative agreement between the tests (EVH benchmarked against CPET, latter as reference) was 0.61 [(0.56–0.65)), I^2 81%], as shown in Fig. 3. The pooled positive agreement between CPET and EVH (CPET benchmarked against EVH, latter as reference) was 0.36 [(0.30–0.42), I^2 93%], as shown in Fig. 4. The pooled negative agreement between the tests (CPET benchmarked against EVH, latter as reference) was 0.82 [(0.77–0.86), I^2 78%], as shown in Fig. 5. Based on Table 3, the average of positive test results across all studies with EVH was greater than that of CPETs (58.84% vs. 39.51%). Heterogeneity was assessed by sub-group and sensitivity analyses. Sub-group analyses based on age (adults or children), performance level (athletes or non-athletes), and type of exercise protocol used (treadmill running or bicycle ergometry) are reported in Table 4a, b and c. Based on these analyses, the source of heterogeneity observed in the main analyses could be from the difference in performance level and type of exercise protocols used, as the 2 studies [17, 23] on children were mostly homogenous. However, I^2 index (showing level of heterogeneity) did not fall below 25% in all of the sub-group analyses, suggesting that additional sources of heterogeneity were present which could not be explored further.

Table 1 Contingency table of test agreement

	CPET positive	CPET negative
EVH positive	a + / +	b − / +
EVH negative	c ±	d − / −

EVH indicates eucapnic voluntary hyperventilation test; CPET indicates cardiopulmonary exercise test; a, b, c, d are cell values; ± refers to test positive or test negative

Discussion

This meta-analysis shows that positive agreement between EVH and CPET (with either model—EVH or CPET as reference) is poor. While there is some discrepancy between the two models on negative agreement, the negative agreement with the first model (CPET as reference) needs to be interpreted with caution due to the known false negative rate of EIB with CPETs. Overall, the results show that because of poor positive agreement between the two tests, they cannot be used inter-changeably for 'ruling out' EIB ("SnOut"), but given the high negative agreement (trusting the second model—using EVH as reference), the two tests would generate few false positive results for EIB and either test could be used to 'rule in' EIB ("SpIn"). Though not compared statistically, Table 3 shows that EVH diagnosed more cases of EIB as compared to CPET. While this may be a safe conclusion

Table 2 Baseline characteristics of study population

Study (Author)	Age Mean (SD)/range	Total *N*	Males (%)	Fitness level	History of asthma/ EIB	EVH protocol	CPET protocol
Schmidt (1983) [24]	11.6 (8–14)	14	N.R	N.R	100%	Gas 5% CO_2, 21% O_2, 74% N_2 Room temp: 21 °C, Relative humidity 40% Hyperventilation vigorously for 6 min	Treadmill test of 6 min. HR target 170 bpm after 2 min running (80% max HR) Follow-up FEV1: 0, 1, 3, 5, 10, 15 min post-test
Rosenthal (1984) [10]	N.R	10	N.R	N.R	100%	Gas (dry) 5% CO_2, 21% O_2, 74% N_2 room temp Hyperventilation at 40 L/min	Treadmill test Dry air and room temp
Eliasson (1992) [20]	22–42	20	82.5%	Clinic patients with suspected EIB	N.R	Gas (dry) compressed air 5% CO_2 5 min at 80% MVV	Bicycle ergometer at 60 rpm. Resistance increased by 25 watts/min Follow-up FEV1: 0, 5, 10, 15 min post-test
Rundell (2004) [23]	22.8 (6.8)	38	72.7	Athletes ranging in accomplishment from Developmental to Senior National Team levels	21%	Gas (dry) 5% CO_2, 21% O_2, 74% N_2 Room temp: 21 °C, Relative humidity: 40%; 85% MVV with FEV1 measurement at 5, 10, 15 min post challenge	Bicycle ergometry for 6 min Relative humidity: 50%, Temp: 2 ± 5.6 °C Follow-up FEV1: 5, 10, 15 min post-test
Dickinson (2006) [19]	22.6 (5.7)	14	N.R	Athlete volunteers from Great Britain short track speed skating and biathlon teams	14%	Gas (dry) 5% CO_2, 21% O_2, N_2, 74% Room temp: 19.1 °C, Relative humidity: < 2%; Hyperventilation for 6 min (30 × baseline FEV1) Follow-up FEV1: 3, 5, 10, 15 min post-test	Treadmill for 8 min with 90% max HR for the final 4 min Relative humidity: 56% Room temp: 18 °C Follow-up FEV1: 3, 5, 10, 15 min post-test
Castricum (2008) [17]	18.2 (4.9)	33	69.7%	Thirty-three volunteer elite swimmers, defined as State Level or above	39%	Gas (dry) 5% CO_2, 21% O_2, 74% N_2; Room temp: 21 ± 0.8 °C, Relative humidity: 60.5 ± 2.1% 85% MVV for 6 min Follow-up FEV1: 1, 3, 5, 7, 10 min post-test	Bicycle ergometry for 8 min Relative humidity: 60.5 ± 2.1%, Room temp: 21 ± 0.8 °C Follow-up FEV1: 1, 3, 5, 7, 10 min post-test

Table 2 (continued)

Study (Author)	Age Mean (SD)/range	Total *N*	Males (%)	Fitness level	History of asthma/ EIB	EVH protocol	CPET protocol
Pedersen (2008) [22]	18.3 (2.7)	16	0%	Elite swimmers	0%	Gas (dry) 5% CO_2, 21% O_2, 74% N_2 Room temp 85% MVV for 6 min Follow-up FEV1: 0, 1, 3, 5, 10, 15, 20 min post-test	Treadmill at constant speed; incline increased 2% every 2 min for 5 min average Follow-up FEV1: 0, 1, 3, 5, 10, 15, 20 min post-test
Chateaubriand (2015) [18]	11.9 (2.4)	34	55.8%	Children from outpatient clinic	100%	Gas (dry) 5% CO_2, 21% O_2, 74% N_2 MVV (baseline FEV1 × 21) for 6 min Follow-up FEV1: 3, 5, 7, 10, 15, and 30 min post-test	Treadmill running for 8 min (HR kept between 80–90% of maximum predicted for the last 6 min) Follow-up FEV1: 3, 5, 7, 10, 15, and 30 min
Filho (2018)	16 (13–20)	35	62%	Children from outpatient clinic with allergic rhinitis	0%	Gas (dry) 5% CO_2, 21% O_2, 74% N_2 MVV (baseline FEV1 × 21) for 6 min Follow-up FEV1: 3, 5, 7, 10, 15 and 30 min post test	The TR test lasted 8 to 9 min. Gradual increase in velocity in the first 2 min, up to target HR (80–90% of maximum predicted) for the final 6 min

N.R indicates data not reported; EIB indicates exercise-induced bronchoconstriction; EVH indicates eucapnic voluntary hyperventilation; CPET indicates cardiopulmonary exercise test; CO_2 indicates carbon dioxide; O_2 indicates oxygen; N_2 indicates nitrogen gas; bpm indicates beats per min; HR indicates heart rate; temp indicates temperature; FEV1 indicates forced expiratory volume in 1-s; MVV indicates maximum voluntary ventilation

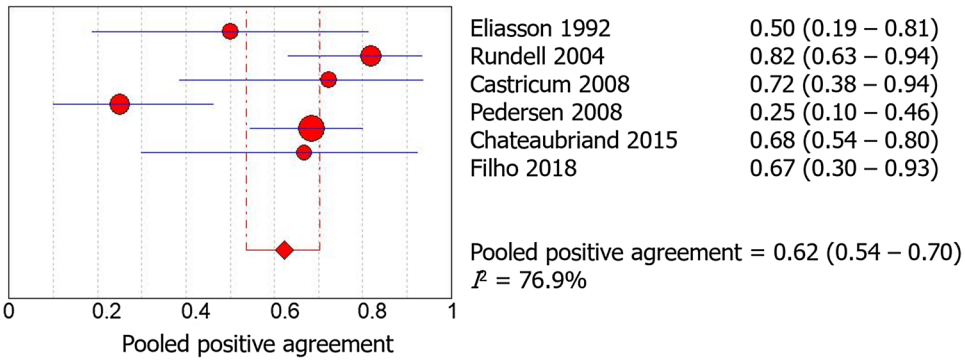

Fig. 2 Pooled positive agreement between EVH and CPET (CPET as reference). The size of the circle indicates the weight of the effect size as determined by the number of studies and participants. The diamond indicates the pooled effect. Pooled estimate is reported with 95% confidence intervals in parentheses

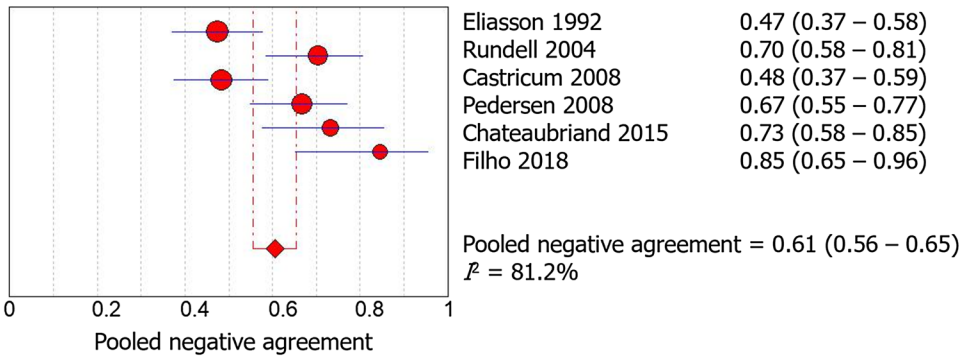

Fig. 3 Pooled negative agreement between EVH and CPET (CPET as reference). The size of the circle indicates the weight of the effect size as determined by the number of studies and participants. The diamond indicates the pooled effect. Pooled estimate is reported with 95% confidence intervals in parentheses

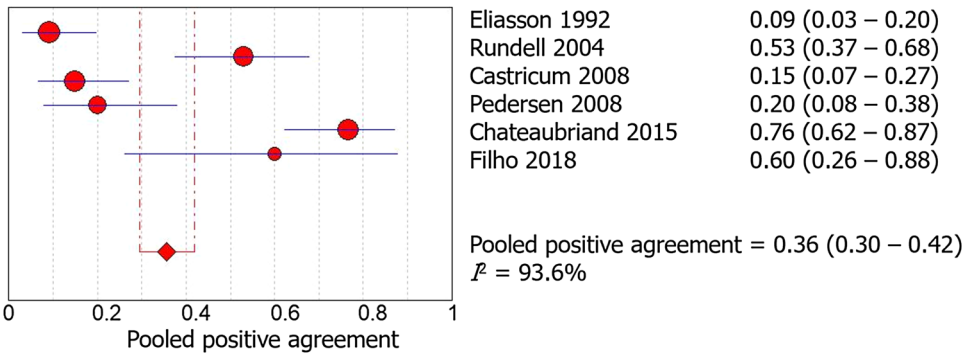

Fig. 4 Pooled positive agreement between CPET and EVH (EVH as reference). The size of the circle indicates the weight of the effect size as determined by the number of studies and participants. The diamond indicates the pooled effect. Pooled estimate is reported with 95% confidence intervals in parentheses

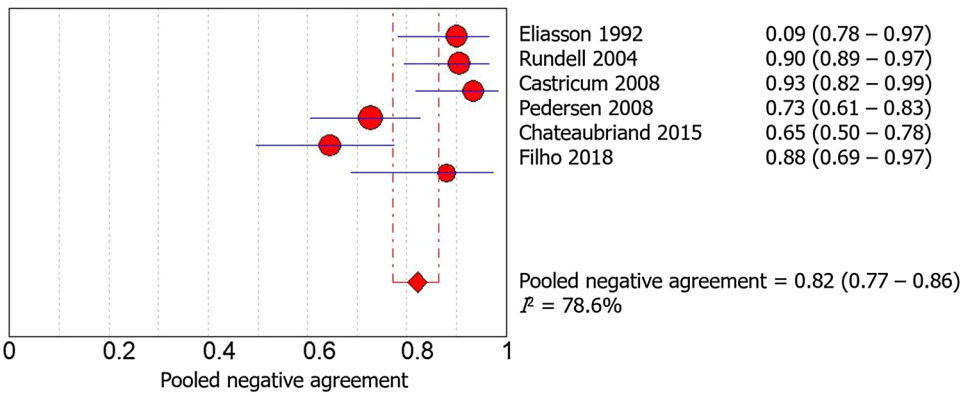

Fig. 5 Pooled negative agreement between CPET and EVH (EVH as reference). The size of the circle indicates the weight of the effect size as determined by the number of studies and participants. The diamond indicates the pooled effect. Pooled estimate is reported with 95% confidence intervals in parentheses

Table 3 Comparison of test positivity between EVH and CPET

Studies	% Positive with EVH	% Positive with CPET
Schmidt (1983)	71	100
Rosenthal (1984)	90	100
Eliasson (1992)	55	10
Rundell (2004)	44.6	29
Dickinson (2006)	71	0
Castricum (2008)	53	11
Pedersen (2008)	31	25
Chateaubriand (2015)	50	55.6
Filho (2018)	28	25
Average	54.84	39.51

from these analyses, in the clinical setting, indications of each test would differ based on different clinical situations. Nevertheless, this is the first meta-analysis that has summarized the agreements between EVH and CPET.

Currently, EVH is the recommended bronchoprovocation test endorsed by the International Olympic Committee Medical Commission (IOC-MC) to diagnose EIB in Olympic athletes [24]. However, the American Thoracic Society (ATS) clinical practice guidelines [9] do not state any preference of EVH or CPET. Instead the guidelines, recommend an 'exercise'—based challenge and terms EVH as a surrogate test to diagnose EIB [9]. Although, both IOC-MC and ATS guidelines use FEV1 as the marker for the diagnosis of EIB, other measurements such as forced vital capacity (FVC) and peak expiratory flow rate (PEFR) can be complementary in the diagnosis of EIB as using only one criterion could result in an underestimation in the prevalence of EIB in a certain subset of population. Parsons et al. in a study using EVH to document EIB in college athletes, showed that when [25] in a study using EVH to document EIB in college athletes, showed that when FEV1 was used as the only criterion for a positive test, the prevalence of EIB was 19%, but when additional criteria were examined- i.e., decline if FEV1 by 5%, and decline in PEFR by 20%, the prevalence increased to 38%.

While it seems intuitive to measure the change in FEV_1, before and after a field- exercise challenge, to diagnose EIB, difficulties with 'field' exercise settings, specifically, inability to control alterations in exercise load and intensity and other ambient conditions, inherently limit the application of this approach [26]. Therefore, while field-testing may be specific, it has poor diagnostic sensitivity [21]. Standardized laboratory-based exercise challenges, despite the efforts to standardize these and their interpretation, may also fail to properly reproduce the broncho-provocative stimulus experienced by athletes when practicing in their own sports environment [27]. It could also be debated whether EVH represents an appropriate diagnostic test for both athletes and non-athletes, who might not easily maintain high ventilation rates for a prolonged time. While one of the major hallmarks of EVH testing in athletes is the ability to achieve a minute ventilation rate that is somewhat identical to the demands of high-intensity exercise, the increased ventilation may not reflect real life, i.e., EVH may cause a greater degree of airway wall desiccation than typical exercise in normal subjects would. This is particularly important in the evaluation of EIB in swimmers, in whom the prevalence of EIB is known to be high [28]. Given the obvious differences between the dry and cold air during EVH and the humid and warm environment in the pool setting, even though it seems that pool-side challenge test would diagnose more EIB cases, evidence from the studies suggests to the contrary [16]. Another aspect of EVH that contrasts with CPET is that in general, in most healthy subjects after a CPET, an increase in FEV_1 (of ~ 10%, physiological bronchodilation) would normally be expected, therefore, a decrease in FEV_1 of at least 10% in any subject would clearly be a significant finding. However, it is not entirely clear if a decline of similar magnitude of FEV_1 after EVH would have the same physiological significance. This may partly explain why EVH would diagnose more EIB cases.

Whether or not a cold air or dry air challenge in EVH should be used, is unclear. At least theoretically, inhalation of cold and dry air, as compared to warm and humidified air, would place a greater burden on the airways to warm and humidify the air quickly. Therefore, inspiration of cold and dry air during EVH would theoretically increase the likelihood of dehydration in airways. McFadden et al. showed that 8 min of EVH with warm air decreased FEV_1 to the same extent as 4 min of EVH with cold air in eight study participants [29]. In a study by Eliasson et al., EVH maneuvers while inhaling air either at room temperature or cooled to − 18 to − 26 °C, showed no difference in FEV_1 responses [19].

This meta-analysis has strengths as well as limitations. In the absence of more robust methods of assessing test accuracy for EIB, this meta-analysis presents positive and negative agreements with both EVH and CPET in different models as the references, as recommended by FDA guidance [14]. The pooled results from our meta-analysis should be viewed as 'exploratory' in nature. It is also important to note that test agreement is not a measure of test accuracy as it neither considers any given test as the 'true' test (gold standard), nor does it consider if both tests, even if they agree, are 'false'. Instead, by using both tests as references in different models, different views about the level of agreement between the tests can be explored and discussed in results. The results of this meta-analysis show high heterogeneity. We attempted to explore this high heterogeneity by several sub-group analyses. However, some of the sub-group analyses still showed

Table 4 Sub-group analyses to explore heterogeneity

(a) Sub-group analysis based on age (adults)			
Pooled positive agreement between EVH and CPET (CPET as reference)		Pooled positive agreement between CPET and EVH (EVH as reference)	
Eliasson (1992)	0.50 (0.18–0.81)	Eliasson (1992)	0.09 (0.03–0.20)
Rundell (2004)	0.81 (0.63–0.93)	Rundell (2004)	0.52 (0.37–0.68)
Pedersen (2008)	0.25 (0.10–0.46)	Pedersen (2008)	0.20 (0.07–0.38)
Castricum (2008)	0.72 (0.38–0.93)	Castricum (2008)	0.14 (0.06–0.27)
Pooled positive agreement	0.57 (0.45–0.68), I^2 **85%**	Pooled positive agreement	0.23 (0.17–0.30), I^2 **89%**
Pooled negative agreement between EVH and CPET (CPET as reference)		Pooled negative agreement between CPET and EVH (EVH as reference)	
Eliasson (1992)	0.47 (0.37–0.57)	Eliasson (1992)	0.90 (0.78–0.96)
Rundell (2004)	0.70 (0.58–0.80)	Rundell (2004)	0.90 (0.79–0.96)
Pedersen (2008)	0.66 (0.54–0.77)	Pedersen (2008)	0.72 (0.60–0.82)
Castricum (2008)	0.48 (0.37–0.59)	Castricum (2008)	0.93 (0.81–0.98)
Pooled negative agreement	0.57 (0.51–0.62), I^2 **79%**	Pooled negative agreement	0.85 (0.80–0.89), I^2 **75%**
(b) Sub-group analysis based on age (children)			
Pooled positive agreement between EVH and CPET (CPET as reference)		Pooled positive agreement between CPET and EVH (EVH as reference)	
Chateaubriand (2015)	0.68 (54–0.80)	Chateaubriand (2015)	0.76 (0.62–0.87)
Filho (2018)	0.66 (0.29–0.92)	Filho (2018)	0.60 (0.26–0.87)
Pooled positive agreement	0.68 (0.55–0.79), I^2 **0%**	Pooled positive agreement	0.73 (0.60–0.84), I^2 **8%**
Pooled negative agreement between EVH and CPET (CPET as reference)		Pooled negative agreement between CPET and EVH (EVH as reference)	
Chateaubriand (2015)	0.73 (0.57–0.85)	Chateaubriand (2015)	0.64 (0.49–0.77)
Filho (2018)	0.84 (0.65–0.95)	Filho (2018)	0.88 (0.68–0.97)
Pooled negative agreement	0.77 (0.65–0.86), I^2 **21%**	Pooled negative agreement	0.72 (0.60–0.82), I^2 **80%**
(c) Sub-group analysis based on performance (athletes)			
Pooled positive agreement between EVH and CPET (CPET as reference)		Pooled positive agreement between CPET and EVH (EVH as reference)	
Rundell (2004)	0.81 (0.63–0.93)	Rundell (2004)	0.52 (0.37–0.68)
Castricum (2008)	0.72 (0.38–0.93)	Castricum (2008)	0.14 (0.06–0.27)
Pedersen (2008)	0.25 (0.10–0.46)	Pedersen (2008)	0.20 (0.07–0.38)
Pooled positive agreement	0.58 (0.45–0.70), I^2 **89%**	Pooled positive agreement	0.29 (0.21–0.37), I^2 **89%**
Pooled negative agreement between EVH and CPET (CPET as reference)		Pooled negative agreement between CPET and EVH (EVH as reference)	
Rundell (2004	0.70 (0.58–0.80)	Rundell (2004)	0.90 (0.79–0.96)
Castricum (2008)	0.48 (0.37–0.59)	Castricum (2008)	0.93 (0.81–0.98)
Pedersen (2008)	0.66 (0.54–0.77)	Pedersen (2008)	0.72 (0.60–0.82)
Pooled negative agreement	0.60 (0.54–0.67), I^2 **79%**	Pooled negative agreement	0.84 (0.77–0.89), I^2 **82%**
(d) Sub-group analysis based on performance (non-athletes)			
Pooled positive agreement between EVH and CPET (CPET as reference)		Pooled positive agreement between CPET and EVH (EVH as reference)	
Eliasson (1992)	0.50 (0.18–0.81)	Eliasson (1992)	0.09 (0.03–0.20)
Chateaubriand (2015)	0.68 (0.54–0.80)	Chateaubriand (2015)	0.76 (0.62–0.87)
Filho (2018)	0.66 (0.29–0.92)	Filho (2018)	0.60 (0.26–0.87)
Pooled positive agreement	0.65 (0.53–0.76), I^2 **0%**	Pooled positive agreement	0.42 (0.33–0.52), I^2 **96%**
Pooled negative agreement between EVH and CPET (CPET as reference)		Pooled negative agreement between CPET and EVH (EVH as reference)	
Eliasson (1992)	0.47 (0.37–0.57)	Eliasson (1992)	0.90 (0.78–0.96)
Chateaubriand (2015)	0.73 (0.57–0.85)	Chateaubriand (2015)	0.64 (0.49–0.77)

Table 4 (continued)

(d) Sub-group analysis based on performance (non-athletes)			
Filho (2018)	0.84 (0.65–0.95)	Filho (2018)	0.88 (0.68–0.97)
Pooled negative agreement	0.60 (0.52–0.67), I^2 **88%**	Pooled negative agreement	0.79 (0.71–0.86), I^2 **82%**
(e) Sub-group analysis based on type of exercise test (bicycle ergometer)			
Pooled positive agreement between EVH and CPET (CPET as reference)		Pooled positive agreement between CPET and EVH (EVH as reference)	
Eliasson (1992)	0.50 (0.18–0.81)	Eliasson (1992)	0.09 (0.03–0.20)
Rundell (2004)	0.81 (0.63–0.93)	Rundell (2004)	0.52 (0.37–0.68)
Castricum (2008)	0.72 (0.38–0.93)	Castricum (2008)	0.14 (0.06–0.27)
Pooled positive agreement	0.73 (0.58–0.84), I^2 **44%**	Pooled positive agreement	0.23 (0.17–0.31), I^2 **92%**
Pooled negative agreement between EVH and CPET (CPET as reference)		Pooled negative agreement between CPET and EVH (EVH as reference)	
Eliasson (1992)	0.47 (0.37–0.57)	Eliasson (1992)	0.90 (0.78–0.96)
Rundell (2004)	0.70 (0.58–0.80)	Rundell (2004)	0.90 (0.79–0.96)
Castricum (2008)	0.48 (0.37–0.59)	Castricum (2008)	0.93 (0.81–0.98)
Pooled negative agreement	0.54 (0.47–0.60), I^2 **81%**	Pooled negative agreement	0.91 (0.85–0.95), I^2 **0%**
(f) Sub-group analysis based on type of exercise test (treadmill running)			
Pooled positive agreement between EVH and CPET (CPET as reference)		Pooled positive agreement between CPET and EVH (EVH as reference)	
Pedersen (2008)	0.25 (0.10–0.46)	Pedersen (2008)	0.20 (0.07–0.38)
Chateaubriand (2015)	0.68 (0.54–0.80)	Chateaubriand (2015)	0.76 (0.62–0.87)
Filho (2018)	0.66 (0.29–0.92)	Filho (2018)	0.60 (0.26–0.87)
Pooled positive agreement	0.56 (0.45–0.66), I^2 **85%**	Pooled positive agreement	0.55 (0.44–0.65), I^2 **92%**
Pooled negative agreement between EVH and CPET (CPET as reference)		Pooled negative agreement between CPET and EVH (EVH as reference)	
Pedersen (2008)	0.66 (0.54–0.77)	Pedersen (2008)	0.72 (0.60–0.82)
Chateaubriand (2015)	0.73 (0.57–0.85)	Chateaubriand (2015)	0.64 (0.49–0.77)
Filho (2018)	0.84 (0.65–0.95)	Filho (2018)	0.88 (0.68–0.97)
Pooled negative agreement	0.71 (0.63–0.79), I^2 **40%**	Pooled negative agreement	0.72 (0.64–0.79), I^2 **60%**

EVH indicates eucapnic voluntary hyperventilation test; CPET indicates cardiopulmonary exercise test; I^2 indicates heterogeneity; $I^2 < 25\%$ indicates low heterogeneity, 25–70% indicates moderate heterogeneity, and > 70% indicates high heterogeneity

high heterogeneity, suggesting that other sources of heterogeneity were present that could not be adequately explored, which is a limitation of this meta-analysis. Aside from what has been reported, some heterogeneity could also be due to spectrum bias (studies conducted in different sub-sets of population with varying levels of prevalence). Accepting these limitations, we believe findings of this first-ever meta-analysis on test-agreement between EVH and CPET are qualitatively (if not quantitatively) informative.

In conclusion, the results of this meta-analysis show poor positive agreement between EVH and CPET but the negative agreement (using EVH as reference) is high which means either test can be relied upon for correctly identifying those without EIB. Future studies should evaluate not only test–retest repeatability of EVH (especially in those with mild bronchial hyper-responsiveness), but also determine the role of EVH in predicting treatment outcome in patients with EIB.

Author Contributions IHI had full access to all of the extracted data in the network meta-analysis and takes responsibility for the integrity of the data and the accuracy of the data analysis. IHI conceptualized and designed the study protocol, conducted the analyses, and wrote the first draft manuscript. IHI and MG contributed to assessment of study quality. All authors contributed substantially to the interpretation of analyses and in revisions of manuscript.

Funding This study was not sponsored by any funding agency or pharmaceutical company.

Compliance with Ethical Standards

Conflicts of interest The authors declare that they have no conflicts of interest.

References

1. Weiler JM, Anderson SD, Randolph C, Bonini S, Craig TJ, Pearlman DS, Rundell KW, Silvers WS, Storms WW, Bernstein DI et al (2010) Pathogenesis, prevalence, diagnosis, and management of exercise-induced bronchoconstriction: a practice parameter. Ann Allergy Asthma Immunol 105:S1–47
2. Price OJ, Hull JH, Backer V, Hostrup M, Ansley L (2014) The impact of exercise-induced bronchoconstriction on athletic performance: a systematic review. Sports Med 44:1749–1761
3. Parsons JP (2014) Exercise-induced bronchoconstriction. Otolaryngol Clin North Am 47:119–126
4. van Leeuwen JC, Driessen JM, Kersten ET, Thio BJ (2013) Assessment of exercise-induced bronchoconstriction in adolescents and young children. Immunol Allergy Clin North Am 33:381–394
5. Ansley L, Kippelen P, Dickinson J, Hull JH (2012) Misdiagnosis of exercise-induced bronchoconstriction in professional soccer players. Allergy 67:390–395
6. Lund TK, Pedersen L, Anderson SD, Sverrild A, Backer V (2009) Are asthma-like symptoms in elite athletes associated with classical features of asthma? Br J Sports Med 43:1131–1135
7. Rundell KW, Im J, Mayers LB, Wilber RL, Szmedra L, Schmitz HR (2001) Self-reported symptoms and exercise-induced asthma in the elite athlete. Med Sci Sports Exerc 33:208–213
8. Bonini M, Lapucci G, Petrelli G, Todaro A, Pamich T, Rasi G, Bonini S (2007) Predictive value of allergy and pulmonary function tests for the diagnosis of asthma in elite athletes. Allergy 62:1166–1170
9. Parsons JP, Hallstrand TS, Mastronarde JG, Kaminsky DA, Rundell KW, Hull JH, Storms WW, Weiler JM, Cheek FM, Wilson KC et al (2013) An official American Thoracic Society clinical practice guideline: exercise-induced bronchoconstriction. Am J Respir Crit Care Med 187:1016–1027
10. Rosenthal RR (1984) Simplified eucapnic voluntary hyperventilation challenge. J Allergy Clin Immunol 73:676–679
11. Anderson SD, Argyros GJ, Magnussen H, Holzer K (2001) Provocation by eucapnic voluntary hyperpnoea to identify exercise induced bronchoconstriction. Br J Sports Med 35:344–347
12. Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, Leeflang MM, Sterne JA, Bossuyt PM (2011) Group Q-: QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med 155:529–536
13. Moher D, Liberati A, Tetzlaff J, Altman DG, Group P (2009) Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. J Clin Epidemiol 62:1006–1012
14. USFDA (2007) Statistical guidance on reporting results from studies evaluating diagnostic tests. https://www.fda.gov/RegulatoryInformation/Guidances/ucm071148.htm. Accessed 15 April 2018
15. Freeman K, Tsertsvadze A, Taylor-Phillips S, McCarthy N, Mistry H, Manuel R, Mason J (2017) Agreement between gastrointestinal panel testing and standard microbiology methods for detecting pathogens in suspected infectious gastroenteritis: test evaluation and meta-analysis in the absence of a reference standard. PLoS ONE 12:e0173196
16. Castricum A, Holzer K, Brukner P, Irving L (2010) The role of the bronchial provocation challenge tests in the diagnosis of exercise-induced bronchoconstriction in elite swimmers. Br J Sports Med 44:736–740
17. Chateaubriand do Nascimento Silva Filho MJ, Goncalves AV, Tavares Viana M, Peixoto DM, Cavalcanti Sarinho ES, Rizzo JA (2015) Exercise-induced bronchoconstriction diagnosis in asthmatic children: comparison of treadmill running and eucapnic voluntary hyperventilation challenges. Ann Allergy Asthma Immunol 115:277–281
18. Dickinson JW, Whyte GP, McConnell AK, Harries MG (2006) Screening elite winter athletes for exercise induced asthma: a comparison of three challenge methods. Br J Sports Med 40:179–182 (**discussion 179–182**)
19. Eliasson AH, Phillips YY, Rajagopal KR, Howard RS (1992) Sensitivity and specificity of bronchial provocation testing. An evaluation of four techniques in exercise-induced bronchospasm. Chest 102:347–355
20. Pedersen L, Winther S, Backer V, Anderson SD, Larsen KR (2008) Airway responses to eucapnic hyperpnea, exercise, and methacholine in elite swimmers. Med Sci Sports Exerc 40:1567–1572
21. Rundell KW, Anderson SD, Spiering BA, Judelson DA (2004) Field exercise vs laboratory eucapnic voluntary hyperventilation to identify airway hyperresponsiveness in elite cold weather athletes. Chest 125:909–915
22. Schmidt A, Bach-Mortensen N, Bundgaard A (1983) Comparison of isocapnic hyperventilation and treadmill exercise in children with exercise-induced asthma. Scand J Clin Lab Invest 43:203–206
23. AlbuquerqueRodriguesFilho E, Rizzo JA, Goncalves AV, CorreiaJunior MAV, Sarinho ESC, Medeiros D (2018) Exercise-induced bronchospasm in children and adolescents with allergic rhinitis by treadmill and hyperventilation challenges. Respir Med 138:102–106
24. Fitch KD, Sue-Chu M, Anderson SD, Boulet LP, Hancox RJ, McKenzie DC, Backer V, Rundell KW, Alonso JM, Kippelen P et al (2008) Asthma and the elite athlete: summary of the International Olympic Committee's consensus conference, Lausanne, Switzerland, January 22–24, 2008. J Allergy Clin Immunol 122:254–260
25. Parsons JP, Kaeding C, Phillips G, Jarjoura D, Wadley G, Mastronarde JG (2007) Prevalence of exercise-induced bronchospasm in a cohort of varsity college athletes. Med Sci Sports Exerc 39:1487–1492
26. Dryden DM, Spooner CH, Stickland MK, Vandermeer B, Tjosvold L, Bialy L, Wong K, Rowe BH (2010) Exercise-induced bronchoconstriction and asthma. Evid Rep Technol Assess (Full Rep) 189:1–154
27. Anderson SD, Lambert S, Brannan JD, Wood RJ, Koskela H, Morton AR, Fitch KD (2001) Laboratory protocol for exercise asthma to evaluate salbutamol given by two devices. Med Sci Sports Exerc 33:893–900
28. Bougault V, Turmel J, Levesque B, Boulet LP (2009) The respiratory health of swimmers. Sports Med 39:295–312
29. McFadden ER Jr, Nelson JA, Skowronski ME, Lenner KA (1999) Thermally induced asthma and airway drying. Am J Respir Crit Care Med 160:221–226