

# BMJ Open

BMJ Open is committed to open peer review. As part of this commitment we make the peer review history of every article we publish publicly available.

When an article is published we post the peer reviewers' comments and the authors' responses online. We also post the versions of the paper that were used during peer review. These are the versions that the peer review comments apply to.

The versions of the paper that follow are the versions that were submitted during the peer review process. They are not the versions of record or the final published versions. They should not be cited or distributed as the published version of this manuscript.

BMJ Open is an open access journal and the full, final, typeset and author-corrected version of record of the manuscript is available on our site with no access controls, subscription charges or pay-per-view fees (<http://bmjopen.bmj.com>).

If you have any questions on BMJ Open's open peer review process please email [editorial.bmjopen@bmj.com](mailto:editorial.bmjopen@bmj.com)

# BMJ Open

## Diagnostic Value of Transthoracic Echocardiography for Pulmonary Hypertension: a Systematic Review and Meta-Analysis

Journal:	<i>BMJ Open</i>
Manuscript ID	bmjopen-2019-033084
Article Type:	Research
Date Submitted by the Author:	20-Jul-2019
Complete List of Authors:	Ni, Jin-Rong; The First Hospital of Lanzhou University Yan, Pei-Jing; Institute of Clinical Research and Evidence Based Medicine, Gansu Provincial Hospital Liu, Shi-Dong; The First Hospital of Lanzhou University Hu, Yuan; the First Hospital of Lanzhou University Yang, Kehu; evidence based meidicine center, lanzhou university Lei, Jun-Qiang; The First Hospital of Lanzhou University Song, Bing; The First Hospital of Lanzhou University
Keywords:	Hypertension < CARDIOLOGY, Echocardiography < CARDIOLOGY, Diagnostic radiology < RADIOLOGY & IMAGING

SCHOLARONE™  
Manuscripts

## Diagnostic Value of Transthoracic Echocardiography for Pulmonary Hypertension: a Systematic Review and Meta-Analysis

Jin-Rong Ni<sup>1,3,8,9</sup>, Pei-Jing Yan<sup>2,5,6#</sup>, Shi-Dong Liu<sup>1,3</sup>, Yuan Hu<sup>3</sup>, Ke-Hu Yang<sup>2,4,5,6</sup>, Jun-Qiang Lei<sup>1,7,8,9\*</sup>, Bing Song<sup>3\*</sup>

<sup>1</sup>The First Hospital (the First Clinical Medical School) of Lanzhou University, Lanzhou 730000, China;

<sup>2</sup>Institute of Clinical Research and Evidence Based Medicine, Gansu Provincial Hospital, Lanzhou, 730000, China; <sup>3</sup>Department of Cardiovascular Surgery, the First Hospital of Lanzhou University,

Lanzhou 730000, China; <sup>4</sup>Evidence-Based Medicine Center, School of Basic Medical Sciences, Lanzhou

University, Lanzhou 730000, China; <sup>5</sup>Evidence Based Social Science Research Center, Lanzhou

University, Lanzhou, 730000, China; <sup>6</sup>Key Laboratory of Evidence-based Medicine and Knowledge

Translation of Gansu Province, Lanzhou, 730000, China. <sup>7</sup>Department of Radiology, the First Hospital

of Lanzhou University, Lanzhou 730000, China. <sup>8</sup>Intelligent Imaging Medical Engineering Research

Center of Gansu province, Lanzhou 730000, China; <sup>9</sup>Precision Image and Collaborative Innovation

International Scientific and Technological Cooperation Base of Gansu province, Lanzhou 730000, China.

#Joint first author

\*Joint corresponding authors

**Corresponding Author:** 1. Jun-Qiang Lei, No.11 Donggang West Road, Chengguan District,

Lanzhou, China.

Tel: +8613919289040

E-mail: leijunqiangldyy@163.com

2. Bing Song, No.11 Donggang West Road, Chengguan District,

Lanzhou, China.

Tel: +8613993118707

E-mail: songbingldyyxwk@163.com

**Funding sources:** This study was supported by the Key Laboratory of Evidence Based Medicine and Knowledge Translation Foundation of Gansu Province (Grant No. GSXZYZH2018006).

**Statement on potential conflicts of interest:** None declared.

**Patient and Public Involvement:** There is no patient or the public involved in our work.

**Data Availability:** All data relevant to the study are included in the article.

**Key Words:** Echocardiography; Pulmonary Hypertension; Catheterization; Diagnosis; Accuracy; Meta-analysis.

## Abstract

**Objective:** To evaluate the diagnostic value of TTE in patients with PH and preliminarily explore the factors that may affect the diagnostic accuracy of TTE.

**Design:** Systematic review and meta-analysis.

**Data sources and eligibility criteria:** EMBASE, Cochrane Library for clinical trials, PubMed and Web of Science were searched from inception to June 19, 2019. Studies using both TTE and RHC to diagnose PH were included.

**Mean results:** 27 studies involving 4386 subjects were eligible for analysis. TTE had a pooled sensitivity 85% (95% CI 81–90%), a pooled specificity of 74% (95% CI 64–81%), a pooled positive likelihood ratio of 3.2 (95% CI 2.3–4.4), a pooled negative likelihood ratio of 0.20 (95% CI 0.15–0.26), a pooled diagnostic odds ratio of 16 (95% CI 10–27), and finally an area under the SROC curve of 0.88 (95%CI 0.85–0.90).

**Conclusion:** The value of TTE in diagnosing PH is certain, although it cannot yet replace RHC as the gold standard at this stage. The time interval between TTE and RHC, the threshold value of TTE and the disease composition of the study population may all be factors affecting the diagnostic value of TTE.

**Review registration number:** PROSPERO CRD42019123289.

### Strengths and limitations of this study

1. We conducted a systematic and comprehensive search of the main database, included more studies, and obtained a large sample size.
2. Detailed subgroup analysis and sensitivity analysis were performed.
3. The types of pulmonary hypertension included in the studies could not be distinguished.
4. No unified procedure for measuring pulmonary artery pressure by transthoracic echocardiography.

## 1 Introduction

2 The prevalence of pulmonary hypertension (PH) is estimated 1% in the general population,  
3 and as high as 10% in the 600 million people older than 65<sup>1</sup>. With the aggravation of population  
4 aging, PH will become a global health problem<sup>2</sup>. Early detection and accurate assessment is  
5 vital for improved outcomes for PH patients<sup>3</sup>. Right heart catheterization (RHC) is the gold  
6 standard for accurate measurements of pulmonary pressures and for the diagnosis of PH<sup>4</sup>. But  
7 RHC is invasive and it cannot be used frequently or repeatedly<sup>5</sup>. Transthoracic  
8 echocardiography (TTE) is a noninvasive test recommended for use in screening for PH<sup>4</sup>. The  
9 feasibility of TTE in evaluating systolic pulmonary artery pressure (SPAP) through tricuspid  
10 regurgitation has been confirmed in previous studies<sup>6 7</sup>. TTE is also able to provide crucial  
11 information on heart size and function<sup>8</sup>.

12 Despite the frequent use of TTE for screening for PH, its diagnostic accuracy and clinical  
13 value has courted much controversy. The latest guideline<sup>4</sup> for PH suggest that peak tricuspid  
14 regurgitation velocity (TRV) can be used to determine the likelihood of PH. But clinicians often  
15 expect to get a specific numerical value by echocardiography to evaluate the condition, observe  
16 the curative effect and judge the prognosis. Therefore echocardiographic method for estimating  
17 SPAP by tricuspid regurgitation are still adopted in clinic.

18 From 2010 to 2013, three reviews<sup>9-11</sup> unanimously concluded that TTE could only be used  
19 as a crude screening tool and was not suitable for the diagnosis of PH. However, the studies  
20 they included were published before 2010. In recent years, many new original studies related  
21 to this topic have emerged. In addition, an ideal evidence system should integrate and evaluate  
22 all important research evidence related to specific clinical problems<sup>12</sup>. So high quality meta-

23 analysis has been increasingly regarded as one of the key tools for achieving evidence<sup>13 14</sup>. Thus,  
24 the purpose of this study is to undertake an updated systematic review and quantitative meta-  
25 analysis on the value of TTE for diagnosing PH and preliminarily explore the factors that may  
26 affect the diagnostic accuracy of TTE.

## 27 **Methods**

28 The present study is reported as per the Preferred Reporting Items for Systematic Reviews  
29 and Meta-analyses (PRISMA) statement<sup>15</sup> and the published recommendations<sup>16</sup>. The detailed  
30 protocol is accessible in PROSPERO (CRD42019123289)<sup>17 18</sup>.

### 31 **Data sources and search**

32 We performed a systematic search in EMBASE, Cochrane Library, PubMed and Web of  
33 Science for relevant literatures from inception to June 19, 2019. Subject words were combined  
34 with free words, and the search strategies were developed and adapted for each database  
35 (appendix 1). For unpublished trials, we searched ClinicalTrials.gov and the trials registers on  
36 the World Health Organization International Clinical Trials Registry Platform. We also  
37 reviewed the references of included studies and other systematic reviews and meta-analysis to  
38 obtain a comprehensive list of included studies.

### 39 **Study selection**

40 Studies were selected based on the following inclusion criteria: studies diagnosing PH by  
41 TTE; the study population was patients with suspected PH; TTE measurements of SPAP were  
42 performed using tricuspid regurgitation; RHC was used as the gold standard for the diagnosis  
43 of PH.



1  
2  
3  
4 44 Exclusion criteria were: insufficient data to construct a 2×2 table; the tricuspid regurgitation  
5  
6 45 method was not used to calculate pulmonary artery pressure; studies with fewer than 20 subjects;  
7  
8  
9 46 duplicate data is used (in which case, select the largest sample or the latest study).  
10

11 47 Two reviewers (JR.N and PJ.Y) independently screened eligible studies for suitability.  
12  
13  
14 48 Disagreements were resolved by consensus. If consensus could not be reached, a third reviewer  
15  
16  
17 49 (SD.L) was deferred to for arbitration. No language restriction was applied.  
18  
19

### 20 50 **Data extraction**

21  
22  
23 51 Two reviewers (JR.N and PJ.Y) extracted data independently as per a predefined data  
24  
25 52 extraction sheet. The following variables were extracted from included studies: lead author,  
26  
27  
28 53 publication year, country of study, study design, study population demographics, sample size,  
29  
30  
31 54 mean age, male ratio, the time interval between TTE and RHC, the cut-off threshold levels for  
32  
33 55 TTE and RHC, and the number of true-positive (TP), false-negative (FN), true-negative (TN)  
34  
35  
36 56 and false-positive (FP) observations. Extracted data was cross-checked and disagreements were  
37  
38  
39 57 resolved via discussion or referral to a third reviewer (Y.H).  
40

### 41 58 **Quality assessment**

42  
43  
44 59 The Quality Assessment of Diagnostic Accuracy Studies QUADAS-2 tool was used to assess  
45  
46  
47 60 the risk of bias and clinical applicability concerns of included studies as per the Cochrane  
48  
49  
50 61 Collaboration recommendation<sup>19 20</sup>. Two reviewers (JR.N and PJ.Y) independently evaluated  
51  
52  
53 62 QUADAS-2 items, and all emerging conflicts were resolved by consensus.  
54

### 55 63 **Data synthesis and statistical analysis**

56  
57  
58 64 All statistical analyses were performed with STATA/SE version 15.1 (Stata Corp, College  
59  
60

65 Station, TX) and Review Manager Version 5.3 software (Copenhagen, Denmark, Nordic  
66 Cochrane Centre, Cochrane Collaboration, 2014). All tests were two-tailed. A  $P$  value < 0.05  
67 was considered statistically significant.

68 The correlation coefficient between the logarithm of sensitivity and logarithm of one minus  
69 specificity (i.e. the false-positive rate) was calculated to test whether the threshold effect was  
70 one of the sources of heterogeneity<sup>21</sup>. Deeks' test was used to test for publication bias<sup>22</sup>. The  
71 bivariate model for diagnostic meta-analysis was used to obtain pooled estimates of sensitivity  
72 and specificity<sup>23</sup>. Statistical heterogeneity among studies was explored using the  $I^2$  statistic.

73 Pooled sensitivity, specificity, diagnostic odds ratio (DOR), positive (PLR) , negative  
74 (PLR) likelihood ratios, and the area under the summary receiver operating curve (SROC) were  
75 calculated from the number of TPs, FNs, FPs, and TNs. The 95% confidence intervals (CIs)  
76 were estimated for each metric.

77 Subgroup analyses were undertaken based on the following variables: the time interval  
78 between TTE and RHC; disease classification of the study population; publication year of the  
79 study; study design (prospective or retrospective studies) and cut-off threshold of TTE  
80 diagnostic PH. Sensitivity analysis was undertaken by excluding low-quality studies (according  
81 to the QUADAS-2 quality assessment) or trials with characteristics different from the others.

## 82 **Results**

### 83 **Studies Retrieved and Characteristics**

84 Figure 1 presents the PRISMA flow chart of literature screening. A total 27 publications<sup>7-24-</sup>  
85 <sup>49</sup> involving 4386 subjects met our inclusion criteria (Table 1). One study<sup>28</sup> was divided into  
86 two independent parts because of the differences between the case group and the control group.

1  
2  
3  
4 87 Of the 27 eligible studies, 14 (52%) were published during 2010–2019, and 13 (48%) were  
5  
6 88 published before 2010. 12 (44%) studies were undertaken in the Europe, nine (30%) in the  
7  
8  
9 89 USA, two in the Asia (8%), three in the Middle East (12%) and one in the Australia (4%). Most  
10  
11 90 studies were prospective in design (56%, 15/27) versus 44% (12/27) retrospective.

12  
13  
14 91 All included studies used the tricuspid maximal regurgitation velocity (TRVmax) to estimate  
15  
16 92 SPAP; the majority of these studies using classical methods ( $4\text{TRVmax}^2 + \text{RAP}$ ) to calculate  
17  
18 93 SPAP. The right atrial pressure (RAP) was calculated through the diameter and collapse of the  
19  
20 94 inferior vena cava (IVC) during spontaneous respiration in 16(59%) studies, through the jugular  
21  
22 95 vein pressure (JVP) in one study (4%), and using a fixed value (5 or 10 mm Hg) in three studies  
23  
24 96 (11%). Three studies (11%) did not report their methods for calculating RAP. Four studies (15%)  
25  
26 97 used a tricuspid gradient ( $4\text{TRVmax}^2$ ) instead of SPAP.

27  
28  
29 98 The majority of studies (81%) reported the time interval (mean or maximum) between TTE  
30  
31 99 and RHC while five (19%) did not. 10 studies (37%) had time intervals greater than one week,  
32  
33 100 13 studies (48%) had time intervals of less than one week. The time interval between TTE and  
34  
35 101 RHC ranged from four hours to three months.

## 102 **Quality Assessment**

103 The quality assessment of included studies as per the QUADAS-2 inventory is presented in  
104 Figure 2. In 20 (74%) study protocols, consecutive subjects were enrolled, with no  
105 inappropriate exclusions. The risk of bias during patient recruitment was unclear in the  
106 remaining seven (26%) studies, as patient recruitment was not reported. In seven (26%) studies  
107 investigators designed the single-blind methods for TTE or RHC. Double blinding in imaging  
108 assessment was not mentioned in any of the articles of the studies. Risk of bias on flow and  
9

1  
2  
3  
4 109 timing between the index test and reference standard was categorized as unclear in 14 (52%)  
5  
6 110 study protocols that did not explicitly state successful investigation with both index and  
7  
8  
9 111 reference tests in all included patients. Overall, there were low concerns for patient selection,  
10  
11  
12 112 index test, and reference standard. Only five studies (19%) described the number of patients  
13  
14 113 excluded for lack of tricuspid regurgitation, while the rest (81%) only indicated excluding  
15  
16  
17 114 patients without tricuspid regurgitation or using contrast media to enhance tricuspid  
18  
19  
20 115 regurgitation signals.

### 116 **Quantitative Analysis**

117 The SROC curve for TTE is presented in Figure 3. The AUC was 0.88 (95%CI 0.85–0.90).  
118 The pooled sensitivity and specificity for TTE were 85% (95%CI 81%–90%) and 74% (95%CI  
119 64%–81%), respectively. There was significant heterogeneity across studies for the specificity  
120 and sensitivity of TTE (Figure 4). The pooled PLR and NLR were 3.2 (95%CI 2.3–4.4) and  
121 0.20 (95%CI 0.15–0.26). The pooled DOR for TTE is 16 (95%CI 10–27).

122 The results of the subgroup analysis are presented in Table 2. The sensitivity (87%, 95%CI  
123 81%–91%), specificity (74%, 95%CI 62%–83%) and AUC (0.89, 95%CI 0.86–0.91) of TTE  
124 for diagnosing PH was higher for studies published in 2010 and later relative to those published  
125 before 2010. Among the time interval subgroups, the group with the shortest time interval  
126 between TTE and RHC had the best diagnostic effect, with sensitivity specificity and AUC of  
127 88% (95%CI 73%–95%), 90% (95%CI 53%–99%) and 0.94(95%CI 0.92–0.96), respectively.  
128 The disease composition of the study population also affected the diagnostic accuracy of TTE.  
129 Compared with patients of other diseases, TTE had lower sensitivity (81%, 95%CI 70%–88%),  
130 specificity (61%, 95%CI 53%–69%) and AUC (0.73, 95%CI 0.69–0.77) in the subgroup of  
10

1  
2  
3  
4 131 patients with definite lung diseases.

5  
6 132 Subgroup analysis of different cut-off thresholds for diagnosing PH based on TTE showed  
7  
8  
9 133 that subgroup applied a cut-off threshold of 35mmHg had superior performance than at  
10  
11 134 40mmHg. The sensitivity, specificity and AUC of the former were respectively 92% (95%CI  
12  
13 135 88%–94%), 65% (95%CI 43%–83%) and 0.92 (95%CI 0.89–0.94), while the sensitivity,  
14  
15 136 specificity and AUC at 40mmHg were 84% (95%CI 75%–91%), 52% (95%CI 31%–71%) and  
16  
17 137 0.80 (95%CI 76%–83%). Tricuspid regurgitation pressure gradient (TRPG) calculated by  
18  
19 138  $4TRV_{max}^2$  has good specificity (81%, 95%CI 70%–89%) in the diagnosis of PH.

20  
21  
22  
23  
24  
25 139 The sensitivity analysis results are shown in Table 3. After excluding low-quality studies and  
26  
27 140 studies with specific characteristics, sensitivity analysis did not reveal a source for the  
28  
29 141 heterogeneity for the diagnostic accuracy analysis. Overall, the pooled meta-analysis results for  
30  
31 142 outcomes were robust to our sensitivity analyses.

### 32 33 34 35 143 **Discussion**

36  
37  
38 144 Our study found that TTE has better sensitivity but moderate specificity for the detection of  
39  
40 145 PH. The pooled sensitivity of the TTE was 0.85 at a specificity of 0.74. The AUC of TTE was  
41  
42 146 0.88. These findings altogether suggest that TTE has clinical value for diagnosing PH. In  
43  
44 147 addition, the time intervals between TTE and RHC, the cut-off threshold of TTE, the basic  
45  
46 148 diseases of the tested patients and other factors may affect the accuracy of TTE to diagnose PH.

47  
48  
49  
50 149 There was significant heterogeneity in our study, threshold test proves that threshold effect  
51  
52 150 is not the source of heterogeneity ( $r=0.34$ ,  $P=0.12$ ). From SROC diagram, we can see that only  
53  
54 151 four studies fall within the 95% confidence contour. Deeks' test for funnel plot asymmetry  
55  
56 152 suggested no publication bias ( $P=0.69$ ). We speculate the source of heterogeneity to a

1  
2  
3  
4 153 combination of variation, including study design, interval time between TTE and RHC, and  
5  
6 154 biases in the review, natural history of PH and population. Despite the significant heterogeneity  
7  
8  
9 155 of the studies included in this paper, our results can still serve as the adamant evidence for the  
10  
11  
12 156 diagnosis of TTE with PH.

13  
14 157 Guidelines recommend the use of inferior vena cava width and collapse rate to estimate RAP,  
15  
16 158 which was not used in some of the included studies. Sensitivity analysis for this point showed  
17  
18  
19 159 that studies of using IVC to calculate RAP did not seem to have a higher diagnostic performance.  
20  
21  
22 160 In order to avoid errors caused by RAP estimation, TRVmax is also considered as an indicator  
23  
24 161 to evaluate the possibility of PH. Four studies using TRPG replacing SPAP were grouped into  
25  
26  
27 162 a subgroup. The results showed that this subgroup had high diagnostic specificity, but the  
28  
29  
30 163 overall diagnostic effect was general.

31  
32 164 At last year's 6th World Symposium on Pulmonary Hypertension, the new definition of PH  
33  
34 165 surprised the audience and provoked lasting discussion. Scholars represented by Gérald  
35  
36 166 Simonneau<sup>50</sup> pointed out that the normal mean pulmonary artery pressure (MPAP) was  
37  
38 167  $14.0 \pm 3.3$  mmHg, and two standard deviations higher than the mean value would be  $MPAP > 20$   
39  
40  
41 168 mmHg should be defined as the upper limit of normal values (above 97.5 percentage points)  
42  
43  
44 169 based on scientific methods. However, some opponents<sup>51</sup> argue that it is too early to reduce the  
45  
46  
47 170 MPAP threshold to 20 mmHg because of the risk of over-diagnosis, unclear treatment  
48  
49  
50 171 implications and additional psychological burden on patients. Therefore, we still follow the  
51  
52  
53 172 previous standard that PH was defined as  $MPAP \geq 25$  mmHg at rest. After excluding 7 studies  
54  
55  
56 173 that RHC threshold was not 25 mmHg, sensitivity analysis results showed that sensitivity  
57  
58  
59 174 decreased slightly, specificity increased slightly. It should be noted that we must face up to the

1  
2  
3  
4 175 increased risk of misdiagnosis due to the reduction of the threshold value of PH diagnostic gold  
5  
6 176 standard. More research is needed to determine whether lowering standards does more harm  
7  
8  
9 177 than good.

10  
11 178 Studies included in previous systematic reviews and meta-analysis similar to this study all  
12  
13  
14 179 published before 2010, therefore we conducted subgroup analysis according to the year of  
15  
16  
17 180 publication. The data confirm that studies published after 2010 have a slightly higher diagnostic  
18  
19  
20 181 accuracy than previous studies. We surmised that this could be due to improvement of  
21  
22 182 ultrasonic equipment, the enhancement of ultrasonic operation or clear diagnostic gold standard  
23  
24  
25 183 of RHC.

26  
27 184 Does the time interval between the diagnostic test and the standard test affect the diagnostic  
28  
29  
30 185 accuracy? There is no clear consensus. Although PH is a chronic disease, we still believe that  
31  
32  
33 186 the shortest possible time interval is more conducive. Otherwise, changes in the patient's  
34  
35  
36 187 condition and the application of intervention measures will lead to an increase in the deviation  
37  
38 188 of the results of the two examinations. Since most studies (25/27) did not mention whether the  
39  
40  
41 189 subjects were hospitalized patients, we conducted a detailed subgroup analysis according to the  
42  
43 190 time interval between TTE and RHC. As expected, the diagnostic accuracy was highest when  
44  
45  
46 191 the time interval was less than or equal to 24 hours.

47  
48 192 Analysis restricted to subgroups disease composition of the population suggested that the  
49  
50  
51 193 diagnostic accuracy of TTE is lower for patients with lung diseases. Changes associated with  
52  
53  
54 194 chronic pulmonary disease, including marked increase in intrathoracic gas, consolidation of  
55  
56  
57 195 lung tissue, expansion of the thoracic cage, and alterations in the position of the heart, adversely  
58  
59 196 affect the imaging quality and parameters measurement of TTE<sup>52</sup>. Therefore, using TTE to

197 measure pulmonary pressure in patients with lung diseases may not be an ideal choice.

198 The cut-off threshold of the diagnostic test often affects its sensitivity and specificity.  
199 However, guidelines in Europe and the United States differ slightly in the threshold value for  
200 the echocardiographic diagnosis of PH. The American guideline advocates for a 40mmHg<sup>53</sup>,  
201 while the European equivalent recommends 36mmHg<sup>54</sup>. In our review, four studies assessed  
202 PH using TRPG (4TRVmax<sup>2</sup>) without adding RAP. In the remaining 22 studies, the cut-off  
203 thresholds of SPAP ranged from 30–50mmHg. Subgroup analysis showed that the diagnostic  
204 accuracy of the group of 35 mmHg was higher. Sensitivity analysis results of studies that  
205 excluded high TTE cut-off value showed that high cut-off value increased the specificity and  
206 reduced the sensitivity of TTE. Due to the small sample size of the subgroup in this study, the  
207 value of the cut-off threshold still needs to be determined by further prospective studies of  
208 multi-center and large samples.

209 There were two types of included literature, prospective studies and retrospective studies.  
210 Subgroup analysis presented that TTE had higher diagnostic value in prospective studies than  
211 in retrospective studies. A prospective study has predetermined details of experimental  
212 implementation, and the quality of research can be improved through multiple controllable links.  
213 It is more rigorous and scientific than a retrospective study<sup>55</sup>.

214 The previous three meta-analyses have their own limitations. The studies of Mohammed  
215 Taleb<sup>9</sup> and Rui Feng Zhang<sup>11</sup> included few references and the sample size was relatively small.  
216 Although Surinder Janda et al<sup>10</sup> included 28 studies, only 13 of them had sufficient data for  
217 meta-analysis. Most importantly, they only synthesized the diagnostic accuracy data in their  
218 study, but did not conduct subgroup analysis. In this study, we comprehensively included more



1  
2  
3  
4 219 literatures and conducted detailed subgroup analysis according to different research parameters  
5  
6 220 to explore the factors that may affect the diagnostic accuracy of TTE. It was expected that more  
7  
8  
9 221 researchers related to PH will benefit from the results of our detailed subgroup analysis.  
10

## 11 222 **Limitations**

12  
13  
14 223 Systematic review and meta-analysis is a secondary research method based on multiple  
15  
16  
17 224 original studies. The quality of the included studies will have an impact on the results of meta-  
18  
19  
20 225 analysis. In addition, there are several limitations in our study. Firstly, the studies included in  
21  
22 226 this review involved several different types of PH, and some studies did not describe the basic  
23  
24  
25 227 disease and PH typing in detail. It is obvious that pulmonary lesions can affect the quality of  
26  
27 228 TTE imaging, leading to underestimated results. Secondly, the blinding procedures were  
28  
29  
30 229 unreported for some of the included studies. Finally, in order to obtain more original studies,  
31  
32 230 we did not specify the type of echocardiography equipment and specific requirements in the  
33  
34  
35 231 operation process. Differences in equipment and ultrasonic inspection implementation process  
36  
37 232 could be a source of interference.  
38

## 39 233 **Conclusion**

40  
41  
42  
43 234 The value of TTE in diagnosing PH is certain, although it cannot yet replace RHC as the gold  
44  
45  
46 235 standard. The diagnostic accuracy of TTE improved with shortening time interval between TTE  
47  
48  
49 236 and RHC and appropriate cut-off threshold. TTE may not be suitable for assessing pulmonary  
50  
51 237 arterial pressure in patients with pulmonary disease. More multicenter, large-sample  
52  
53 238 randomized controlled trials are needed to verify our findings.  
54  
55  
56  
57  
58  
59  
60

### Author's contribution

The joint corresponding authors (JQ.L and B.S) are responsible for the design and implementation of the study. As a professional consultant, Y.H carried out quality control on the links of literature inclusion and data extraction. KH.Y provided guidance in literature retrieval and data processing methodology and was responsible for the quality evaluation part. JR.N and PJ.Y have done the systematic review of the literature and extracted data. JR.N has conducted the meta-analyses, and two authors (JR.N, PJ.Y) have substantially contributed to interpretation of data and co-authored the article. All the authors have made repeated revisions to the article. The corresponding authors (JQ.L and B.S) and JR.N take responsibility of the integrity of the analyses.

## Reference

1. Hoepfer MM, Humbert M, Souza R, et al. A global view of pulmonary hypertension. *The Lancet Respiratory medicine* 2016;4:306-322.
2. Beard JR, Officer A, de Carvalho IA, et al. The World report on ageing and health: a policy framework for healthy ageing. *The Lancet* 2016;387:2145-2154.
3. Deano RC, Glassner-Kolmin C, Rubenfire M, et al. Referral of patients with pulmonary hypertension diagnoses to tertiary pulmonary hypertension centers: the multicenter RePHerral study. *JAMA Intern Med* 2013;173:887-893.
4. Galie N, Humbert M, Vachiery JL, et al. 2015 ESC/ERS Guidelines for the diagnosis and treatment of pulmonary hypertension. *Kardiol Pol* 2015;73:1127-1206.
5. Hoepfer MM, Lee SH, Voswinckel R, et al. Complications of right heart catheterization procedures in patients with pulmonary hypertension in experienced centers. *J Am Coll Cardiol* 2006;48:2546-2552.
6. Yock PG, Popp RL. Noninvasive estimation of right ventricular systolic pressure by Doppler ultrasound in patients with tricuspid regurgitation. *Circulation* 1984;70:657-662.
7. Greiner S, Jud A, Aurich M, et al. Reliability of noninvasive assessment of systolic pulmonary artery pressure by Doppler echocardiography compared to right heart catheterization: analysis in a large patient population. *J Am Heart Assoc* 2014;3.
8. Bossone E, D'Andrea A, D'Alto M, et al. Echocardiography in pulmonary arterial hypertension: from diagnosis to prognosis. *Journal of the American Society of Echocardiography : official publication of the American Society of Echocardiography*

- 2013;26:1-14.
9. Taleb M, Khuder S, Tinkel J, et al. The diagnostic accuracy of Doppler echocardiography in assessment of pulmonary artery systolic pressure: a meta-analysis. *Echocardiography* 2013;30:258-265.
  10. Janda S, Shahidi N, Gin K, et al. Diagnostic accuracy of echocardiography for pulmonary hypertension: a systematic review and meta-analysis. *Heart* 2011;97:612-622.
  11. Zhang RF, Zhou L, Ma GF, et al. Diagnostic value of transthoracic Doppler echocardiography in pulmonary hypertension: a meta-analysis. *American journal of hypertension* 2010;23:1261-1264.
  12. Haynes RB. Of studies, syntheses, synopses, summaries, and systems: the "5S" evolution of information services for evidence-based healthcare decisions. *Evidence-based medicine* 2006;11:162-164.
  13. Tian J, Zhang J, Ge L, et al. The methodological and reporting quality of systematic reviews from China and the USA are similar. *Journal of clinical epidemiology* 2017;85:50-58.
  14. Xiu-xia L, Ya Z, Yao-long C, et al. The reporting characteristics and methodological quality of Cochrane reviews about health policy research. *Health policy (Amsterdam, Netherlands)* 2015;119:503-510.
  15. Moher D, Liberati A, Tetzlaff J, et al. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ (Clinical research ed)* 2009;339:b2535.

16. Shamseer L, Moher D, Clarke M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. *BMJ (Clinical research ed)* 2015;350:g7647.
17. Wang X, Chen Y, Yao L, et al. Reporting of declarations and conflicts of interest in WHO guidelines can be further improved. *Journal of clinical epidemiology* 2018;98:1-8.
18. Ge L, Tian JH, Li YN, et al. Association between prospective registration and overall reporting and methodological quality of systematic reviews: a meta-epidemiological study. *Journal of clinical epidemiology* 2018;93:45-55.
19. Whiting P, Rutjes AW, Dinnes J, et al. Development and validation of methods for assessing the quality of diagnostic accuracy studies. *Health technology assessment (Winchester, England)* 2004;8(25):iii, 1-234.
20. Whiting P, Rutjes AW, Reitsma JB, et al. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC medical research methodology* 2003;3:25.
21. Lee J, Kim KW, Choi SH, et al. Systematic Review and Meta-Analysis of Studies Evaluating Diagnostic Test Accuracy: A Practical Review for Clinical Researchers- Part II. Statistical Methods of Meta-Analysis. *Korean journal of radiology* 2015;16:1188-1196.
22. Deeks JJ, Macaskill P, Irwig L. The performance of tests of publication bias and other sample size effects in systematic reviews of diagnostic test accuracy was assessed. *Journal of clinical epidemiology* 2005;58:882-893.

- 1  
2  
3  
4 23. Reitsma JB, Glas AS, Rutjes AW, et al. Bivariate analysis of sensitivity and specificity  
5 produces informative summary measures in diagnostic reviews. *Journal of clinical*  
6  
7  
8  
9  
10  
11  
12 24. Ahmed M, Elshinawy O, Agmy G, et al. Tricuspid regurgitation velocity versus right  
13  
14  
15  
16  
17  
18  
19  
20 25. Shujaat A, Bajwa AA, Al-Saffar F, et al. Diagnostic accuracy of echocardiography  
21  
22  
23  
24  
25  
26  
27  
28 26. Schneider M, Pistritto AM, Gerges C, et al. Multi-view approach for the diagnosis of  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43 27. Keir GJ, Wort SJ, Kokosi M, et al. Pulmonary hypertension in interstitial lung disease:  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60
- 20
- 13.

- 1  
2  
3  
4 30. Balci MK, Arl E, Vayvada M, et al. Assessment of Pulmonary Hypertension in Lung  
5  
6 Transplantation Candidates: Correlation of Doppler Echocardiography With Right  
7  
8 Heart Catheterization. *Transplantation Proceedings* 2016;48:2797-2802.  
9  
10  
11 31. Nagel C, Henn P, Ehlken N, et al. Stress Doppler echocardiography for early detection  
12  
13 of systemic sclerosis-associated pulmonary arterial hypertension. *Arthritis research &*  
14  
15 *therapy* 2015;17:165.  
16  
17 32. Raevens S, Colle I, Reyntjens K, et al. Echocardiography for the detection of  
18  
19 portopulmonary hypertension in liver transplant candidates: An analysis of cutoff  
20  
21 values. *Liver Transplantation* 2013;19:602-610.  
22  
23  
24 33. Lange TJ, Baumgartner S, Arzt M, et al. Qualitative echocardiography parameters for  
25  
26 prediction of pulmonary hypertension. *International Journal of Clinical Practice*  
27  
28 2013;67:5-12.  
29  
30  
31 34. Lafitte S, Pillois X, Reant P, et al. Estimation of pulmonary pressures and diagnosis  
32  
33 of pulmonary hypertension by doppler echocardiography: A retrospective comparison  
34  
35 of routine echocardiography and invasive hemodynamics. *Journal of the American*  
36  
37 *Society of Echocardiography* 2013;26:457-463.  
38  
39  
40 35. Rajaram S, Swift AJ, Capener D, et al. Comparison of the diagnostic utility of cardiac  
41  
42 magnetic resonance imaging, computed tomography, and echocardiography in  
43  
44 assessment of suspected pulmonary arterial hypertension in patients with connective  
45  
46 tissue disease. *Journal of Rheumatology* 2012;39:1265-1274.  
47  
48  
49 36. Parsaee M, Amin A, Nematollahi MR, et al. Comparison of transthoracic  
50  
51 echocardiography and right heart catheterization for assessing pulmonary arterial  
52  
53  
54  
55  
56  
57  
58  
59  
60

- 1  
2  
3  
4 pressure in patients with congenital or valvular heart defects. *Iranian Heart Journal*  
5  
6 2012;12:54-61.  
7  
8  
9 37. Hua R, Sun YW, Wu ZY, et al. Role of 2-dimensional Doppler echo-cardiography in  
10  
11 screening portopulmonary hypertension in portal hypertension patients. *Hepatobiliary*  
12  
13 & pancreatic diseases international : HBDP INT 2009;8:157-61.  
14  
15  
16  
17 38. Nathan SD, Shlobin OA, Barnett SD, et al. Right ventricular systolic pressure by  
18  
19 echocardiography as a predictor of pulmonary hypertension in idiopathic pulmonary  
20  
21 fibrosis. *Respiratory Medicine* 2008;102:1305-1310.  
22  
23  
24  
25 39. Mogollon Jimenez MV, Escosca Ortega AM, Cabeza Letran ML, et al. Correlation  
26  
27 of Echocardiographic and Hemodynamic Parameters in Pulmonary Hypertension  
28  
29 Assessment Prior to Heart Transplantation. *Transplantation Proceedings*  
30  
31 2008;40:3023-3024.  
32  
33  
34  
35 40. Hsu VM, Moreyra AE, Wilson AC, et al. Assessment of pulmonary arterial hypertension  
36  
37 in patients with systemic sclerosis: comparison of noninvasive tests with results of  
38  
39 right-heart catheterization. *The Journal of rheumatology* 2008;35:458-465.  
40  
41  
42  
43 41. Fisher MR, Criner GJ, Fishman AP, et al. Estimating pulmonary artery pressures by  
44  
45 echocardiography in patients with emphysema. *European Respiratory Journal*  
46  
47 2007;30:914-21.  
48  
49  
50  
51 42. Lanzarini L, Fontana A, Campana C, et al. Two simple echo-doppler measurements  
52  
53 can accurately identify pulmonary hypertension in the large majority of patients with  
54  
55 chronic heart failure. *Journal of Heart and Lung Transplantation* 2005;24:745-754.  
56  
57  
58  
59 43. Mukerjee D, St George D, Knight C, et al. Echocardiography and pulmonary function  
60

22



- 1  
2  
3  
4 as screening tests for pulmonary arterial hypertension in systemic sclerosis.  
5  
6  
7 *Rheumatology (Oxford)* 2004;43:461-466.  
8  
9  
10 44. Arcasoy SM, Christie JD, Ferrari VA, et al. Echocardiographic assessment of pulmonary  
11  
12 hypertension in patients with advanced lung disease. *American Journal of Respiratory*  
13  
14 *and Critical Care Medicine* 2003;167:735-740.  
15  
16  
17 45. Penning S, Robinson KD, Major CA, et al. A comparison of echocardiography and  
18  
19 pulmonary artery catheterization for evaluation of pulmonary artery pressures in  
20  
21 pregnant patients with suspected pulmonary hypertension. *American Journal of*  
22  
23 *Obstetrics and Gynecology* 2001;184:1568-1570.  
24  
25  
26  
27 46. Matsuyama W, Ohkubo R, Michizono K, et al. Usefulness of transcutaneous Doppler  
28  
29 jugular venous echo to predict pulmonary hypertension in COPD patients. *Eur Respir*  
30  
31 *J* 2001;17:1128-1131.  
32  
33  
34  
35 47. Kim WR, Krowka MJ, Plevak DJ, et al. Accuracy of doppler echocardiography in the  
36  
37 assessment of pulmonary hypertension in liver transplant candidates. *Liver*  
38  
39 *Transplantation* 2000;6:453-458.  
40  
41  
42  
43 48. Denton CP, Cailles JB, Phillips GD, et al. Comparison of Doppler echocardiography  
44  
45 and right heart catheterization to assess pulmonary hypertension in systemic sclerosis.  
46  
47 *Br J Rheumatol* 1997;36:239-243.  
48  
49  
50  
51 49. Laaban JP, Diebold B, Zelinski R, et al. Noninvasive Estimation of Systolic  
52  
53 Pulmonary-Artery Pressure Using Doppler Echocardiography in Patients with Chronic  
54  
55 Obstructive Pulmonary-Disease. *Chest* 1989;96:1258-1262.  
56  
57  
58  
59  
60

- 1  
2  
3  
4 50. Simonneau G, Montani D, Celermajer DS, et al. Haemodynamic definitions and  
5  
6 updated clinical classification of pulmonary hypertension. *Eur Respir J* 2019;53.  
7  
8  
9 51. Gibbs JSR, Torbicki A. Proposed new pulmonary hypertension definition: is 4 mm(Hg)  
10  
11 worth re-writing medical textbooks? *Eur Respir J* 2019;53.  
12  
13  
14 52. Galderisi M, Cosyns B, Edvardsen T, et al. Standardization of adult transthoracic  
15  
16 echocardiography reporting in agreement with recent chamber quantification, diastolic  
17  
18 function, and heart valve disease recommendations: an expert consensus document of  
19  
20 the European Association of Cardiovascular Imaging. *Eur Heart J Cardiovasc Imaging*  
21  
22 2017;18:1301-1310.  
23  
24  
25 53. McLaughlin VV, Archer SL, Badesch DB, et al. ACCF/AHA 2009 expert consensus  
26  
27 document on pulmonary hypertension: a report of the American College of Cardiology  
28  
29 Foundation Task Force on Expert Consensus Documents and the American Heart  
30  
31 Association: developed in collaboration with the American College of Chest Physicians,  
32  
33 American Thoracic Society, Inc., and the Pulmonary Hypertension Association.  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43 54. Galie N, Hoeper MM, Humbert M, et al. Guidelines for the diagnosis and treatment of  
44  
45 pulmonary hypertension. *Eur Respir J* 2009;34:1219-1263.  
46  
47  
48 55. Ukai T, Shikata S, Nakayama T, et al. A comparison of the results of prospective and  
49  
50 retrospective cohort studies in the field of digestive surgery. *Surgery today*  
51  
52 2017;47:789-794.  
53  
54  
55  
56  
57  
58  
59  
60

Table 1 Characteristics of Studies Included in this Meta-analysis

Study	Year	Country	Design	N	Disease Composition of the Population	Mean Age (Years)	Male (%)	Time Interval	TTE Threshold (mmHg)	RHC Threshold (mmHg)	TTE Method
Ahmed	2019	USA	Retrospective	136	Multiple diseases	59±20	35	<3m	SPAP≥40	MPAP≥25	4TRVmax <sup>2</sup> +RAP(IVC)
Keir	2018	Australia	Prospective	265	Interstitial lung disease	60.8±10	46	–	TRPG>46	MPAP≥25	4TRVmax <sup>2</sup>
Habash- 1	2018	USA	Retrospective	31	Liver transplantation candidates	57±11	42	36.8±13.4d	SPAP>47	MPAP≥25	4TRVmax <sup>2</sup> +RAP(IVC)
Habash-2	2018	USA	Retrospective	49	Multiple diseases	59±15	31	16.0±11.6d	SPA>43	MPAP≥25	4TRVmax <sup>2</sup> +RAP(IVC)
Schneider	2018	Austria	Prospective	65	Cardiac and lung diseases	67.2	43	<48h	TRPG>32	MPAP≥25	4TRVmax <sup>2</sup>
Balci	2016	Turkey	Prospective	103	Lung transplantation candidates	47.6±10	66	<72h	SPAP>35	MPAP≥25	4TRVmax <sup>2</sup> +RAP (NR)
Shujaat	2016	USA	Retrospective	87	Multiple diseases	54.3±10	29	13d*	SPAP>40	MPAP>25	4TRVmax <sup>2</sup> +RAP (NR)
Sohrabi	2016	Iran	Prospective	300	Rheumatic mitral stenosis	59.9	31	<24h	SPAP≥35	MPAP≥25	4TRVmax <sup>2</sup> +RAP(IVC)
Nagel	2015	Germany	Prospective	76	Systemic sclerosis	58±14	16	–	SPAP>40	MPAP≥25	4TRVmax <sup>2</sup> +RAP(IVC)
Greiner	2014	Germany	Retrospective	1695	Cardiac disease	63±15	67	<5d	SPAP≥36	MPAP≥25	4TRVmax <sup>2</sup> +RAP(IVC)
Lafitte	2013	France	Retrospective	114	Cardiac and lung disease	64.8±10	52	<48h	SPAP≥38	MPAP>25	4TRVmax <sup>2</sup> +RAP(IVC)
Lange	2013	Germany	Retrospective	231	Multiple diseases	62±13	43	5±4d	SPAP>50	MPAP≥25	4TRVmax <sup>2</sup> +RAP (5)
Raevens	2013	Belgium	Retrospective	152	Liver transplantation candidates	58±11	66	–	SPAP>38	MPAP≥25	4TRVmax <sup>2</sup> +RAP (NR)
Parsaee	2012	Iran	Prospective	103	Cardiac diseases	41.0±10	44	<4h	SPAP≥35	MPAP>25	4TRVmax <sup>2</sup> +RAP(IVC)
Rajaram	2012	UK	Retrospective	81	Connective tissue disease	62±14	15	<48h	TRPG≥40	MPAP≥25	4TRVmax <sup>2</sup>
Hua	2009	China	Prospective	105	Liver transplantation candidates	49.5±10	63	4.2±2.0 d	SPAP≥30	MPAP≥25	4TRVmax <sup>2</sup> +RAP(IVC)
Nathan	2008	USA	Retrospective	60	Idiopathic pulmonary fibrosis	62.9±8	55	32±78d	SPAP≥40	MPAP>25	4TRVmax <sup>2</sup> +RAP(IVC)
Hsu	2008	USA	Prospective	49	Systemic Sclerosis	55	18	<4h	SPAP>47	MPAP≥25	4TRVmax <sup>2</sup> +RAP (10)
Mogollon	2008	Spain.	Retrospective	67	Heart transplantation candidates	–	–	–	SPAP>40	MPAP>35	4TRVmax <sup>2</sup> +RAP(IVC)
Fisher	2007	USA	Retrospective	63	Emphysema patients	65.6±6	60	23d	SPAP>40	MPAP≥25	4TRVmax <sup>2</sup> +RAP(IVC)
Lanzarini	2005	Italy	Prospective	57	Heart failure	52±11	74	<24h	SPAP≥32	SPAP≥35	4TRVmax <sup>2</sup> +RAP (IVC)
Mukerjee	2004	UK	Prospective	137	Systemic sclerosis	63	–	<3 m	TRPG>40	MPAP≥25	4TRVmax <sup>2</sup>
Arcasoy	2003	USA	Prospective	166	COPD 68%, ILD 28%, PVD 4%	51	43	<72h	SPAP≥45	SPAP≥45	4TRVmax <sup>2</sup> +RAP(IVC)
Penning	2001	USA	Retrospective	27	Pregnant women with cardiac diseases	28.6	0	25.8d	SPAP≥40	SPAP≥35	4TRVmax <sup>2</sup> +RAP(IVC)
Matsuyama	2001	Japan	Prospective	35	COPD	66	94	–	SPAP≥40	MPAP>25	4TRVmax <sup>2</sup> +RAP(IVC)
Kim	2000	USA	Prospective	74	Liver transplantation candidates	54	50	59d	SPAP>50	MPAP≥35	4TRVmax <sup>2</sup> +RAP(IVC)
Denton	1997	UK	Prospective	20	COPD	48.6±10	30	1.8±2.3m	SPAP≥30	SPAP≥30	4TRVmax <sup>2</sup> +RAP(JVP)
Laaban	1989	France	Prospective	27	COPD	63±9	78	<2d	SPAP≥35	SPAP≥35	4TRVmax <sup>2</sup> +RAP(5)

TTE, Transthoracic echocardiography; RHC, right cardiac catheter; SPAP, systolic pulmonary artery pressure; MPAP, mean pulmonary artery pressure; TRPG, tricuspid regurgitation pressure gradient; TRV, tricuspid regurgitation velocity; RAP, right atrial pressure; IVC, Inferior vena cava; JVP, jugular vein pressure; COPD, chronic obstructive pulmonary disease; ILD, Interstitial lung disease; PVD, peripheral vascular disease.

\* The median time (other terms are mean time).

**Table 2 Subgroup analysis**

Group	N	<i>I</i> <sup>2</sup> (95%CI)	AUC (95%CI)	Sensitivity (95%CI)	Specificity (95%CI)	PLR (95%CI)	NLR (95%CI)	DOR (95%CI)
<b>All Studies</b>	28	98(97–99)	0.88(0.85–0.90)	0.85(0.81–0.90)	0.74(0.64–0.81)	3.2(2.3–4.4)	0.20(0.15–0.26)	16(10–27)
<b><i>Time Interval</i></b>								
≤24h	4	95(90–99)	0.94(0.92–0.96)	0.88(0.73–0.95)	0.90(0.53–0.99)	8.9(1.5–54.5)	0.13(0.06–0.29)	68(13–348)
≤48h*	7	95(90–99)	0.94(0.91–0.95)	0.88(0.81–0.93)	0.89(0.71–0.96)	7.8(2.8–21.3)	0.13(0.09–0.21)	59(23–148)
≤72h*	9	94(89–99)	0.91(0.89–0.93)	0.87(0.82–0.91)	0.83(0.65–0.93)	5.2(2.4–11.2)	0.15(0.11–0.21)	34(14–82)
≤1 week	13	93(87–99)	0.91(0.88–0.93)	0.87(0.84–0.90)	0.80(0.68–0.88)	4.3(2.7–6.9)	0.16(0.12–0.21)	27(15–48)
>1 week	10	97(95–99)	0.82(0.78–0.85)	0.85(0.73–0.92)	0.60(0.40–0.77)	2.1(1.3–3.4)	0.25(0.14–0.45)	9(4–21)
Unclear	5	82(63–100)	0.85(0.81–0.88)	0.79(0.63–0.99)	0.76(0.61–0.87)	3.4(1.9–5.9)	0.27(0.15–0.51)	12(5–33)
<b><i>Population Disease</i></b>								
cardiac diseases	6	94(89–99)	0.90(0.87–0.92)	0.90(0.86–0.93)	0.67(0.29–0.91)	2.7(0.9–8.1)	0.15(0.08–0.30)	18(3–95)
lung diseases	8	90(81–100)	0.73(0.69–0.77)	0.81(0.70–0.88)	0.61(0.53–0.69)	2.1(1.8–2.4)	0.32(0.21–0.48)	7(4–10)
multiple diseases <sup>#</sup>	6	93(87–99)	0.90(0.87–0.92)	0.89(0.84–0.92)	0.70(0.40–0.89)	3.0(1.3–7.1)	0.16(0.11–0.23)	19(6–60)
Unclear <sup>&amp;</sup>	8	88(77–100)	0.88(0.85–0.90)	0.80(0.64–0.90)	0.85(0.80–0.89)	5.3(4.0–7.0)	0.23(0.12–0.45)	23(10–51)
<b><i>Published Year</i></b>								
≥2010	15	97(95–99)	0.89(0.86–0.91)	0.87(0.81–0.91)	0.74(0.62–0.83)	3.3(2.3–4.9)	0.18(0.13–0.25)	19(11–13)
<2010	13	96(93–99)	0.86(0.83–0.89)	0.84(0.74–0.90)	0.73(0.56–0.85)	3.1(1.8–5.3)	0.22(0.14–0.37)	14(6–33)
<b><i>Study Design</i></b>								
Prospective	15	97(95–99)	0.90(0.87–0.92)	0.86(0.77–0.91)	0.79(0.69–0.87)	4.2(2.7–6.4)	0.18(0.11–0.28)	23(12–45)
Retrospective	13	96(92–99)	0.86(0.83–0.89)	0.86(0.80–0.90)	0.65(0.49–0.78)	2.5(1.6–3.7)	0.22(0.15–0.32)	11(6–22)
<b><i>Cut-off Threshold</i></b>								

SPAP $\geq$ 40 mmHg	8	96(93–99)	0.80(0.76–0.83)	0.84(0.75–0.91)	0.52(0.31–0.71)	1.7(1.2–2.5)	0.30(0.21–0.44)	6(3–11)
SPAP $\geq$ 35 mmHg	4	76(47–100 )	0.92(0.89–0.94 )	0.92(0.88–0.94)	0.65(0.43–0.83)	2.6(1.4–4.9)	0.13(0.08–0.22)	16(9–28)
TRPG	4	0(0–100)	0.85(0.82–0.88)	0.75(0.58–0.86)	0.81(0.70–0.89)	4.0(2.2–7.3)	0.31(0.17–0.57)	13(4–40)

AUC, Area under curve; PLR, positive likelihood ratio; NLR, negative likelihood ratio; DOR, diagnostic odds ratio; TRPG, tricuspid regurgitation pressure gradient.

\*Studies with time intervals less than or equal to 24 hours were included in this subgroup.

\* Studies with time intervals less than or equal to 24 hours and 48 hours were included in this subgroup.

# Studies included a variety of diseases, including heart disease and lung disease.

&Diseases were not specifically identified in the studies (transplant candidates).

**Table 3 Sensitivity analysis**

Study characteristic	N	<i>I</i> <sup>2</sup>	AUC	Sensitivity	Specificity	PLR	NLR	DOR
		(95%CI)	(95%CI)	(95%CI)	(95%CI)	(95%CI)	((95%CI)	(95%CI)
all included studies	28	98(97–99)	0.88(0.85–0.90)	0.85(0.81–0.90)	0.74(0.64–0.81)	3.2(2.3–4.4)	0.20(0.15–0.26)	16(10–27)
excluding study of Penning	27	98(97–99)	0.88(0.85–0.91)	0.86(0.81–0.89)	0.75(0.66–0.82)	3.4(2.5–4.6)	0.19(0.14–0.26)	18(11–28)
MPAP $\geq$ 25	21	98(97–99)	0.87(0.84–0.90)	0.83(0.77–0.88)	0.76(0.67–0.83)	3.5(2.5–4.8)	0.22(0.16–0.30)	16(10–26)
RAP method(IVC)	17	96(93–99)	0.89(0.86–0.91)	0.86(0.82–0.90)	0.73(0.59–0.84)	3.2(2.0–5.1)	0.19(0.13–0.27)	17(8–35)
Excluding high cut-off value*	21	97(95–99)	0.90(0.87–0.92)	0.88(0.85–0.91)	0.72(0.59–0.82)	3.2(2.1–4.8)	0.16(0.12–0.22)	20(11–36)

AUC, Area under curve; PLR, positive likelihood ratio; NLR, negative likelihood ratio; DOR, diagnostic odds ratio; SPAP, systolic pulmonary artery pressure; MPAP, mean pulmonary artery pressure; RAP, right atrial pressure; IVC, Inferior vena cava.

\* High cut-off value was defined as SPAP greater than 45mmHg or tricuspid regurgitation pressure gradient (TRPG) greater than 40mmHg.

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

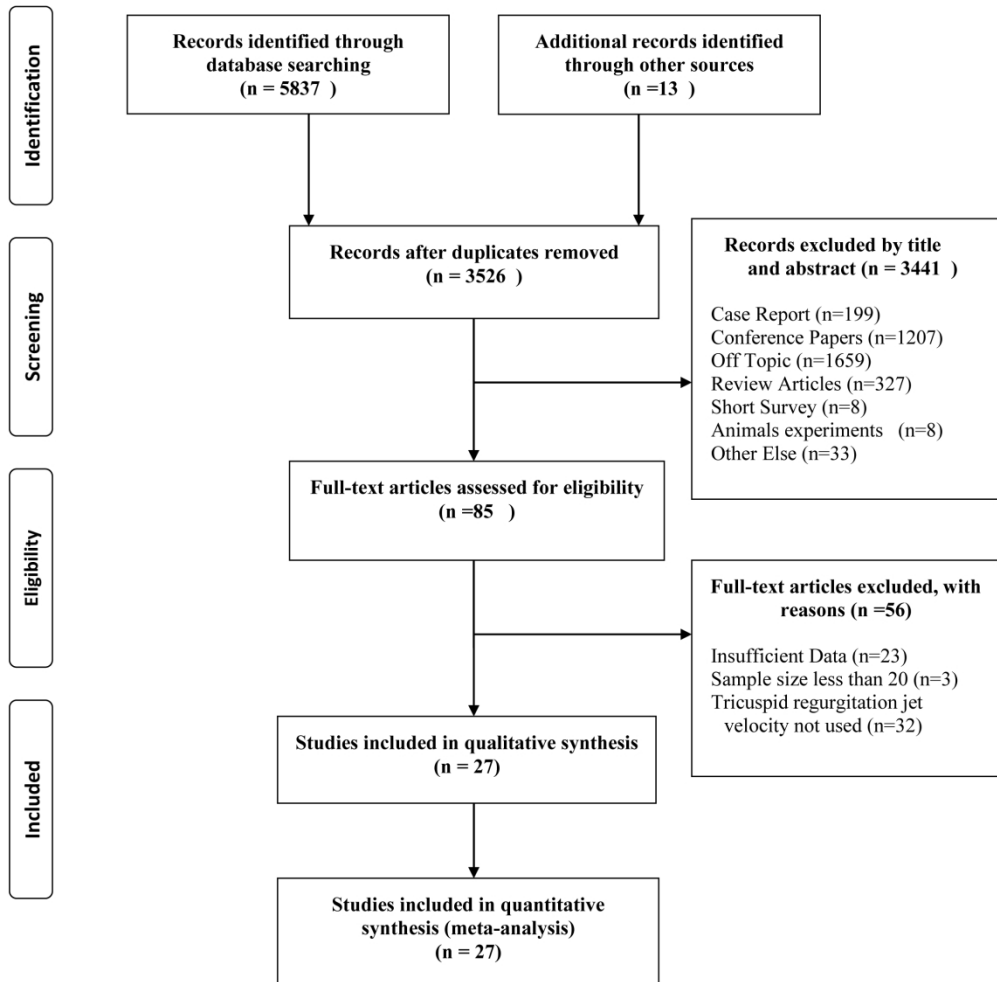


Figure 1 Flowchart for identification of studies

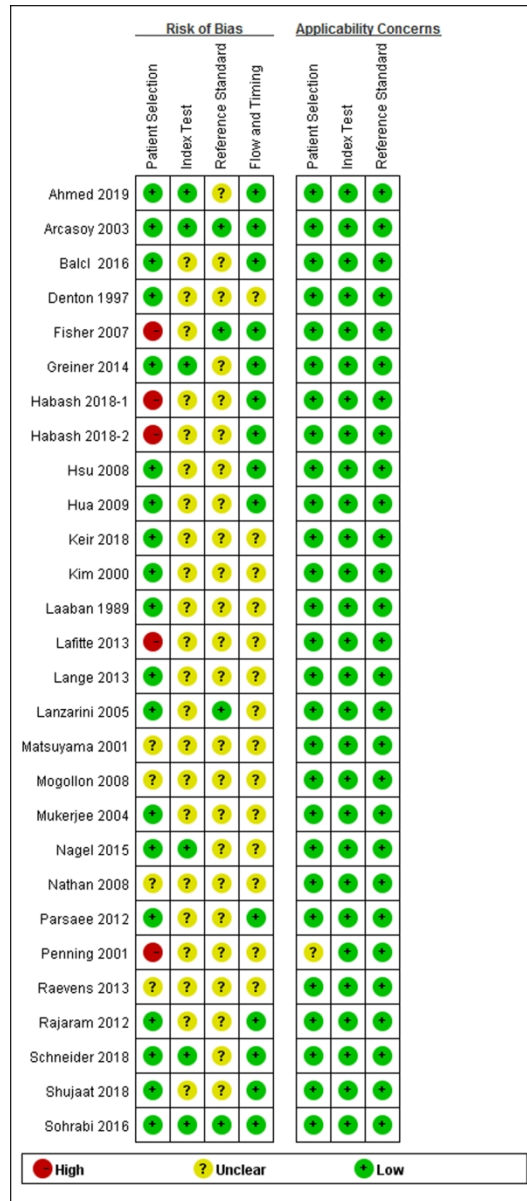


Figure 2 Risk of bias and applicability concerns summary: review authors' judgements regarding each domain for each included study (n=28).



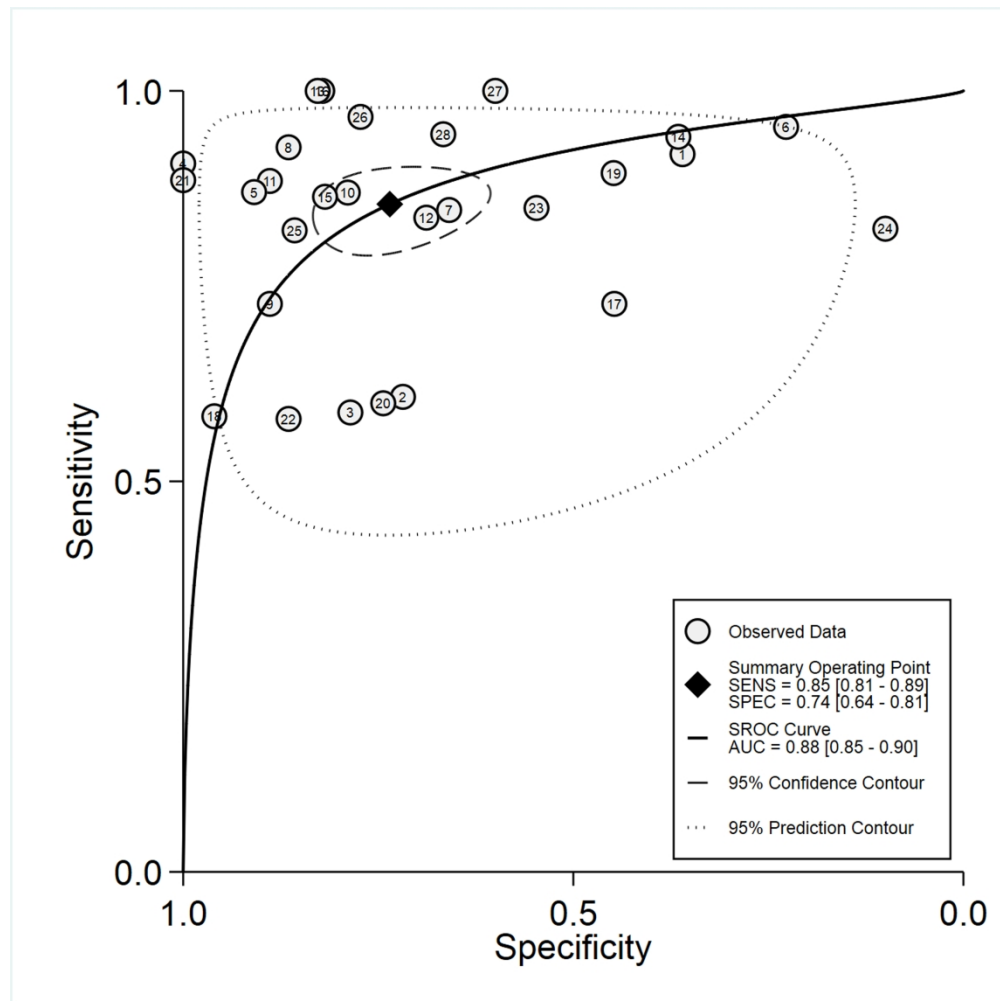


Figure 3 Summary receiver operating characteristic (SROC) graph with 95% confidence region and 95% prediction region for TTE in the diagnosis of pulmonary hypertension (n=28).

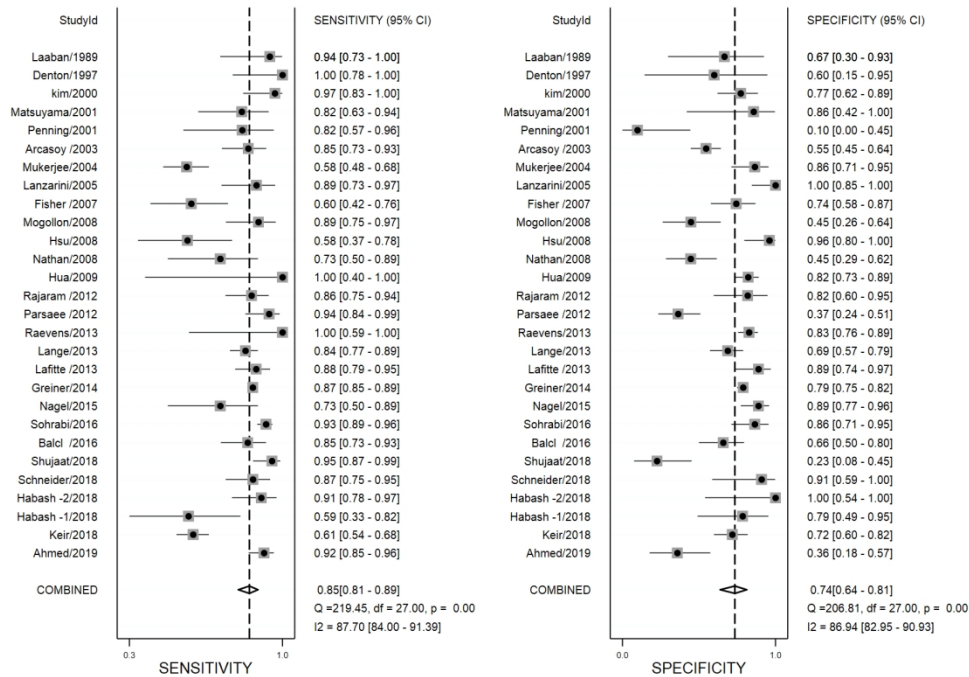


Figure 4 Forest plot of the sensitivity and specificity of each individual study, summary sensitivity and specificity and I2 statistic for heterogeneity (n=28).



# PRISMA-DTA Checklist

Section/topic	#	PRISMA-DTA Checklist Item	Reported on page #
<b>TITLE / ABSTRACT</b>			
Title	1	Identify the report as a systematic review (+/- meta-analysis) of diagnostic test accuracy (DTA) studies.	1
Abstract	2	Abstract: See PRISMA-DTA for abstracts.	3
<b>INTRODUCTION</b>			
Rationale	3	Describe the rationale for the review in the context of what is already known.	6
Clinical role of index test	D1	State the scientific and clinical background, including the intended use and clinical role of the index test, and if applicable, the rationale for minimally acceptable test accuracy (or minimum difference in accuracy for comparative design).	6
Objectives	4	Provide an explicit statement of question(s) being addressed in terms of participants, index test(s), and target condition(s).	7
<b>METHODS</b>			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	7
Eligibility criteria	6	Specify study characteristics (participants, setting, index test(s), reference standard(s), target condition(s), and study design) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	7
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	7
Search	8	Present full search strategies for all electronic databases and other sources searched, including any limits used, such that they could be repeated.	Supplementary Data
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	30, Figure 1
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	8
Definitions for data extraction	11	Provide definitions used in data extraction and classifications of target condition(s), index test(s), reference standard(s) and other characteristics (e.g. study design, clinical setting).	8
Risk of bias and applicability	12	Describe methods used for assessing risk of bias in individual studies and concerns regarding the applicability to the review question.	9
Diagnostic accuracy measures	13	State the principal diagnostic accuracy measure(s) reported (e.g. sensitivity, specificity) and state the unit of assessment (e.g. per-patient, per-lesion).	9
Synthesis of results	14	Describe methods of handling data, combining results of studies and describing variability between studies. This could include, but is not limited to: a) handling of multiple definitions of target condition. b) handling of multiple thresholds of test positivity. c) handling multiple index test readers. d) handling of indeterminate test results. e)	9



# PRISMA-DTA Checklist

grouping and comparing tests, f) handling of different reference standards

Page 1 of 2

Section/topic	#	PRISMA-DTA Checklist Item	Reported on page #
Meta-analysis	D2	Report the statistical methods used for meta-analyses, if performed.	9
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	9
<b>RESULTS</b>			
Study selection	17	Provide numbers of studies screened, assessed for eligibility, included in the review (and included in meta-analysis, if applicable) with reasons for exclusions at each stage, ideally with a flow diagram.	30, Figure 1
Study characteristics	18	For each included study provide citations and present key characteristics including: a) participant characteristics (presentation, prior testing), b) clinical setting, c) study design, d) target condition definition, e) index test, f) reference standard, g) sample size, h) funding sources	26 Table 1
Risk of bias and applicability	19	Present evaluation of risk of bias and concerns regarding applicability for each study.	31 Figure 2
Results of individual studies	20	For each analysis in each study (e.g. unique combination of index test, reference standard, and positivity threshold) report 2x2 data (TP, FP, FN, TN) with estimates of diagnostic accuracy and confidence intervals, ideally with a forest or receiver operator characteristic (ROC) plot.	32 Figure 3
Synthesis of results	21	Describe test accuracy, including variability; if meta-analysis was done, include results and confidence intervals.	33 Figure 4
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression); analysis of index test: failure rates, proportion of inconclusive results, adverse events).	27-29 Table 2-3
<b>DISCUSSION</b>			
Summary of evidence	24	Summarize the main findings including the strength of evidence.	12
Limitations	25	Discuss limitations from included studies (e.g. risk of bias and concerns regarding applicability) and from the review process (e.g. incomplete retrieval of identified research).	16
Conclusions	26	Provide a general interpretation of the results in the context of other evidence. Discuss implications for future research and clinical practice (e.g. the intended use and clinical role of the index test).	16
<b>FUNDING</b>			
Funding	27	For the systematic review, describe the sources of funding and other support and the role of the funders.	2



# PRISMA-DTA Checklist

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47

Accuracy Studies: The PRISMA-DTA Statement. JAMA. 2018 Jan 23;319(4):388-396. doi: 10.1001/jama.2017.19163.  
For more information, visit: [www.prisma-statement.org](http://www.prisma-statement.org).

Page 2 of 2

For peer review only

Open-2019-033084 on 22 December 2019. Downloaded from <http://bmjopen.bmj.com/> on April 17, 2024 by guest. Protected by copyright.

## 1. Pubmed

(((((((((((((((Echocardiography[MeSH Terms]) OR Echocardiography[Title/Abstract]) OR  
 ECHO[Title/Abstract]) OR echocardiogram\*[Title/Abstract]) OR ultrasound  
 cardiography[Title/Abstract]) OR heart echography[Title/Abstract]) OR cardiac  
 echography[Title/Abstract]) OR Cardiac ultrasound[Title/Abstract]) OR heart  
 ultrasound[Title/Abstract]) OR cardial echography[Title/Abstract]) OR cardio  
 echography[Title/Abstract]) OR echo cardiogram\*[Title/Abstract]) OR echo  
 cardiography[Title/Abstract]) OR echo sounding, heart[Title/Abstract]) OR heart echo  
 sounding[Title/Abstract])) AND (((((((((((Hypertension, Pulmonary/diagnosis[MeSH Terms])  
 OR Hypertension, Pulmonary[Title/Abstract]) OR pulmonary hypertension[Title/Abstract])  
 OR hypertension, lung[Title/Abstract]) OR hypertensive pulmonary[Title/Abstract]) OR  
 pulmonary arter\* hypertension[Title/Abstract]) OR lung arter\* hypertension[Title/Abstract])  
 OR lung hypertension[Title/Abstract]) OR pulmonary fixed hypertension[Title/Abstract]) OR  
 PH[Title/Abstract])) OR pulmonary hypertensive[Title/Abstract])) AND (((((((((((Right heart  
 catheterization[Title/Abstract]) OR Right heart catheterisation[Title/Abstract]) OR right heart  
 catheter\*[Title/Abstract]) OR right cardiac catheter\*[Title/Abstract]) OR right cardiac  
 catheterization[Title/Abstract]) OR right cardiac catheterisation[Title/Abstract]) OR  
 RHC[Title/Abstract])) OR right - sided heart catheterization[Title/Abstract])) OR right - sided  
 heart catheterisation[Title/Abstract]) Sort by: Best Match

## 2. Cochrane library

- 1  
2  
3  
4  
5  
6  
7 #1 MeSH descriptor: [Echocardiography] explode all trees  
8  
9 #2 Echocardiography  
10  
11 #3 echocardiogram\*  
12  
13  
14 #4 Echo  
15  
16 #5 Cardiac ultrasound  
17  
18 #6 heart ultrasound  
19  
20 #7 ultrasound cardiography  
21  
22  
23 #8 heart echography  
24  
25 #9 cardi\* echography  
26  
27 #10 echo cardiogra\*  
28  
29 #11 echo sounding, heart  
30  
31 #12 echography, heart  
32  
33 #13 heart echo sounding  
34  
35 #14 #1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7 OR #8 OR #9 OR #10 OR #11 OR #12  
36  
37 OR #13  
38  
39 #15 MeSH descriptor: [Hypertension, Pulmonary] explode all trees  
40  
41 #16 pulmonary hypertension  
42  
43 #17 hypertension, pulmonary  
44  
45 #18 lung hypertension  
46  
47 #19 hypertension, lung  
48  
49 #20 pulmonary arterial hypertension  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

- 1  
2  
3  
4 #21 lung arterial hypertension  
5  
6 #22 lung artery hypertension  
7  
8  
9 #23 pulmonary artery hypertension  
10  
11 #24 pulmonary fixed hypertension  
12  
13  
14 #25 pulmonary hypertensive  
15  
16  
17 #26 hypertensive pulmonary vascular disease\*  
18  
19  
20 #27 #15 OR #16 OR #17 OR 15 OR #19 #20 OR #21 OR #22 OR #23 OR #24 OR #25  
21  
22 OR #26  
23  
24  
25 #28 right cardiac catheterization  
26  
27 #29 Right heart catheterization  
28  
29  
30 #30 right heart catheter\*  
31  
32 #31 right cardiac catheter\*  
33  
34  
35 #32 right heart catheterisation  
36  
37 #33 right cardiac catheterisation  
38  
39  
40 #34 right - sided heart catheterization  
41  
42 #35 right - sided heart catheterisation  
43  
44  
45 #36 RHC  
46  
47  
48 #37 #28 OR #29 OR #30 #31 #32 OR #33 OR #34 OR #35 OR #36  
49  
50  
51 #38 #14 AND #27 AND #37  
52  
53  
54  
55  
56  
57  
58  
59  
60



## 3. EMBASE

('echocardiography'/exp OR 'echocardiography':ab OR 'ECHO':ab OR 'echocardiogram\*':ab OR 'ultrasound cardiography':ab OR 'heart echography':ab OR 'cardi\* echography':ab OR 'cardiac ultrasound':ab OR 'heart ultrasound':ab OR 'echo cardiogra\*':ab OR 'echo sounding, heart':ab OR 'echography, heart':ab OR 'heart echo sounding':ab) AND ('pulmonary hypertension'/exp OR 'pulmonary hypertension':ab OR 'hypertension, pulmonary':ab OR 'hypertension, lung':ab OR 'hypertensive pulmonary':ab OR 'pulmonary arter\* hypertension':ab OR 'lung arter\* hypertension':ab OR 'lung hypertension':ab OR 'pulmonary fixed hypertension':ab OR 'pulmonary hypertensive':ab OR 'PH':ab) AND ('right cardiac catheterisation':ab OR 'right cardiac catheterization':ab OR 'right heart catheterization':ab OR 'right heart catheter\*':ab OR 'right cardiac catheter\*':ab OR 'RHC':ab)

#### 4. Web of Science

#4 #3 AND #2 AND #1

DocType=All document types; Language=All languages;

#3 TS= ( “right cardiac catheterization” or “right heart catheterization” or “right heart catheter\*” or “right cardiac catheter\*” or “right heart catheterisation” or “right cardiac catheterisation” or “RHC” or “right - sided heart catheterization” or “right - sided heart catheterisation” )

DocType=All document types; Language=All languages;

#2 TS= ( “Pulmonary Hypertension” or “PH” or “lung hypertension” or “pulmonary arter\*hypertension” or “lung arter\* hypertension” or “pulmonary hypertensive” or “hypertensive pulmonary” )

DocType=All document types; Language=All languages;

#1 TS= ( echocardiography or ECHO or echocardiogram\* or “ultrasound cardiography” or “heart echography” or “Cardiac ultrasound” or “heart ultrasound” or “cardi\* echography” or “echo cardiogra\*” )

DocType=All document types; Language=All languages;

# BMJ Open

## Diagnostic Accuracy of Transthoracic Echocardiography for Pulmonary Hypertension: a Systematic Review and Meta-Analysis

Journal:	<i>BMJ Open</i>
Manuscript ID	bmjopen-2019-033084.R1
Article Type:	Original research
Date Submitted by the Author:	18-Oct-2019
Complete List of Authors:	Ni, Jin-Rong; The First Hospital of Lanzhou University Yan, Pei-Jing; Institute of Clinical Research and Evidence Based Medicine, Gansu Provincial Hospital Liu, Shi-Dong; The First Hospital of Lanzhou University Hu, Yuan; the First Hospital of Lanzhou University Yang, Kehu; evidence based medicine center, lanzhou university Song, Bing; The First Hospital of Lanzhou University Lei, Jun-Qiang; The First Hospital of Lanzhou University
<b>Primary Subject Heading</b>:	Diagnostics
Secondary Subject Heading:	Radiology and imaging, Qualitative research, Global health, Cardiovascular medicine
Keywords:	Hypertension < CARDIOLOGY, Echocardiography < CARDIOLOGY, Diagnostic radiology < RADIOLOGY & IMAGING

SCHOLARONE™  
Manuscripts

## Diagnostic Accuracy of Transthoracic Echocardiography for Pulmonary Hypertension: a Systematic Review and Meta-Analysis

Jin-Rong Ni<sup>1,3,8,9</sup>, Pei-Jing Yan<sup>2,5,6#</sup>, Shi-Dong Liu<sup>1,3</sup>, Yuan Hu<sup>3</sup>, Ke-Hu Yang<sup>2,4,5,6</sup>, Bing Song<sup>3\*</sup>, Jun-Qiang Lei<sup>1,7,8,9\*</sup>

<sup>1</sup>The First Hospital (the First Clinical Medical School) of Lanzhou University, Lanzhou 730000, China;

<sup>2</sup>Institute of Clinical Research and Evidence Based Medicine, Gansu Provincial Hospital, Lanzhou,

730000, China; <sup>3</sup>Department of Cardiovascular Surgery, the First Hospital of Lanzhou University,

Lanzhou 730000, China; <sup>4</sup>Evidence-Based Medicine Center, School of Basic Medical Sciences, Lanzhou

University, Lanzhou 730000, China; <sup>5</sup>Evidence-Based Social Science Research Center, Lanzhou

University, Lanzhou, 730000, China; <sup>6</sup>Key Laboratory of Evidence-based Medicine and Knowledge

Translation of Gansu Province, Lanzhou, 730000, China. <sup>7</sup>Department of Radiology, the First Hospital

of Lanzhou University, Lanzhou 730000, China. <sup>8</sup>Intelligent Imaging Medical Engineering Research

Center of Gansu province, Lanzhou 730000, China; <sup>9</sup>Precision Image and Collaborative Innovation

International Scientific and Technological Cooperation Base of Gansu province, Lanzhou 730000, China.

#Joint first author

\*Joint corresponding authors

**Corresponding Author:** 1. Jun-Qiang Lei, No.11 Donggang West Road, Chengguan District,

Lanzhou, China.

Tel: +8613919289040

E-mail: leijunqiangldyy@163.com

2. Bing Song, No.11 Donggang West Road, Chengguan District,

Lanzhou, China.

Tel: +8613993118707

E-mail: songbingldyyxwk@163.com

**Funding sources:** This study was supported by the Key Laboratory of Evidence Based Medicine and Knowledge Translation Foundation of Gansu Province (Grant No. GSXZYZH2018006).

**Statement on potential conflicts of interest:** None declared.

**Patient and Public Involvement:** There is no patient or the public involved in our work.

**Data Availability:** All data relevant to the study are included in the article.

**Key Words:** Echocardiography; Pulmonary Hypertension; Catheterization; Diagnosis; Accuracy; Meta-analysis.

## Abstract

**Objective:** To evaluate the diagnostic accuracy of transthoracic echocardiography (TTE) in patients with pulmonary hypertension (PH).

**Design:** Systematic review and meta-analysis.

**Data sources and eligibility criteria:** EMBASE, Cochrane Library for clinical trials, PubMed and Web of Science were searched from inception to June 19, 2019. Studies using both TTE and right heart catheterization (RHC) to diagnose PH were included.

**Main results:** A total of 27 studies involving 4386 subjects were considered as eligible for analysis. TTE had a pooled sensitivity of 85% (95% CI 81–90%), a pooled specificity of 74% (95% CI 64–81%), a pooled positive likelihood ratio of 3.2 (95% CI 2.3–4.4), a pooled negative likelihood ratio of 0.20 (95% CI 0.15–0.26), a pooled diagnostic odds ratio of 16 (95% CI 10–27), and finally an area under the summary receiver operating characteristic (SROC) curve of 0.88 (95%CI 0.85–0.90).

**Conclusion:** TTE has clinical value in diagnosing PH, although it cannot yet replace RHC considered as the gold standard. The accuracy of TTE may be improved by shortening the time interval between TTE and RHC and by developing an appropriate threshold. TTE may not be suitable to assess pulmonary arterial pressure in patients with pulmonary diseases.

**Review registration number:** PROSPERO CRD42019123289.

### Strengths and limitations of this study

1. We conducted a comprehensive search of the main database, included more studies, and obtained a large sample size.
2. Detailed subgroup analysis and sensitivity analysis were performed.
3. The types of pulmonary hypertension included in the studies could not be distinguished.
4. Significant heterogeneity in our study limits the interpretation of results.

## 1 Introduction

2 The prevalence of pulmonary hypertension (PH) is estimated at 1% in the general population,  
3 and as high as 10% in the 600 million people older than 65.<sup>1</sup> Early detection and accurate  
4 assessment are vital for improved outcomes for PH patients.<sup>2</sup> Right heart catheterization (RHC)  
5 is the gold standard in the diagnosis of PH,<sup>3</sup> but it is invasive and cannot be used frequently or  
6 repeatedly.<sup>4</sup> The latest guideline for PH recommends transthoracic echocardiography (TTE) as  
7 a noninvasive test for screening.<sup>3</sup>

8 High quality meta-analysis has been considered as one of the key tools for achieving  
9 evidence.<sup>5 6</sup> Three systematic reviews and meta-analysis regarding the diagnostic accuracy of  
10 TTE for PH were published between 2010 and 2013.<sup>7-9</sup> The studies included in these meta-  
11 analysis were all published before 2010. In addition, two of them included fewer studies and  
12 performed a simple diagnostic data synthesis.<sup>8 9</sup> The other included a relatively large number  
13 of studies, but did not assess a detailed subgroup analysis.<sup>7</sup> In recent years, TTE has still been  
14 used in the clinical diagnosis of PH, and many new original studies have been published.<sup>10-13</sup>  
15 Therefore, the purpose of our study was to undertake a comprehensive systematic review and  
16 quantitative meta-analysis on the accuracy of TTE in the diagnosis of PH.

## 17 Methods

18 The present study is reported according to the Preferred Reporting Items for Systematic  
19 Reviews and Meta-analyses (PRISMA) statement and the published recommendations.<sup>14 15</sup> The  
20 detailed protocol is accessible in PROSPERO (CRD42019123289).<sup>16 17</sup>

## 21 Data sources and search



1  
2  
3  
4 22 A systematic search in EMBASE, Cochrane Library for clinical trials, PubMed and Web of  
5  
6 23 Science was performed to find the relevant literature from inception to June 19, 2019. Subject  
7  
8  
9 24 words were combined with free words, and the search strategy was developed and adapted for  
10  
11 25 each database. For unpublished trials, ClinicalTrials.gov and the trials registers on the World  
12  
13 26 Health Organization International Clinical Trials Registry Platform were searched. The  
14  
15 27 references of the included studies and other systematic reviews and meta-analysis were also  
16  
17 28 reviewed to obtain a comprehensive list of included studies.  
18  
19  
20  
21

### 22 29 **Study selection**

23  
24  
25 30 Studies were selected based on the following inclusion criteria: studies that diagnosed PH by  
26  
27 31 TTE; the study population was represented by patients with suspected PH; TTE measurement  
28  
29 32 of systolic pulmonary artery pressure (SPAP) was performed using tricuspid regurgitation;  
30  
31 33 RHC was used as the gold standard for the diagnosis of PH.  
32  
33  
34

35 34 The exclusion criteria were the following: insufficient data to construct a 2×2 table; tricuspid  
36  
37 35 regurgitation method was not used to calculate pulmonary artery pressure; studies with less than  
38  
39 36 20 subjects; duplicate data were used (in this case, the largest sample or the latest study was  
40  
41 37 selected).  
42  
43  
44

45 38 Two reviewers (JR.N and PJ.Y) independently screened the eligible studies for suitability.  
46  
47 39 Disagreements were resolved by consensus. If consensus could not be reached, a third reviewer  
48  
49 40 (SD.L) was deferred to arbitration and consensus. No language restriction was applied. If a  
50  
51 41 study is not conducted in the author's language, professional translation software could be used.  
52  
53  
54  
55

### 56 42 **Data extraction**

Two reviewers (JR.N and PJ.Y) extracted the data independently according to a predefined data extraction sheet. The following variables were extracted from the included studies: lead author, publication year, country of study, study design, study population demographics, sample size, mean age, male ratio, time interval between TTE and RHC, the cut-off threshold levels for TTE and RHC, and number of true-positive (TP), false-negative (FN), true-negative (TN) and false-positive (FP) observations. Extracted data was cross-checked and disagreements were resolved via discussion or referral to a third reviewer (Y.H).

### Quality assessment

The Quality Assessment of Diagnostic Accuracy Studies QUADAS-2 tool was used to assess the risk of bias and clinical applicability concerns of the included studies according to the Cochrane Collaboration recommendation.<sup>18 19</sup> Two reviewers (JR.N and PJ.Y) independently evaluated QUADAS-2 items, and all emerging conflicts were resolved by consensus.

### Data synthesis and statistical analysis

Statistical analysis was performed using STATA/SE version 15.1 (Stata Corp, College Station, TX) and Review Manager Version 5.3 software (Copenhagen, Denmark, Nordic Cochrane Centre, Cochrane Collaboration, 2014). All tests were two-tailed. A  $p$  value < 0.05 was considered statistically significant.

The correlation coefficient between the logarithm of sensitivity and logarithm of one minus specificity was calculated to test whether the threshold effect was one of the sources of heterogeneity.<sup>20</sup> Deeks' test was used to test for publication bias.<sup>21</sup> The bivariate model for diagnostic meta-analysis was used to obtain pooled estimates of sensitivity and specificity.<sup>22</sup>

64 Statistical heterogeneity among studies was explored using the  $I^2$  statistic.

65 Pooled sensitivity, specificity, diagnostic odds ratio (DOR), positive likelihood ratio (PLR),  
66 negative likelihood ratio (NLR), and the area under the summary receiver operating  
67 characteristic (SROC) curve were calculated from the number of TPs, FNs, FPs, and TNs. The  
68 95% confidence interval (CI) was estimated for each metric.

69 Subgroup analyses were undertaken based on the following variables: the time interval  
70 between TTE and RHC; disease classification of the study population; publication year of the  
71 study; study design (prospective or retrospective) and cut-off threshold of TTE to diagnose PH.  
72 Sensitivity analysis was undertaken by excluding low-quality studies (according to the  
73 QUADAS-2 quality assessment) or trials with characteristics different from the others.

## 74 Results

### 75 Studies selection and characteristics

76 Figure 1 shows the PRISMA flow chart of the literature screening. A total of 27 articles  
77 involving 4386 subjects met our inclusion criteria (Table 1).<sup>10-13 23-45</sup> Habash's study was  
78 divided into two independent parts because of the differences between the case group (Habash-1)  
79 and the control group (Habash-2).<sup>27</sup>

80 Of the 27 eligible studies, fourteen (52%) were published during 2010–2019,<sup>10-13 26 27 30 33 34</sup>  
81 <sup>39 41 43-45</sup> and thirteen (48%) were published before 2010.<sup>23-25 28 29 31 32 35-38 40 42</sup> Twelve (44%)  
82 studies were performed in Europe,<sup>12 24 26 32-35 37-39 43 44</sup> nine (30%) in United States of America  
83 (USA),<sup>10 13 23 25 27 28 31 40 42</sup> two (8%) in East Asia,<sup>29 36</sup> three (12%) in Middle East,<sup>11 41 45</sup> and one  
84 (4%) in Australia.<sup>30</sup> Most of the studies (15/27, 56%)<sup>11 12 23 24 28-32 35 36 38 39 41 45</sup> were of prospective  
85 design versus 44% (12/27)<sup>10 13 25-27 33 34 37 40 42-44</sup> retrospective.

1  
2  
3  
4 86 All included studies used the tricuspid maximal regurgitation velocity (TRVmax) to estimate  
5  
6 87 SPAP; the majority of these studies (23/27, 85%) used classical method ( $4\text{TRVmax}^2 + \text{RAP}$ ) to  
7  
8  
9 88 calculate SPAP.<sup>10 11 13 23-28 31-37 39-45</sup> The right atrial pressure (RAP) was calculated through the  
10  
11 89 diameter and collapse of the inferior vena cava (IVC) during spontaneous respiration in  
12  
13  
14 90 sixteen (59%) studies,<sup>10 23 25-27 31 33 35-37 39-42 44 45</sup> through the jugular vein pressure (JVP) in one  
15  
16  
17 91 study (4%),<sup>24</sup> and using a fixed value (5 or 10 mm Hg) in three studies (11%).<sup>28 32 34</sup> Three  
18  
19 92 studies (11%) did not report their method for calculating RAP.<sup>11 13 43</sup> Four studies (15%) used  
20  
21  
22 93 a tricuspid gradient ( $4\text{TRVmax}^2$ ) instead of SPAP.<sup>12 29 30 38</sup>

23  
24  
25 94 The majority of the studies (22/27, 81%) reported the time interval (mean or maximum)  
26  
27 95 between TTE and RHC,<sup>10-13 23-29 31-35 38 40-42 44 45</sup> while five (5/9, 19%) did not.<sup>30 36 37 39 43</sup> Nine  
28  
29 96 studies (33%) considered time intervals greater than one week,<sup>10 13 24 25 27 31 38 40 42</sup> while thirteen  
30  
31 97 studies (48%) considered time intervals of less than one week.<sup>11 12 23 26 29 32-35 37 39 41 44</sup> The time  
32  
33  
34 98 interval between TTE and RHC ranged from four hours to three months.

### 35 36 37 38 99 **Quality Assessment**

39  
40  
41 100 The quality assessment of the included studies according to the QUADAS-2 inventory is  
42  
43 101 shown in Figure 2. In twenty (74%) study protocols,<sup>10-13 23 24 26 28-32 34 35 37-39 41 44 45</sup> consecutive  
44  
45 102 subjects were enrolled, with no inappropriate exclusions. The risk of bias during patient  
46  
47 103 recruitment was unclear in the remaining seven (26%) studies,<sup>25 27 33 36 40 42 43</sup> as patient  
48  
49 104 recruitment was not reported. In six (22%) studies investigators designed the single-blind  
50  
51 105 methods for TTE.<sup>10 12 23 26 39 45</sup> Double blinding in imaging assessment was not mentioned in any  
52  
53 106 studies. The risk of bias on flow and timing between the index test and reference standard was  
54  
55 107 categorized as unclear in 14 (52%) study protocols that did not explicitly state the successful  
56  
57  
58  
59  
60 9

1  
2  
3  
4 108 investigation with both index and reference tests in all included patients.<sup>24 30-40 42 43</sup>  
5  
6

### 7 109 **Quantitative Analysis**

8  
9

10 110 The SROC curve for TTE is shown in Figure 3. Four studies fall within the 95% confidence  
11 interval.<sup>11 26 34 44</sup> The AUC was 0.88 (95%CI 0.85–0.90). The pooled sensitivity and specificity  
12 111 for TTE were 85% (95%CI 81%–90%) and 74% (95%CI 64%–81%), respectively (Figure 4).  
13 112 The pooled PLR and NLR were 3.2 (95%CI 2.3–4.4) and 0.20 (95%CI 0.15–0.26). The pooled  
14 113 DOR for TTE was 16 (95%CI 10–27).  
15 114

16 115 The heterogeneity in our study was significant. The threshold test proved that the threshold  
17 116 effect was not the source of heterogeneity ( $r=0.34$ ,  $P=0.12$ ). Deeks' test for funnel plot  
18 117 asymmetry suggested no publication bias ( $P=0.69$ ). The results of the subgroup analysis are  
19 118 presented in Table 2. The sensitivity (87%, 95%CI 81%–91%), specificity (74%, 95%CI 62%–  
20 119 83%) and AUC (0.89, 95%CI 0.86–0.91) of TTE to diagnose PH was higher for studies  
21 120 published in 2010 and later compared to those published before 2010. Among the time interval  
22 121 subgroups, the group with the shortest time interval between TTE and RHC had the best  
23 122 diagnostic effect, with sensitivity, specificity and AUC of 88% (95%CI 73%–95%), 90%  
24 123 (95%CI 53%–99%) and 0.94(95%CI 0.92–0.96), respectively. The disease composition of the  
25 124 study population also affected the diagnostic accuracy of TTE. Compared with patients with  
26 125 other diseases, TTE had lower sensitivity (81%, 95%CI 70%–88%), specificity (61%, 95%CI  
27 126 53%–69%) and AUC (0.73, 95%CI 0.69–0.77) in the subgroup of patients with definite lung  
28 127 diseases.  
29 128

30 129 Subgroup analysis of different cut-off thresholds to diagnose PH based on TTE showed that  
31 130 the subgroup with a cut-off threshold of 35 mmHg had superior performance than that at 40  
32 131 mmHg.  
33 132  
34 133  
35 134  
36 135  
37 136  
38 137  
39 138  
40 139  
41 140  
42 141  
43 142  
44 143  
45 144  
46 145  
47 146  
48 147  
49 148  
50 149  
51 150  
52 151  
53 152  
54 153  
55 154  
56 155  
57 156  
58 157  
59 158  
60 159  
10

1  
2  
3  
4 130 mmHg. The sensitivity, specificity and AUC of the former were respectively 92% (95%CI  
5  
6 131 88%–94%), 65% (95%CI 43%–83%) and 0.92 (95%CI 0.89–0.94), while the sensitivity,  
7  
8  
9 132 specificity and AUC at 40 mmHg were 84% (95%CI 75%–91%), 52% (95%CI 31%–71%) and  
10  
11  
12 133 0.80 (95%CI 76%–83%).

13  
14 134 The sensitivity analysis results are shown in Table 3. After excluding low-quality studies and  
15  
16  
17 135 studies with specific characteristics, sensitivity analysis did not reveal a source for the  
18  
19  
20 136 heterogeneity in the diagnostic accuracy analysis. Overall, the pooled meta-analysis results for  
21  
22  
23 137 outcomes were in accordance to our sensitivity analyses.

## 24 25 138 Discussion

26  
27 139 Our study found that TTE has a better sensitivity but moderate specificity for the detection  
28  
29  
30 140 of PH. In addition, shortening the time interval between TTE and RHC and developing an  
31  
32  
33 141 appropriate threshold could improve the accuracy of TTE. However, the accuracy of TTE to  
34  
35  
36 142 diagnose PH in patients with lung diseases was low.

37  
38 143 Although PH is a chronic disease, we still believe that the shortest possible time interval  
39  
40  
41 144 between TTE and RHC is more conducive. Otherwise, changes in the patient's condition and  
42  
43  
44 145 the application of intervention measures would lead to an increase in the deviation of the results  
45  
46  
47 146 of the two examinations. A detailed subgroup analysis was performed according to the time  
48  
49  
50 147 interval between TTE and RHC. As expected, the diagnostic accuracy was highest when the  
51  
52  
53 148 time interval was less than or equal to 24 hours. The results also showed that the efficacy of  
54  
55  
56 149 TTE in the diagnosis of PH was gradually reduced with the extension of the time interval.

57  
58 150 Subgroups analysis based on the disease composition of the population suggested that the  
59  
60 151 diagnostic accuracy of TTE was lower in patients with lung diseases. Changes associated with

1  
2  
3  
4 152 chronic pulmonary disease, including marked increase in intrathoracic gas, consolidation of  
5  
6 153 lung tissue, expansion of the thoracic cage, and alterations in the position of the heart, adversely  
7  
8  
9 154 affect the imaging quality and parameters measurement of TTE.<sup>46</sup> Therefore, the use of TTE to  
10  
11  
12 155 measure pulmonary pressure in patients with lung diseases might not be an ideal choice.

13  
14 156 Guideline recommend the use of IVC width and collapse rate to estimate RAP,<sup>3</sup> which was  
15  
16  
17 157 not used in some of the included studies. Sensitivity analysis for this point showed that studies  
18  
19  
20 158 which calculated RAP through IVC do not seem to have a higher diagnostic performance. In  
21  
22  
23 159 order to avoid errors caused by RAP estimation, TRVmax was also considered as an indicator  
24  
25 160 to evaluate the possibility of PH. Four studies using TRPG (4TRVmax2) instead of SPAP were  
26  
27  
28 161 grouped into a subgroup and showed that this subgroup had good diagnostic specificity but poor  
29  
30 162 sensitivity.

31  
32  
33 163 The sensitivity analysis based on the mean pulmonary artery pressure (MPAP) threshold of  
34  
35 164 25 mmHg did not result in a higher diagnostic value than the whole, indicating that the overall  
36  
37  
38 165 results were stable. It has been suggested that MPAP threshold of 25 mmHg is arbitrarily chosen  
39  
40  
41 166 and that lowering it to 20 mmHg (two standard deviations higher than MPAP for the population)  
42  
43 167 is considered a scientific methods.<sup>47</sup> However, some scientists insist that it is premature to  
44  
45  
46 168 reduce the MPAP threshold to 20 mmHg because of the risk of over-diagnosis, unclear  
47  
48 169 treatment implications and additional psychological burden on patients.<sup>48</sup> Since none of the  
49  
50  
51 170 study we included used MPAP>20mmHg as the diagnostic threshold for RHC, subgroup  
52  
53  
54 171 analysis on the two thresholds of 20mmHg and 25mmHg could not be performed. Therefore,  
55  
56  
57 172 we expect that more studies may be conducted in the future to verify the appropriate threshold  
58  
59 173 of RHC.

1  
2  
3  
4 174 In our review, the cut-off thresholds of SPAP ranged from 30 to 50 mmHg. Subgroup  
5  
6 175 analysis showed that the diagnostic accuracy of the group of 35 mmHg was higher. Sensitivity  
7  
8  
9 176 analysis results of studies that excluded high TTE cut-off value showed that a high cut-off value  
10  
11  
12 177 increased the specificity and reduced the sensitivity of TTE. Due to the small sample size of the  
13  
14 178 subgroup in this study, the value of the cut-off threshold still needs to be determined by further  
15  
16  
17 179 prospective studies of multi-center and large samples.

18  
19 180 Subgroup analysis according to the publication year confirmed that studies published after  
20  
21  
22 181 2010 had only a slightly higher diagnostic accuracy than previous studies. With the  
23  
24 182 improvement of TTE technology and instruments in the past ten years, the diagnostic efficiency  
25  
26  
27 183 of PH has not been significantly improved, which forces us to pay attention to other TTE  
28  
29  
30 184 parameters.<sup>49,50</sup> Perhaps, this could be a new direction for future studies on PH diagnosis.

### 31 32 33 185 **Limitations:**

34  
35 186 Several limitations are present in our study. Firstly, the systematic review and meta-analysis  
36  
37  
38 187 is a secondary research method based on original research and the quality of the included study  
39  
40  
41 188 affects the results. In addition, the possibility of missing relevant articles objectively exists, and  
42  
43 189 significant heterogeneity may limit the interpretation of the results. Secondly, the accuracy of  
44  
45  
46 190 echocardiography relies heavily on the operator's ability, experience, and operational discipline.  
47  
48 191 In order to obtain more original studies, we did not strictly controlled this aspect. Thirdly, the  
49  
50  
51 192 studies included in this review involve several different types of PH, and some studies do not  
52  
53  
54 193 describe the basic disease and PH type in detail. It is obvious that pulmonary lesions can affect  
55  
56  
57 194 the quality of TTE imaging, leading to underestimated results.

### 58 59 195 **Conclusion**



1  
2  
3  
4 196 TTE has clinical value in the diagnosis of PH thanks to its better sensitivity and moderate  
5  
6 197 specificity, but it cannot yet replace RHC considered as the gold standard. Shortening the time  
7  
8  
9 198 interval between TTE and RHC and developing an appropriate threshold can improve the  
10  
11 199 accuracy of TTE. TTE may not be suitable to assess pulmonary arterial pressure in patients with  
12  
13 200 pulmonary disease. It may be necessary to combine multiple TTE parameters and conduct  
14  
15  
16 201 multi-center, large-sample studies to further improve the accuracy of TTE in the diagnosis of  
17  
18  
19 202 PH in future research.  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

### Author's contribution

The joint corresponding authors (JQ.L and B.S) are responsible for the design and implementation of the study. SD.L is responsible for the quality control of study selection. Y.H carried out quality control on the links of data extraction. KH.Y provided guidance in literature retrieval and data processing methodology and was responsible for the quality evaluation part. JR.N and PJ.Y have done the systematic review of the literature and extracted data. JR.N has conducted the meta-analyses, and two authors (JR.N, PJ.Y) have substantially contributed to interpretation of data and co-authored the article. All the authors have made repeated revisions to the article. The corresponding authors (JQ.L and B.S) and JR.N take responsibility of the integrity of the analyses.

## Reference

1. Hoeper MM, Humbert M, Souza R, et al. A global view of pulmonary hypertension. *Lancet Respir Med* 2016;4(4):306-22.
2. Lau EM, Humbert M, Celermajer DS. Early detection of pulmonary arterial hypertension. *Nat Rev Cardiol* 2015;12(3):143-55.
3. Galie N, Humbert M, Vachiery JL, et al. 2015 ESC/ERS Guidelines for the diagnosis and treatment of pulmonary hypertension. *Kardiol Pol* 2015;73(12):1127-206.
4. Hoeper MM, Lee SH, Voswinckel R, et al. Complications of right heart catheterization procedures in patients with pulmonary hypertension in experienced centers. *J Am Coll Cardiol* 2006;48(12):2546-52.
5. Tian J, Zhang J, Ge L, et al. The methodological and reporting quality of systematic reviews from China and the USA are similar. *J Clin Epidemiol* 2017;85:50-58.
6. Xiu-xia L, Ya Z, Yao-long C, et al. The reporting characteristics and methodological quality of Cochrane reviews about health policy research. *Health Policy* 2015;119(4):503-10.
7. Janda S, Shahidi N, Gin K, et al. Diagnostic accuracy of echocardiography for pulmonary hypertension: a systematic review and meta-analysis. *Heart* 2011;97(8):612-22.
8. Taleb M, Khuder S, Tinkel J, et al. The diagnostic accuracy of Doppler echocardiography in assessment of pulmonary artery systolic pressure: a meta-analysis. *Echocardiography* 2013;30(3):258-65.
9. Zhang RF, Zhou L, Ma GF, et al. Diagnostic value of transthoracic Doppler

- 1  
2  
3  
4 echocardiography in pulmonary hypertension: a meta-analysis. *Am J Hypertens*  
5  
6 2010;23(12):1261-64.  
7  
8  
9 10. Ahmed M, Elshinawy O, Agmy G, et al. Tricuspid regurgitation velocity versus right  
10  
11 ventricular systolic pressure in the echocardiographic evaluation of pulmonary  
12  
13 hypertension. *Egypt J Chest Dis Tuberc* 2019;68(2):203-08.  
14  
15  
16 11. Balci MK, Ari E, Vayvada M, et al. Assessment of Pulmonary Hypertension in Lung  
17  
18 Transplantation Candidates: Correlation of Doppler Echocardiography With Right  
19  
20 Heart Catheterization. *Transplant Proc* 2016;48(8):2797-802.  
21  
22  
23 12. Schneider M, Pistritto AM, Gerges C, et al. Multi-view approach for the diagnosis of  
24  
25 pulmonary hypertension using transthoracic echocardiography. *Int J Cardiovasc*  
26  
27 *Imaging* 2018;34(5):695-700.  
28  
29  
30 13. Shujaat A, Bajwa AA, Al-Saffar F, et al. Diagnostic accuracy of echocardiography  
31  
32 combined with chest CT in pulmonary hypertension. *Clin Respir J* 2018;12(3):948-52.  
33  
34  
35 14. Moher D, Liberati A, Tetzlaff J, et al. Preferred reporting items for systematic reviews  
36  
37 and meta-analyses: the PRISMA statement. *BMJ* 2009;339:b2535.  
38  
39  
40 15. Shamseer L, Moher D, Clarke M, et al. Preferred reporting items for systematic review  
41  
42 and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. *BMJ*  
43  
44 2015;350:g7647.  
45  
46  
47 16. Ge L, Tian JH, Li YN, et al. Association between prospective registration and overall  
48  
49 reporting and methodological quality of systematic reviews: a meta-epidemiological  
50  
51 study. *J Clin Epidemiol* 2018;93:45-55.  
52  
53  
54 17. Wang X, Chen Y, Yao L, et al. Reporting of declarations and conflicts of interest in  
55  
56  
57  
58  
59  
60

17

- WHO guidelines can be further improved. *J Clin Epidemiol* 2018;98:1-8.
18. Whiting P, Rutjes AW, Dinnes J, et al. Development and validation of methods for assessing the quality of diagnostic accuracy studies. *Health Technol Assess* 2004;8(25):iii, 1-234.
19. Whiting P, Rutjes AW, Reitsma JB, et al. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Med Res Methodol* 2003;3:25.
20. Lee J, Kim KW, Choi SH, et al. Systematic Review and Meta-Analysis of Studies Evaluating Diagnostic Test Accuracy: A Practical Review for Clinical Researchers- Part II. Statistical Methods of Meta-Analysis. *Korean J Radiol* 2015;16(6):1188-96.
21. Deeks JJ, Macaskill P, Irwig L. The performance of tests of publication bias and other sample size effects in systematic reviews of diagnostic test accuracy was assessed. *J Clin Epidemiol* 2005;58(9):882-93.
22. Reitsma JB, Glas AS, Rutjes AW, et al. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. *J Clin Epidemiol* 2005;58(10):982-90.
23. Arcasoy SM, Christie JD, Ferrari VA, et al. Echocardiographic assessment of pulmonary hypertension in patients with advanced lung disease. *Am J Respir Crit Care Med* 2003;167(5):735-40.
24. Denton CP, Cailles JB, Phillips GD, et al. Comparison of Doppler echocardiography and right heart catheterization to assess pulmonary hypertension in systemic sclerosis. *Br J Rheumatol* 1997;36(2):239-43.

- 1  
2  
3  
4 25. Fisher MR, Criner GJ, Fishman AP, et al. Estimating pulmonary artery pressures by  
5 echocardiography in patients with emphysema. *Eur Respir J* 2007;30(5):914-21.  
6  
7  
8  
9 26. Greiner S, Jud A, Aurich M, et al. Reliability of noninvasive assessment of systolic  
10 pulmonary artery pressure by Doppler echocardiography compared to right heart  
11 catheterization: analysis in a large patient population. *J Am Heart Assoc* 2014;3(4).  
12  
13  
14  
15 27. Habash F, Gurram P, Almomani A, et al. Correlation between echocardiographic  
16 pulmonary artery pressure estimates and right heart catheterization measurement in  
17 liver transplant candidates. *J Cardiovasc Imaging* 2018;26(2):75-84.  
18  
19  
20  
21  
22  
23 28. Hsu VM, Moreyra AE, Wilson AC, et al. Assessment of pulmonary arterial  
24 hypertension in patients with systemic sclerosis: comparison of noninvasive tests with  
25 results of right-heart catheterization. *J Rheumatol* 2008;35(3):458-65.  
26  
27  
28  
29  
30  
31 29. Hua R, Sun YW, Wu ZY, et al. Role of 2-dimensional Doppler echo-cardiography in  
32 screening portopulmonary hypertension in portal hypertension patients. *Hepatobiliary*  
33 *Pancreat. Hepatobiliary Pancreat Dis Int* 2009;8(2):157-61.  
34  
35  
36  
37  
38  
39 30. Keir GJ, Wort SJ, Kokosi M, et al. Pulmonary hypertension in interstitial lung disease:  
40 Limitations of echocardiography compared to cardiac catheterization. *Respirology*  
41 2018;23(7):687-94.  
42  
43  
44  
45  
46  
47 31. Kim WR, Krowka MJ, Plevak DJ, et al. Accuracy of doppler echocardiography in the  
48 assessment of pulmonary hypertension in liver transplant candidates. *Liver Transpl*  
49 2000;6(4):453-58.  
50  
51  
52  
53  
54  
55 32. Laaban JP, Diebold B, Zelinski R, et al. Noninvasive Estimation of Systolic  
56 Pulmonary-Artery Pressure Using Doppler Echocardiography in Patients with Chronic  
57  
58  
59  
60

- 1  
2  
3  
4 Obstructive Pulmonary-Disease. *Chest* 1989;96(6):1258-62.
- 5  
6  
7 33. Lafitte S, Pillois X, Reant P, et al. Estimation of pulmonary pressures and diagnosis of  
8  
9 pulmonary hypertension by doppler echocardiography: A retrospective comparison of  
10  
11 routine echocardiography and invasive hemodynamics. *J Am Soc Echocardiogr*  
12  
13 2013;26(5):457-63.
- 14  
15  
16 34. Lange TJ, Baumgartner S, Arzt M, et al. Qualitative echocardiography parameters for  
17  
18 prediction of pulmonary hypertension. *Int J Clin Pract* 2013;67(SUPPL. 179):5-12.
- 19  
20  
21 35. Lanzarini L, Fontana A, Campana C, et al. Two simple echo-doppler measurements  
22  
23 can accurately identify pulmonary hypertension in the large majority of patients with  
24  
25 chronic heart failure. *J Heart Lung Transplant* 2005;24(6):745-54.
- 26  
27  
28 36. Matsuyama W, Ohkubo R, Michizono K, et al. Usefulness of transcutaneous Doppler  
29  
30 jugular venous echo to predict pulmonary hypertension in COPD patients. *Eur Respir*  
31  
32 *J* 2001;17(6):1128-31.
- 33  
34  
35 37. Mogollon Jimenez MV, Escorcesca Ortega AM, Cabeza Letran ML, et al. Correlation  
36  
37 of Echocardiographic and Hemodynamic Parameters in Pulmonary Hypertension  
38  
39 Assessment Prior to Heart Transplantation. *Transplant Proc* 2008;40(9):3023-24.
- 40  
41  
42 38. Mukerjee D, St George D, Knight C, et al. Echocardiography and pulmonary function  
43  
44 as screening tests for pulmonary arterial hypertension in systemic sclerosis.  
45  
46 *Rheumatology (Oxford)* 2004;43(4):461-6.
- 47  
48  
49 39. Nagel C, Henn P, Ehlken N, et al. Stress Doppler echocardiography for early detection  
50  
51 of systemic sclerosis-associated pulmonary arterial hypertension. *Arthritis Res Ther*  
52  
53 2015;17(1):165.
- 54  
55  
56  
57  
58  
59  
60

- 1  
2  
3  
4 40. Nathan SD, Shlobin OA, Barnett SD, et al. Right ventricular systolic pressure by  
5 echocardiography as a predictor of pulmonary hypertension in idiopathic pulmonary  
6 fibrosis. *Respir Med* 2008;102(9):1305-10.  
7  
8  
9  
10  
11 41. Parsaee M, Amin A, Nematollahi MR, et al. Comparison of transthoracic  
12 echocardiography and right heart catheterization for assessing pulmonary arterial  
13 pressure in patients with congenital or valvular heart defects. *Iran Heart J*  
14 2012;12(4):54-61.  
15  
16  
17 42. Penning S, Robinson KD, Major CA, et al. A comparison of echocardiography and  
18 pulmonary artery catheterization for evaluation of pulmonary artery pressures in  
19 pregnant patients with suspected pulmonary hypertension. *Am J Obstet Gynecol*  
20 2001;184(7):1568-70.  
21  
22  
23 43. Raevens S, Colle I, Reyntjens K, et al. Echocardiography for the detection of  
24 portopulmonary hypertension in liver transplant candidates: An analysis of cutoff  
25 values. *Liver Transpl* 2013;19(6):602-10.  
26  
27  
28 44. Rajaram S, Swift AJ, Capener D, et al. Comparison of the diagnostic utility of cardiac  
29 magnetic resonance imaging, computed tomography, and echocardiography in  
30 assessment of suspected pulmonary arterial hypertension in patients with connective  
31 tissue disease. *J Rheumatol* 2012;39(6):1265-74.  
32  
33  
34 45. Sohrabi B, Kazemi B, Mehryar A, et al. Correlation between pulmonary artery pressure  
35 measured by echocardiography and right heart catheterization in patients with  
36 rheumatic mitral valve stenosis (A prospective study). *Echocardiography*  
37 2016;33(1):7-13.  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60



- 1  
2  
3  
4 46. Galderisi M, Cosyns B, Edvardsen T, et al. Standardization of adult transthoracic  
5  
6 echocardiography reporting in agreement with recent chamber quantification, diastolic  
7  
8 function, and heart valve disease recommendations: an expert consensus document of  
9  
10 the European Association of Cardiovascular Imaging. *Eur Heart J Cardiovasc Imaging*  
11  
12 2017;18(12):1301-10.  
13  
14  
15  
16  
17 47. Simonneau G, Montani D, Celermajer DS, et al. Haemodynamic definitions and  
18  
19 updated clinical classification of pulmonary hypertension. *Eur Respir J* 2019;53(1).  
20  
21  
22 48. Gibbs JSR, Torbicki A. Proposed new pulmonary hypertension definition: is 4 mm(Hg)  
23  
24 worth re-writing medical textbooks? *Eur Respir J* 2019;53(3).  
25  
26  
27 49. Misra D, Kendes A, Sulica R, et al. Exercise-induced pulmonary hypertension by stress  
28  
29 echocardiography: Prevalence and correlation with right heart hemodynamics. *Int J*  
30  
31 *Cardiol* 2017;228:518-22.  
32  
33  
34  
35 50. Vitarelli A, Mangieri E, Terzano C, et al. Three-dimensional echocardiography and  
36  
37 2D-3D speckle-tracking imaging in chronic pulmonary hypertension: diagnostic  
38  
39 accuracy in detecting hemodynamic signs of right ventricular (RV) failure. *J Am Heart*  
40  
41 *Assoc* 2015;4(3):e001584.  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

**Table 1 Characteristics of each study included in this meta-analysis.**

Study	Year	Country	Design	N	Disease Composition of the Population	Mean Age (Years)	Male (%)	Time Interval	TTE Threshold (mmHg)	RHC Threshold (mmHg)	TTE Method
<b>Ahmed Keir</b>	2019	USA	Retrospective	136	Multiple diseases	59±20	35	<3m	SPAP≥40	MPAP≥25	4TRVmax <sup>2</sup> +RAP(IVC)
<b>Habash-1</b>	2018	Australia	Prospective	265	Interstitial lung disease	60.8±16.5	46	–	TRPG>46	MPAP≥25	4TRVmax <sup>2</sup>
<b>Habash-2</b>	2018	USA	Retrospective	31	Liver transplantation candidates	57±11	42	36.8±13.4d	SPAP>47	MPAP≥25	4TRVmax <sup>2</sup> +RAP(IVC)
<b>Schneider</b>	2018	USA	Retrospective	49	Multiple diseases	59±15	31	16.0±11.6d	SPA>43	MPAP≥25	4TRVmax <sup>2</sup> +RAP(IVC)
<b>Balci</b>	2018	Austria	Prospective	65	Cardiac and lung diseases	67.2	43	<48h	TRPG>32	MPAP≥25	4TRVmax <sup>2</sup>
<b>Shujaat</b>	2016	Turkey	Prospective	103	Lung transplantation candidates	47.6±10.4	66	<72h	SPAP>35	MPAP≥25	4TRVmax <sup>2</sup> +RAP (NR)
<b>Sohrabi</b>	2016	USA	Retrospective	87	Multiple diseases	54.3±15.9	29	13d*	SPAP>40	MPAP>25	4TRVmax <sup>2</sup> +RAP (NR)
<b>Nagel</b>	2016	Iran	Prospective	300	Rheumatic mitral stenosis	59.9	31	<24h	SPAP≥35	MPAP≥25	4TRVmax <sup>2</sup> +RAP(IVC)
<b>Greiner</b>	2015	Germany	Prospective	76	Systemic sclerosis	58±14	16	–	SPAP>40	MPAP≥25	4TRVmax <sup>2</sup> +RAP(IVC)
<b>Lafitte</b>	2014	Germany	Retrospective	1695	Cardiac disease	63±15	67	<5d	SPAP≥36	MPAP≥25	4TRVmax <sup>2</sup> +RAP(IVC)
<b>Lange</b>	2013	France	Retrospective	114	Cardiac and lung disease	64.8±15.9	52	<48h	SPAP≥38	MPAP>25	4TRVmax <sup>2</sup> +RAP(IVC)
<b>Raevens</b>	2013	Germany	Retrospective	231	Multiple diseases	62±13	43	5±4d	SPAP>50	MPAP≥25	4TRVmax <sup>2</sup> +RAP (5)
<b>Parsaee</b>	2013	Belgium	Retrospective	152	Liver transplantation candidates	58±11	66	–	SPAP>38	MPAP≥25	4TRVmax <sup>2</sup> +RAP (NR)
<b>Rajaram</b>	2012	Iran	Prospective	103	Cardiac diseases	41.0±15.8	44	<4h	SPAP≥35	MPAP>25	4TRVmax <sup>2</sup> +RAP(IVC)
<b>Hua</b>	2012	UK	Retrospective	81	Connective tissue disease	62±14	15	<48h	TRPG≥40	MPAP≥25	4TRVmax <sup>2</sup>
<b>Nathan</b>	2009	China	Prospective	105	Liver transplantation candidates	49.5±11.8	63	4.2±2.0d	SPAP≥30	MPAP≥25	4TRVmax <sup>2</sup> +RAP(IVC)
<b>Hsu</b>	2008	USA	Retrospective	60	Idiopathic pulmonary fibrosis	62.9±8.6	55	32±78d	SPAP≥40	MPAP>25	4TRVmax <sup>2</sup> +RAP(IVC)
<b>Mogollon</b>	2008	USA	Prospective	49	Systemic Sclerosis	55	18	<4h	SPAP>47	MPAP≥25	4TRVmax <sup>2</sup> +RAP (10)
<b>Fisher</b>	2008	Spain	Retrospective	67	Heart transplantation candidates	–	–	–	SPAP>40	MPAP>35	4TRVmax <sup>2</sup> +RAP(IVC)
<b>Lanzarini</b>	2007	USA	Retrospective	63	Emphysema patients	65.6±6.6	60	23d	SPAP>40	MPAP≥25	4TRVmax <sup>2</sup> +RAP(IVC)
	2005	Italy	Prospective	57	Heart failure	52±11	74	<24h	SPAP≥32	SPAP≥35	4TRVmax <sup>2</sup> +RAP(IVC)

<b>Mukerjee</b>	2004	UK	Prospective	137	Systemic sclerosis	63	–	<3 m	TRPG>40	MPAP≥25	4TRVmax <sup>2</sup>
<b>Arcasoy</b>	2003	USA	Prospective	166	COPD 68%, ILD 28%, PVD 4%	51	43	<72h	SPAP≥45	SPAP≥45	4TRVmax <sup>2</sup> +RAP(IVC)
<b>Penning</b>	2001	USA	Retrospective	27	Pregnant women with cardiac diseases	28.6	0	25.8d	SPAP≥40	SPAP≥35	4TRVmax <sup>2</sup> +RAP(IVC)
<b>Matsuyama</b>	2001	Japan	Prospective	35	COPD	66	94	–	SPAP≥40	MPAP>25	4TRVmax <sup>2</sup> +RAP(IVC)
<b>Kim</b>	2000	USA	Prospective	74	Liver transplantation candidates	54	50	59d	SPAP>50	MPAP≥35	4TRVmax <sup>2</sup> +RAP(JVP)
<b>Denton</b>	1997	UK	Prospective	20	COPD	48.6±11.7	30	1.8±2.3m	SPAP≥30	SPAP≥30	4TRVmax <sup>2</sup> +RAP(JVP)
<b>Laaban</b>	1989	France	Prospective	27	COPD	63±9	78	<2d	SPAP≥35	SPAP≥35	4TRVmax <sup>2</sup> +RAP(5)

USA, United States of America; UK, United Kingdom of Great Britain and Northern Ireland; TTE, Transthoracic echocardiography; RHC, right heart catheterization; SPAP, systolic pulmonary artery pressure; MPAP, mean pulmonary artery pressure; TRPG, tricuspid regurgitation pressure gradient; TRV, tricuspid regurgitation velocity; RAP, right atrial pressure; IVC, Inferior vena cava; JVP, jugular vein pressure; COPD, chronic obstructive pulmonary disease; ILD, Interstitial lung disease; PVD, peripheral vascular disease; NR, not report.

\* The median time (other terms are mean time)

First published as 10.1136/bmjopen-2019-033084 on 22 December 2019. Downloaded from <http://bmjopen.bmj.com/> on April 17, 2024 by guest. Protected by copyright.

**Table 2 Subgroup analysis**

Group	N	<i>I</i> <sup>2</sup> (95%CI)	AUC (95%CI)	Sensitivity (95%CI)	Specificity (95%CI)	PLR (95%CI)	NLR (95%CI)	DOR (95%CI)
<b>All Studies</b>	28	98(97–99)	0.88(0.85–0.90)	0.85(0.81–0.90)	0.74(0.64–0.81)	3.2(2.3–4.4)	20(0.15–0.26)	16(10–27)
<b>Time Interval</b>								
≤24h	4	95(90–99)	0.94(0.92–0.96)	0.88(0.73–0.95)	0.90(0.53–0.99)	8.9(1.5–54.5)	13(0.06–0.29)	68(13–348)
≤48h*	7	95(90–99)	0.94(0.91–0.95)	0.88(0.81–0.93)	0.89(0.71–0.96)	7.8(2.8–21.3)	13(0.09–0.21)	59(23–148)
≤72h*	9	94(89–99)	0.91(0.89–0.93)	0.87(0.82–0.91)	0.83(0.65–0.93)	5.2(2.4–11.2)	15(0.11–0.21)	34(14–82)
≤1 week	13	93(87–99)	0.91(0.88–0.93)	0.87(0.84–0.90)	0.80(0.68–0.88)	4.3(2.7–6.9)	16(0.12–0.21)	27(15–48)
>1 week	10	97(95–99)	0.82(0.78–0.85)	0.85(0.73–0.92)	0.60(0.40–0.77)	2.1(1.3–3.4)	25(0.14–0.45)	9(4–21)
Unclear	5	82(63–100)	0.85(0.81–0.88)	0.79(0.63–0.99)	0.76(0.61–0.87)	3.4(1.9–5.9)	27(0.15–0.51)	12(5–33)
<b>Population Disease</b>								
Cardiac diseases	6	94(89–99)	0.90(0.87–0.92)	0.90(0.86–0.93)	0.67(0.29–0.91)	2.7(0.9–8.1)	15(0.08–0.30)	18(3–95)
Lung diseases	8	90(81–100)	0.73(0.69–0.77)	0.81(0.70–0.88)	0.61(0.53–0.69)	2.1(1.8–2.4)	32(0.21–0.48)	7(4–10)
Multiple diseases <sup>#</sup>	6	93(87–99)	0.90(0.87–0.92)	0.89(0.84–0.92)	0.70(0.40–0.89)	3.0(1.3–7.1)	16(0.11–0.23)	19(6–60)
Unclear <sup>&amp;</sup>	8	88(77–100)	0.88(0.85–0.90)	0.80(0.64–0.90)	0.85(0.80–0.89)	5.3(4.0–7.0)	23(0.12–0.45)	23(10–51)
<b>Published Year</b>								
≥2010	15	97(95–99)	0.89(0.86–0.91)	0.87(0.81–0.91)	0.74(0.62–0.83)	3.3(2.3–4.9)	18(0.13–0.25)	19(11–13)
<2010	13	96(93–99)	0.86(0.83–0.89)	0.84(0.74–0.90)	0.73(0.56–0.85)	3.1(1.8–5.3)	22(0.14–0.37)	14(6–33)
<b>Study Design</b>								
Prospective	15	97(95–99)	0.90(0.87–0.92)	0.86(0.77–0.91)	0.79(0.69–0.87)	4.2(2.7–6.4)	18(0.11–0.28)	23(12–45)
Retrospective	13	96(92–99)	0.86(0.83–0.89)	0.86(0.80–0.90)	0.65(0.49–0.78)	2.5(1.6–3.7)	22(0.15–0.32)	11(6–22)
<b>TTE Threshold</b>								

25

SPAP $\geq$ 40 mmHg	8	96(93–99)	0.80(0.76–0.83)	0.84(0.75–0.91)	0.52(0.31–0.71)	1.7(1.2–2.5)	0.30(0.21–0.44)	6(3–11)
SPAP $\geq$ 35 mmHg	4	76(47–100 )	0.92(0.89–0.94 )	0.92(0.88–0.94)	0.65(0.43–0.83)	2.6(1.4–4.9)	0.13(0.08–0.22)	16(9–28)
TRPG	4	0(0–100)	0.85(0.82–0.88)	0.75(0.58–0.86)	0.81(0.70–0.89)	4.0(2.2–7.3)	0.31(0.17–0.57)	13(4–40)

AUC, Area under curve; PLR, positive likelihood ratio; NLR, negative likelihood ratio; DOR, diagnostic odds ratio; TTE, Transthoracic echocardiography; SPAP, systolic pulmonary artery pressure; TRPG, tricuspid regurgitation pressure gradient.

\*Studies with time intervals less than or equal to 24 hours were included in this subgroup.

\* Studies with time intervals less than or equal to 24 hours and 48 hours were included in this subgroup.

# Studies included a variety of diseases, including heart disease and lung disease.

&Diseases were not specifically identified in the studies (transplant candidates).

Downloaded from <http://bmjopen.bmj.com/> on April 17, 2024 by guest. Protected by copyright.

**Table 3 Sensitivity analysis for diagnostic accuracy meta-analysis.**

Study characteristic	N	$I^2$ (95%CI)	AUC (95%CI)	Sensitivity (95%CI)	Specificity (95%CI)	PLR (95%CI)	NLR (95%CI)	DOR (95%CI)
All included studies	28	98(97–99)	0.88(0.85–0.90)	0.85(0.81–0.90)	0.74(0.64–0.81)	3.2(2.3–4.4)	0.20(0.15–0.26)	16(10–27)
Excluding study of Penning*	27	98(97–99)	0.88(0.85–0.91)	0.86(0.81–0.89)	0.75(0.66–0.82)	3.4(2.5–4.6)	0.19(0.14–0.26)	18(11–28)
RHC threshold MPAP $\geq$ 25 mmHg	21	98(97–99)	0.87(0.84–0.90)	0.83(0.77–0.88)	0.76(0.67–0.83)	3.5(2.5–4.8)	0.22(0.16–0.30)	16(10–26)
RAP method(IVC) <sup>‡</sup>	17	96(93–99)	0.89(0.86–0.91)	0.86(0.82–0.90)	0.73(0.59–0.84)	3.2(2.0–5.1)	0.19(0.13–0.27)	17(8–35)
Excluding high TTE threshold*	21	97(95–99)	0.90(0.87–0.92)	0.88(0.85–0.91)	0.72(0.59–0.82)	3.2(2.1–4.8)	0.16(0.12–0.22)	20(11–36)

AUC, Area under curve; PLR, positive likelihood ratio; NLR, negative likelihood ratio; DOR, diagnostic odds ratio; MPAP, mean pulmonary artery pressure; TTE, Transthoracic echocardiography; RHC, right heart catheterization; RAP, right atrial pressure; IVC, Inferior vena cava.

\* Study of Penning was excluded because only pregnant women with cardiac disease were included as subjects.

‡ The RAP was calculated through the diameter and collapse of IVC during spontaneous respiration. Habash's study was divided into two independent parts, so the results section showed 16 studies, but 17 sets of data were analyzed.

\* High TTE threshold was defined as SPAP greater than 45 mmHg or tricuspid regurgitation pressure gradient (TRPG) greater than 40 mmHg.

1  
2  
3  
4  
5 **Figure legend/caption**  
6  
7  
8  
9

10 **Figure 1** Flowchart for identification of studies.

11  
12  
13 ·Habash's study was divided into two independent parts because of the differences between the  
14 case group (Habash-1) and the control group (Habash-2). 27 studies were included, but 28 sets  
15 of data were analyzed.  
16  
17  
18  
19

20  
21  
22  
23 **Figure 2** Risk of bias and applicability concerns summary: review authors' judgements about  
24 each domain for each included study (n=28).  
25  
26  
27  
28  
29

30  
31 **Figure 3** Summary receiver operating characteristic (SROC) graph with 95% confidence region  
32 and 95% prediction region for TTE in the diagnosis of pulmonary hypertension (n=28).  
33  
34  
35  
36  
37  
38

39 **Figure 4** Forrest plot of the sensitivity and specificity of each individual study, summary  
40 sensitivity and specificity and  $I^2$  statistic for heterogeneity (n=28).  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

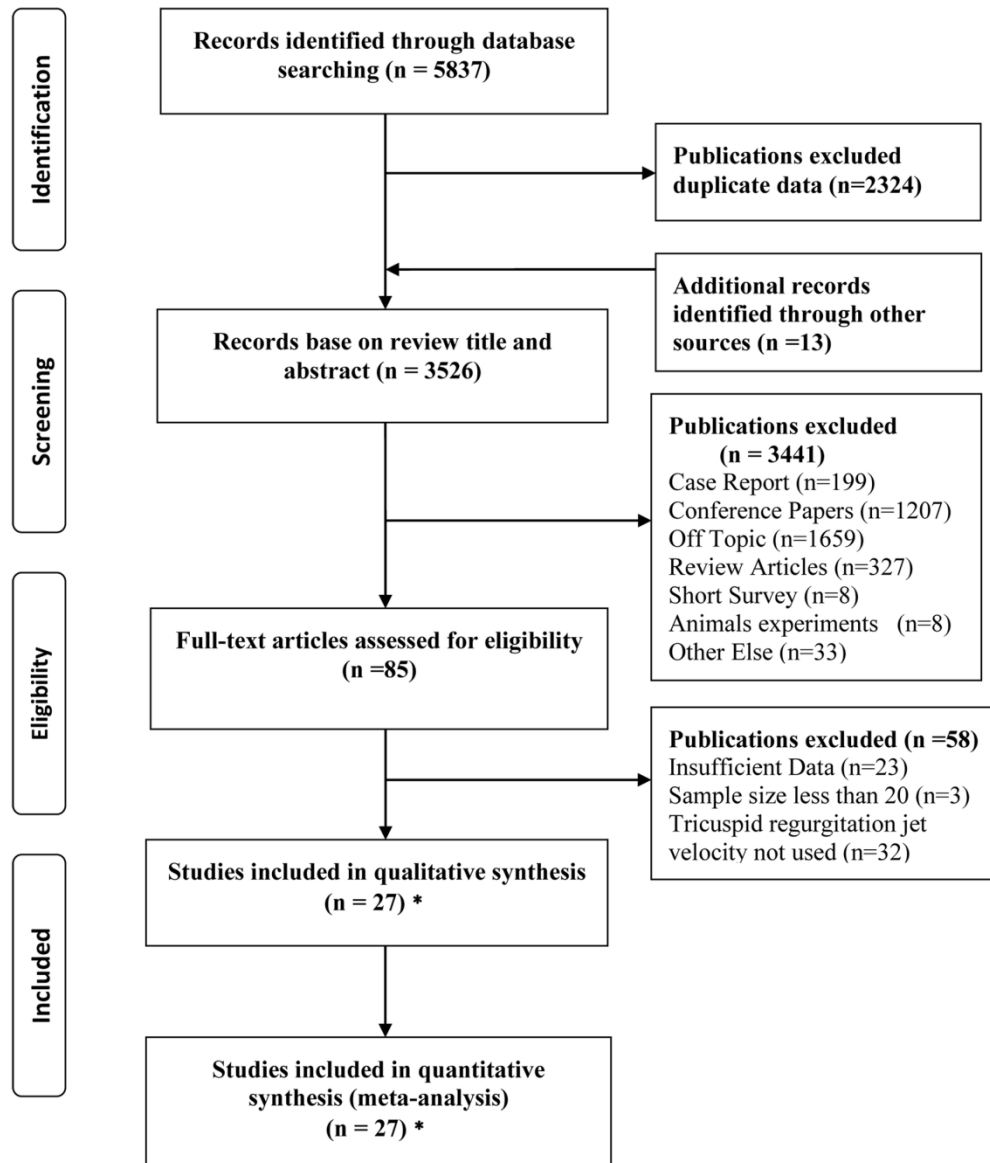


Figure 1 Flowchart for identification of studies.

·Habash's study was divided into two independent parts because of the differences between the case group (Habash-1) and the control group (Habash-2). 27 studies were included, but 28 sets of data were analyzed.

157x188mm (300 x 300 DPI)



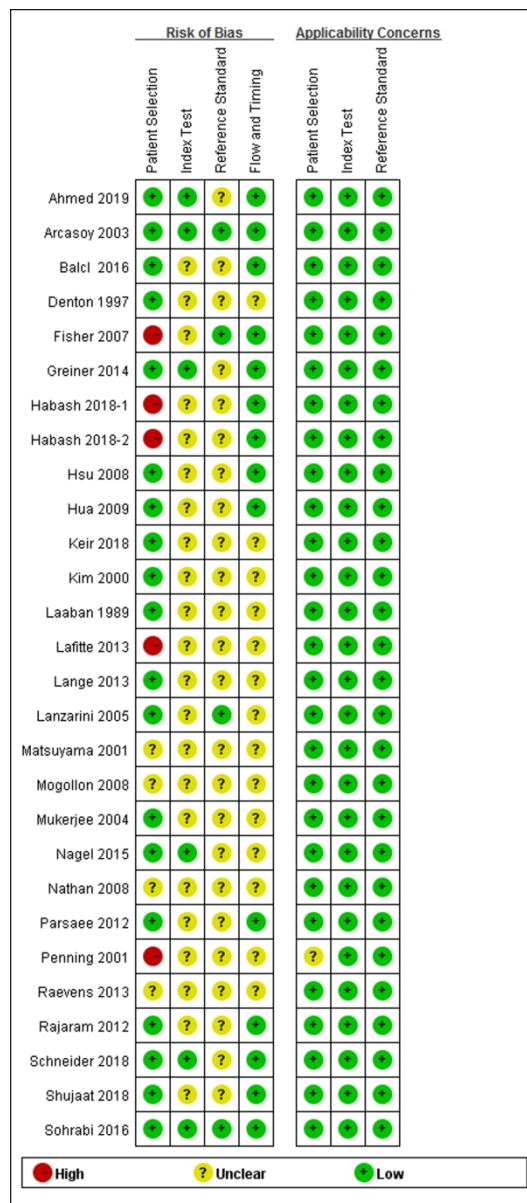


Figure 2 Risk of bias and applicability concerns summary: review authors' judgements about each domain for each included study (n=28).

BMJ Open: first published as 10.1136/bmjopen-2019-033084 on 22 December 2019. Downloaded from <http://bmjopen.bmj.com/> on April 17, 2024 by guest. Protected by copyright.

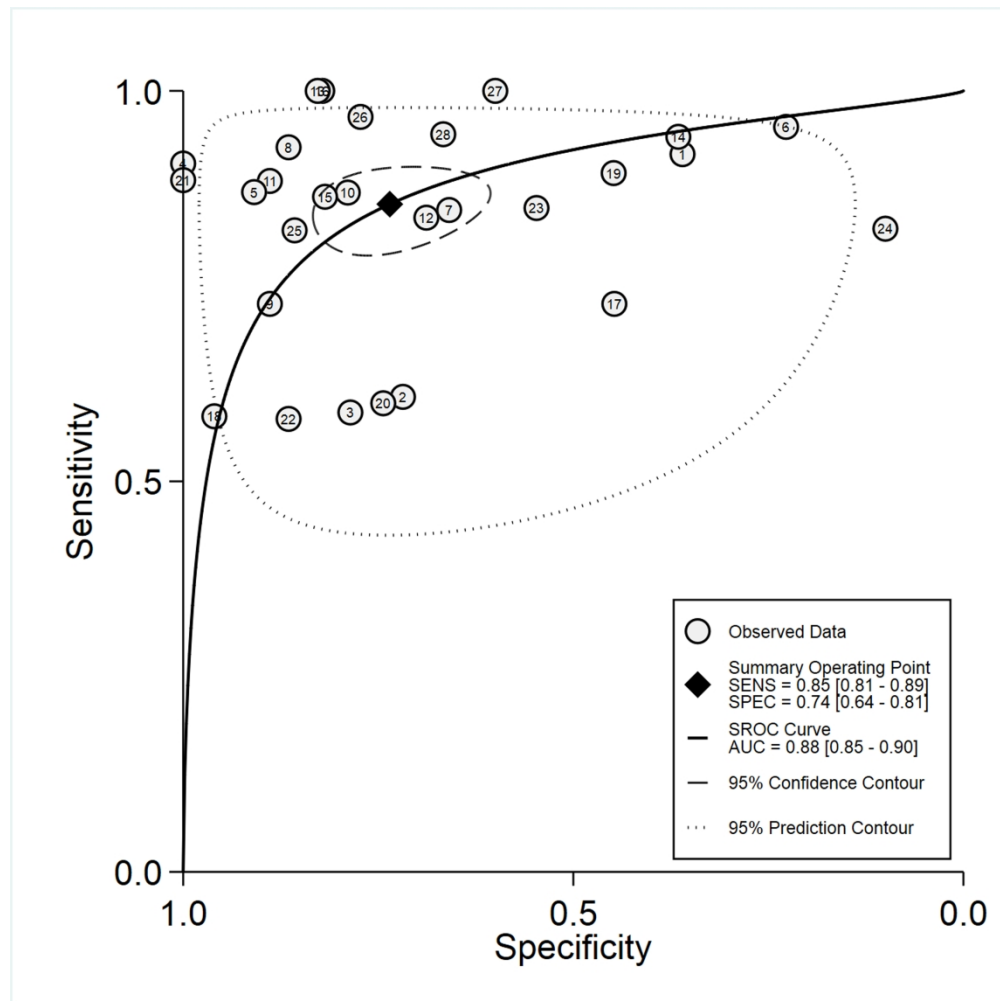


Figure 3 Summary receiver operating characteristic (SROC) graph with 95% confidence region and 95% prediction region for TTE in the diagnosis of pulmonary hypertension (n=28). The study of Habash et al was divided into two independent items, because their case group and control group provided different diagnostic data.

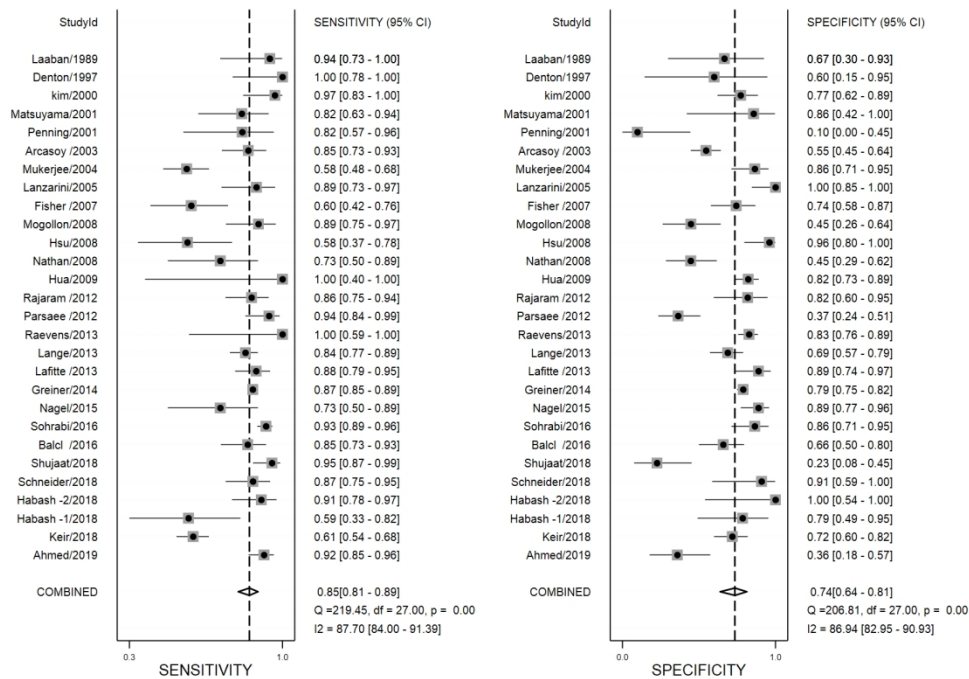


Figure 4 Forrest plot of the sensitivity and specificity of each individual study, summary sensitivity and specificity and I2 statistic for heterogeneity (n=28).

BMJ Open: first published as 10.1136/bmjopen-2019-033084 on 22 December 2019. Downloaded from <http://bmjopen.bmj.com/> on April 17, 2024 by guest. Protected by copyright.



# PRISMA-DTA Checklist

Section/topic	#	PRISMA-DTA Checklist Item	Reported on page #
<b>TITLE / ABSTRACT</b>			
Title	1	Identify the report as a systematic review (+/- meta-analysis) of diagnostic test accuracy (DTA) studies.	1
Abstract	2	Abstract: See PRISMA-DTA for abstracts.	3
<b>INTRODUCTION</b>			
Rationale	3	Describe the rationale for the review in the context of what is already known.	5
Clinical role of index test	D1	State the scientific and clinical background, including the intended use and clinical role of the index test, and if applicable, the rationale for minimally acceptable test accuracy (or minimum difference in accuracy for comparative design).	5
Objectives	4	Provide an explicit statement of question(s) being addressed in terms of participants, index test(s), and target condition(s).	5
<b>METHODS</b>			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	5
Eligibility criteria	6	Specify study characteristics (participants, setting, index test(s), reference standard(s), target condition(s), and study design) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	6
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	6
Search	8	Present full search strategies for all electronic databases and other sources searched, including any limits used, such that they could be repeated.	Supplementary Data
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	Figure 1
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	6
Definitions for data extraction	11	Provide definitions used in data extraction and classifications of target condition(s), index test(s), reference standard(s) and other characteristics (e.g. study design, clinical setting).	6-7
Risk of bias and applicability	12	Describe methods used for assessing risk of bias in individual studies and concerns regarding the applicability to the review question.	7
Diagnostic accuracy measures	13	State the principal diagnostic accuracy measure(s) reported (e.g. sensitivity, specificity) and state the unit of assessment (e.g. per-patient, per-lesion).	7
Synthesis of results	14	Describe methods of handling data, combining results of studies and describing variability between studies. This could include, but is not limited to: a) handling of multiple definitions of target condition. b) handling of multiple thresholds of test positivity. c) handling multiple index test readers. d) handling of indeterminate test results, e)	7

open-2019-033084 on 22 December 2019. Downloaded from <http://bmjopen.bmj.com/> on April 7, 2024 by guest. Protected by copyright.



# PRISMA-DTA Checklist

grouping and comparing tests, f) handling of different reference standards

Page 1 of 2

Section/topic	#	PRISMA-DTA Checklist Item	Reported on page #
Meta-analysis	D2	Report the statistical methods used for meta-analyses, if performed.	7
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	8
<b>RESULTS</b>			
Study selection	17	Provide numbers of studies screened, assessed for eligibility, included in the review (and included in meta-analysis, if applicable) with reasons for exclusions at each stage, ideally with a flow diagram.	Figure 1
Study characteristics	18	For each included study provide citations and present key characteristics including: a) participant characteristics (presentation, prior testing), b) clinical setting, c) study design, d) target condition definition, e) index test, f) reference standard, g) sample size, h) funding sources	23 Table 1
Risk of bias and applicability	19	Present evaluation of risk of bias and concerns regarding applicability for each study.	Figure 2
Results of individual studies	20	For each analysis in each study (e.g. unique combination of index test, reference standard, and positivity threshold) report 2x2 data (TP, FP, FN, TN) with estimates of diagnostic accuracy and confidence intervals, ideally with a forest or receiver operator characteristic (ROC) plot.	Figure 3
Synthesis of results	21	Describe test accuracy, including variability; if meta-analysis was done, include results and confidence intervals.	Figure 4
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression; analysis of index test: failure rates, proportion of inconclusive results, adverse events).	24-26 Table 2,3
<b>DISCUSSION</b>			
Summary of evidence	24	Summarize the main findings including the strength of evidence.	11
Limitations	25	Discuss limitations from included studies (e.g. risk of bias and concerns regarding applicability) and from the review process (e.g. incomplete retrieval of identified research).	13
Conclusions	26	Provide a general interpretation of the results in the context of other evidence. Discuss implications for future research and clinical practice (e.g. the intended use and clinical role of the index test).	12-13
<b>FUNDING</b>			
Funding	27	For the systematic review, describe the sources of funding and other support and the role of the funders.	2

Adapted From: McInnes MDF, Moher D, Thoms BD, McGrath TA, Bossuyt PM, The PRISMA-DTA Group (2018). Preferred Reporting Items for a Systematic Review and Meta-analysis of Diagnostic Test Accuracy Studies: The PRISMA-DTA Statement. JAMA. 2018 Jan 23;319(4):388-396. doi: 10.1001/jama.2017.19163.

For more information, visit: [www.prisma-statement.org](http://www.prisma-statement.org).

# BMJ Open

## Diagnostic Accuracy of Transthoracic Echocardiography for Pulmonary Hypertension: a Systematic Review and Meta-Analysis

Journal:	<i>BMJ Open</i>
Manuscript ID	bmjopen-2019-033084.R2
Article Type:	Original research
Date Submitted by the Author:	18-Nov-2019
Complete List of Authors:	Ni, Jin-Rong; The First Hospital of Lanzhou University Yan, Pei-Jing; Institute of Clinical Research and Evidence Based Medicine, Gansu Provincial Hospital Liu, Shi-Dong; The First Hospital of Lanzhou University Hu, Yuan; the First Hospital of Lanzhou University Yang, Kehu; evidence based medicine center, lanzhou university Song, Bing; The First Hospital of Lanzhou University Lei, Jun-Qiang; The First Hospital of Lanzhou University
<b>Primary Subject Heading</b>:	Diagnostics
Secondary Subject Heading:	Radiology and imaging, Qualitative research, Global health, Cardiovascular medicine
Keywords:	Hypertension < CARDIOLOGY, Echocardiography < CARDIOLOGY, Diagnostic radiology < RADIOLOGY & IMAGING

SCHOLARONE™  
Manuscripts

## Diagnostic Accuracy of Transthoracic Echocardiography for Pulmonary Hypertension: a Systematic Review and Meta-Analysis

Jin-Rong Ni<sup>1,3,8,9</sup>, Pei-Jing Yan<sup>2,5,6#</sup>, Shi-Dong Liu<sup>1,3</sup>, Yuan Hu<sup>3</sup>, Ke-Hu Yang<sup>2,4,5,6</sup>, Bing Song<sup>3\*</sup>, Jun-Qiang Lei<sup>1,7,8,9\*</sup>

<sup>1</sup>The First Hospital (the First Clinical Medical School) of Lanzhou University, Lanzhou 730000, China;

<sup>2</sup>Institute of Clinical Research and Evidence Based Medicine, Gansu Provincial Hospital, Lanzhou,

730000, China; <sup>3</sup>Department of Cardiovascular Surgery, the First Hospital of Lanzhou University,

Lanzhou 730000, China; <sup>4</sup>Evidence-Based Medicine Center, School of Basic Medical Sciences, Lanzhou

University, Lanzhou 730000, China; <sup>5</sup>Evidence-Based Social Science Research Center, Lanzhou

University, Lanzhou, 730000, China; <sup>6</sup>Key Laboratory of Evidence-based Medicine and Knowledge

Translation of Gansu Province, Lanzhou, 730000, China. <sup>7</sup>Department of Radiology, the First Hospital

of Lanzhou University, Lanzhou 730000, China. <sup>8</sup>Intelligent Imaging Medical Engineering Research

Center of Gansu province, Lanzhou 730000, China; <sup>9</sup>Precision Image and Collaborative Innovation

International Scientific and Technological Cooperation Base of Gansu province, Lanzhou 730000, China.

#Joint first author

\*Joint corresponding authors

**Corresponding Author:** 1. Jun-Qiang Lei, No.11 Donggang West Road, Chengguan District,

Lanzhou, China.

Tel: +8613919289040

E-mail: leijunqiangldyy@163.com

2. Bing Song, No.11 Donggang West Road, Chengguan District,

Lanzhou, China.

Tel: +8613993118707

E-mail: songbingldyyxwk@163.com

**Funding sources:** This study was supported by the Key Laboratory of Evidence Based Medicine and Knowledge Translation Foundation of Gansu Province (Grant No. GSXZYZH2018006).

**Statement on potential conflicts of interest:** None declared.

**Patient and Public Involvement:** There is no patient or the public involved in our work.

**Data Availability:** All data relevant to the study are included in the article.

**Key Words:** Echocardiography; Pulmonary Hypertension; Catheterization; Diagnosis; Accuracy; Meta-analysis.



## Abstract

**Objective:** To evaluate the diagnostic accuracy of transthoracic echocardiography (TTE) in patients with pulmonary hypertension (PH).

**Design:** Systematic review and meta-analysis.

**Data sources and eligibility criteria:** EMBASE, Cochrane Library for clinical trials, PubMed and Web of Science were used to search studies from inception to June 19, 2019. Studies using both TTE and right heart catheterization (RHC) to diagnose PH were included.

**Main results:** A total of 27 studies involving 4386 subjects were considered as eligible for analysis. TTE had a pooled sensitivity of 85%, a pooled specificity of 74%, a pooled positive likelihood ratio of 3.2, a pooled negative likelihood ratio of 0.20, a pooled diagnostic odds ratio of 16, and finally an area under the summary receiver operating characteristic (SROC) curve of 0.88. The subgroup with the shortest time interval between TTE and RHC had the best diagnostic effect, with sensitivity, specificity and AUC of 88%, 90% and 0.94, respectively. TTE had lower sensitivity (81%), specificity (61%) and AUC (0.73) in the subgroup of patients with definite lung diseases. Subgroup analysis also showed that different thresholds of TTE resulted in a different diagnostic performance in the diagnosis of PH.

**Conclusion:** TTE has a clinical value in diagnosing PH, although it cannot yet replace RHC considered as the gold standard. The accuracy of TTE may be improved by shortening the time interval between TTE and RHC and by developing an appropriate threshold. TTE may not be suitable to assess pulmonary arterial pressure in patients with pulmonary diseases.

**Review registration number:** PROSPERO CRD42019123289.

### Strengths and limitations of this study

1. A comprehensive search was conducted in the main database, more studies were included, and a large sample size was obtained.
2. Detailed subgroup analysis and sensitivity analysis were performed.
3. The types of pulmonary hypertension included in the studies could not be distinguished.
4. Significant heterogeneity in our study limits the interpretation of the results.

## 1 Introduction

2 The prevalence of pulmonary hypertension (PH) is estimated at 1% in the general population,  
3 and as high as 10% in the 600 million people older than 65.<sup>1</sup> Early detection and accurate  
4 assessment are vital to obtain better outcomes for PH patients.<sup>2</sup> Right heart catheterization  
5 (RHC) is the gold standard in the diagnosis of PH,<sup>3</sup> but it is invasive and cannot be used  
6 frequently or repeatedly.<sup>4</sup> The latest guideline for PH recommends transthoracic  
7 echocardiography (TTE) as a noninvasive test for screening.<sup>3</sup>

8 High quality meta-analysis has been considered as one of the key tools for achieving  
9 evidence.<sup>5 6</sup> Three systematic reviews and meta-analysis regarding the diagnostic accuracy of  
10 TTE for PH were published between 2010 and 2013.<sup>7-9</sup> Studies included in these meta-analyses  
11 were all published before 2010. In addition, two of them included fewer studies and performed  
12 a simple diagnostic data synthesis.<sup>8 9</sup> The other included a relatively large number of studies,  
13 but did not assess a detailed subgroup analysis.<sup>7</sup> In recent years, TTE has still been used in the  
14 clinical diagnosis of PH, and many new original studies have been published.<sup>10-13</sup> Therefore,  
15 the purpose of our study was to undertake a comprehensive systematic review and quantitative  
16 meta-analysis on the accuracy of TTE in the diagnosis of PH.

## 17 Methods

18 The present study is reported according to the Preferred Reporting Items for Systematic  
19 Reviews and Meta-analyses (PRISMA) statement and the published recommendations.<sup>14 15</sup> The  
20 detailed protocol is accessible in PROSPERO (CRD42019123289).<sup>16 17</sup>

## 21 Data sources and search

22 A systematic search in EMBASE, Cochrane Library for clinical trials, PubMed and Web of  
23 Science was performed to find the relevant literature from inception to June 19, 2019. Subject  
24 words were combined with free words, and the search strategy was developed and adapted for  
25 each database. ClinicalTrials.gov and the trials registers on the World Health Organization  
26 International Clinical Trials Registry Platform were used to search unpublished trails. The  
27 references of the included studies and other systematic reviews and meta-analysis were also  
28 reviewed to obtain a comprehensive list of included studies.

### 29 **Study selection**

30 Studies were selected based on the following inclusion criteria: studies that diagnosed PH by  
31 TTE; study population represented by patients with suspected PH; TTE measurement of  
32 systolic pulmonary artery pressure (SPAP) performed using tricuspid regurgitation; RHC as the  
33 gold standard for the diagnosis of PH.

34 The exclusion criteria were the following: insufficient data to construct a 2×2 table; studies  
35 with less than 20 subjects; duplicate data were used (in this case, the largest sample or the latest  
36 study was selected).

37 Two reviewers (JR.N and PJ.Y) independently screened the eligible studies for suitability.  
38 Disagreements were resolved by consensus. If consensus could not be reached, a third reviewer  
39 (SD.L) was deferred to arbitration and consensus. No language restriction was applied. If a  
40 study was not conducted in the authors' language, a professional translation software could be  
41 used.

### 42 **Data extraction**

1  
2  
3  
4 43 The data were extracted independently by two reviewers (JR.N and PJ.Y) according to a  
5  
6  
7 44 predefined data extraction sheet. The following variables were extracted from the included  
8  
9  
10 45 studies: lead author, publication year, country of study, study design, study population  
11  
12 46 demographics, sample size, mean age, male ratio, time interval between TTE and RHC, cut-off  
13  
14 47 threshold levels for TTE and RHC, and number of true-positive (TP), false-negative (FN), true-  
15  
16 48 negative (TN) and false-positive (FP) observations. Extracted data were cross-checked and  
17  
18  
19 49 disagreements were resolved via discussion or referral to a third reviewer (Y.H).

### 22 50 **Quality assessment**

23  
24  
25 51 The Quality Assessment of Diagnostic Accuracy Studies QUADAS-2 tool was used to assess  
26  
27  
28 52 the risk of bias and clinical applicability concerns of the included studies according to the  
29  
30  
31 53 Cochrane Collaboration recommendation.<sup>18 19</sup> Two reviewers (JR.N and PJ.Y) independently  
32  
33 54 evaluated QUADAS-2 items, and all emerging conflicts were resolved by consensus.

### 36 55 **Data synthesis and statistical analysis**

37  
38  
39 56 Statistical analysis was performed using STATA/SE version 15.1 (Stata Corp, College  
40  
41  
42 57 Station, TX) and Review Manager Version 5.3 software (Copenhagen, Denmark, Nordic  
43  
44 58 Cochrane Centre, Cochrane Collaboration, 2014). All tests were two-tailed. A  $p$  value < 0.05  
45  
46 59 was considered statistically significant.

47  
48  
49 60 The correlation coefficient between the logarithm of sensitivity and logarithm of one minus  
50  
51  
52 61 specificity was calculated to test whether the threshold effect was one of the sources of  
53  
54  
55 62 heterogeneity.<sup>20</sup> Deeks' test was used to test for publication bias.<sup>21</sup> The bivariate model for  
56  
57 63 diagnostic meta-analysis was used to obtain pooled estimates of sensitivity and specificity.<sup>22</sup>  
58  
59

64 Statistical heterogeneity among studies was explored using the  $I^2$  statistic.

65 Pooled sensitivity, specificity, diagnostic odds ratio (DOR), positive likelihood ratio (PLR),  
66 negative likelihood ratio (NLR), and area under the summary receiver operating characteristic  
67 (SROC) curve were calculated from the number of TPs, FNs, FPs, and TNs. The 95%  
68 confidence interval (CI) was estimated for each metric.

69 Subgroup analyses were performed based on the following variables: the time interval  
70 between TTE and RHC; disease classification of the study population; publication year of the  
71 study; study design (prospective or retrospective) and cut-off threshold of TTE to diagnose PH.  
72 Sensitivity analysis was undertaken by excluding low-quality studies (according to the  
73 QUADAS-2 quality assessment) or trials with characteristics different from the others.

## 74 Results

### 75 Studies selection and characteristics

76 Figure 1 shows the PRISMA flow chart of the literature screening. A total of 27 articles  
77 involving 4386 subjects met our inclusion criteria (Table 1).<sup>10-13 23-45</sup> Habash's study was  
78 divided into two independent parts because of the differences between the case group (Habash-1)  
79 and the control group (Habash-2).<sup>27</sup>

80 Of the 27 eligible studies, fourteen (52%) were published between 2010–2019,<sup>10-13 26 27 30 33 34</sup>  
81 <sup>39 41 43-45</sup> and thirteen (48%) were published before 2010.<sup>23-25 28 29 31 32 35-38 40 42</sup> Twelve (44%)  
82 studies were performed in Europe,<sup>12 24 26 32-35 37-39 43 44</sup> nine (30%) in the United States of America  
83 (USA),<sup>10 13 23 25 27 28 31 40 42</sup> two (8%) in East Asia,<sup>29 36</sup> three (12%) in the Middle East,<sup>11 41 45</sup> and  
84 one (4%) in Australia.<sup>30</sup> Most of the studies (15/27, 56%)<sup>11 12 23 24 28-32 35 36 38 39 41 45</sup> were of  
85 prospective design versus 44% (12/27)<sup>10 13 25-27 33 34 37 40 42-44</sup> retrospective.

1  
2  
3  
4 86 All included studies used the tricuspid maximal regurgitation velocity (TRVmax) to estimate  
5  
6 87 SPAP; the majority of these studies (23/27, 85%) used the classical method to calculate SPAP:  
7  
8  
9 88  $4TRV_{max}^2 +$  right atrial pressure (RAP).<sup>10 11 13 23-28 31-37 39-45</sup> The RAP was calculated through  
10  
11 89 the diameter and collapse rate of the inferior vena cava (IVC) during spontaneous respiration  
12  
13  
14 90 in sixteen(59%) studies,<sup>10 23 25-27 31 33 35-37 39-42 44 45</sup> through the jugular vein pressure (JVP) in one  
15  
16  
17 91 study (4%),<sup>24</sup> and using a fixed value (5 or 10 mm Hg) in three studies (11%).<sup>28 32 34</sup> Three  
18  
19 92 studies (11%) did not report their method for calculating RAP.<sup>11 13 43</sup> Four studies (15%) used  
20  
21  
22 93 a tricuspid gradient ( $4TRV_{max}^2$ ) instead of SPAP.<sup>12 29 30 38</sup>

23  
24  
25 94 The majority of the studies (22/27, 81%) reported the time interval (mean or maximum)  
26  
27 95 between TTE and RHC,<sup>10-13 23-29 31-35 38 40-42 44 45</sup> while five (5/9, 19%) did not.<sup>30 36 37 39 43</sup> Nine  
28  
29 96 studies (33%) considered time intervals greater than one week,<sup>10 13 24 25 27 31 38 40 42</sup> while thirteen  
30  
31 97 studies (48%) considered time intervals of less than one week.<sup>11 12 23 26 29 32-35 37 39 41 44</sup> The time  
32  
33  
34 98 interval between TTE and RHC ranged from four hours to three months.

### 35 36 37 38 99 **Quality Assessment**

39  
40  
41 100 The quality assessment of the included studies according to the QUADAS-2 inventory is  
42  
43 101 shown in Figure 2. Overall, the quality of the included studies was modest. The included studies  
44  
45 102 were of good quality regarding the applicability concerns, but most of them were of low quality  
46  
47 103 in the risk of bias. In twenty (74%) study protocols,<sup>10-13 23 24 26 28-32 34 35 37-39 41 44 45</sup> consecutive  
48  
49 104 subjects were enrolled, with no inappropriate exclusions. The risk of bias during patient  
50  
51 105 recruitment was unclear in the remaining seven (26%) studies,<sup>25 27 33 36 40 42 43</sup> as patient  
52  
53 106 recruitment was not reported. In six (22%) studies investigators designed the single-blind  
54  
55 107 methods for TTE.<sup>10 12 23 26 39 45</sup> Double blinding in imaging assessment was not mentioned in any  
56  
57  
58  
59

1  
2  
3  
4 108 study. The risk of bias on flow and timing between the index test and reference standard was  
5  
6 109 categorized as unclear in 14 (52%) study protocols that did not explicitly state the successful  
7  
8  
9 110 investigation with both index and reference tests in all included patients.<sup>24 30-40 42 43</sup>

### 111 **Quantitative Analysis**

112 The SROC curve for TTE is shown in Figure 3. Four studies fall within the 95% confidence  
113 interval.<sup>11 26 34 44</sup> The AUC was 0.88 (95%CI 0.85–0.90). The pooled sensitivity and specificity  
114 for TTE were 85% (95%CI 81%–90%) and 74% (95%CI 64%–81%), respectively (Figure 4).  
115 The pooled PLR and NLR were 3.2 (95%CI 2.3–4.4) and 0.20 (95%CI 0.15–0.26), respectively.  
116 The pooled DOR for TTE was 16 (95%CI 10–27).

117 The heterogeneity in our study was significant. The threshold test proved that the threshold  
118 effect was not the source of heterogeneity ( $r=0.34$ ,  $P=0.12$ ). Deeks' test for funnel plot  
119 asymmetry suggested no publication bias ( $P=0.69$ ). The results of the subgroup analysis are  
120 presented in Table 2. The sensitivity (87%, 95%CI 81%–91%), specificity (74%, 95%CI 62%–  
121 83%) and AUC (0.89, 95%CI 0.86–0.91) of TTE to diagnose PH were higher for studies  
122 published in 2010 and later compared to those published before 2010. Among the time interval  
123 subgroups, the group with the shortest time interval between TTE and RHC had the best  
124 diagnostic effect, with sensitivity, specificity and AUC of 88% (95%CI 73%–95%), 90%  
125 (95%CI 53%–99%) and 0.94(95%CI 0.92–0.96), respectively. The disease composition of the  
126 study population also affected the diagnostic accuracy of TTE. Compared with patients with  
127 other diseases, TTE had lower sensitivity (81%, 95%CI 70%–88%), specificity (61%, 95%CI  
128 53%–69%) and AUC (0.73, 95%CI 0.69–0.77) in the subgroup of patients with definite lung  
129 diseases.



1  
2  
3  
4 130 Subgroup analysis of different cut-off thresholds to diagnose PH based on TTE showed that  
5  
6 131 the subgroup with a cut-off threshold of 35 mmHg had a higher diagnostic accuracy than that  
7  
8  
9 132 at 40 mmHg. The sensitivity, specificity and AUC of the former were respectively 92% (95%CI  
10  
11 133 88%–94%), 65% (95%CI 43%–83%) and 0.92 (95%CI 0.89–0.94), while the sensitivity,  
12  
13 134 specificity and AUC at 40 mmHg were 84% (95%CI 75%–91%), 52% (95%CI 31%–71%) and  
14  
15 135 0.80 (95%CI 76%–83%), respectively.

16  
17  
18  
19 136 The sensitivity analysis results are shown in Table 3. After excluding low-quality studies and  
20  
21 137 studies with specific characteristics, the sensitivity analysis did not reveal a source for the  
22  
23 138 heterogeneity in the diagnostic accuracy analysis. Overall, the pooled meta-analysis results for  
24  
25 139 outcomes were in accordance to our sensitivity analyses.

## 30 140 **Discussion**

31  
32 141 Our study found that TTE has a better sensitivity but moderate specificity for the detection  
33  
34 142 of PH. In addition, shortening the time interval between TTE and RHC and developing an  
35  
36 143 appropriate threshold could improve the accuracy of TTE. However, the accuracy of TTE to  
37  
38 144 diagnose PH in patients with lung diseases was low.

39  
40  
41  
42 145 Although PH is a chronic disease, we still believe that the shortest possible time interval  
43  
44 146 between TTE and RHC is more favorable. Otherwise, changes in the patient's condition and the  
45  
46 147 application of intervention measures would lead to an increase in the deviation of the results of  
47  
48 148 the two examinations. A detailed subgroup analysis was performed according to the time  
49  
50 149 interval between TTE and RHC. As expected, the diagnostic accuracy was the highest when  
51  
52 150 the time interval was less than or equal to 24 hours. The results also showed that the efficacy  
53  
54 151 of TTE in the diagnosis of PH was gradually reduced with the extension of the time interval.

1  
2  
3  
4 152 Subgroup analysis based on the disease composition of the population suggested that the  
5  
6 153 diagnostic accuracy of TTE was lower in patients with lung diseases. Changes associated with  
7  
8  
9 154 chronic pulmonary disease, including a marked increase in intrathoracic gas, consolidation of  
10  
11  
12 155 lung tissue, expansion of the thoracic cage, and alterations in the position of the heart, adversely  
13  
14 156 affect the imaging quality and the parameter measurement of TTE.<sup>46</sup> Therefore, the use of TTE  
15  
16  
17 157 to measure pulmonary pressure in patients with lung diseases might not be an ideal choice.

18  
19 158 The Guideline recommend the use of IVC width and collapse rate to estimate RAP,<sup>3</sup> which  
20  
21  
22 159 was not used in some of the included studies. The sensitivity analysis for this point showed that  
23  
24  
25 160 studies which calculated RAP through IVC do not seem to have a higher diagnostic  
26  
27  
28 161 performance. In order to avoid errors caused by RAP estimation, TRVmax was also considered  
29  
30  
31 162 as an indicator to evaluate the possibility of PH. Four studies using TRPG (4TRVmax2) instead  
32  
33  
34 163 of SPAP were grouped into a subgroup and showed that this subgroup had good diagnostic  
35  
36  
37 164 specificity but poor sensitivity.

38 165 The sensitivity analysis based on the mean pulmonary artery pressure (MPAP) threshold of  
39  
40  
41 166 25 mmHg did not result in a higher diagnostic value than the whole, indicating that the overall  
42  
43  
44 167 results were stable. A previous work suggested that a MPAP threshold of 25 mmHg is arbitrarily  
45  
46  
47 168 chosen and lowering it to 20 mmHg (two standard deviations higher than MPAP for the  
48  
49  
50 169 population) is considered a scientific method.<sup>47</sup> However, some scientists insist that it is  
51  
52  
53 170 premature to reduce the MPAP threshold to 20 mmHg because of the risk of over-diagnosis,  
54  
55  
56 171 unclear treatment implications and additional psychological burden on patients.<sup>48</sup> Since none  
57  
58  
59 172 of the study we included used MPAP>20mmHg as the diagnostic threshold for RHC, subgroup  
60  
173 analysis on the two thresholds of 20mmHg and 25mmHg could not be performed. Therefore,

1  
2  
3  
4 174 we expect that more studies may be performed in the future to verify the appropriate threshold  
5  
6 175 of RHC.  
7  
8

9 176 In our review, the cut-off thresholds of SPAP ranged from 30 to 50 mmHg. Subgroup  
10  
11 177 analysis showed that the diagnostic accuracy of the group of 35 mmHg was higher. Sensitivity  
12  
13 178 analysis results of studies that excluded high TTE cut-off value showed that a high cut-off value  
14  
15 179 increased the specificity and reduced the sensitivity of TTE. Due to the small sample size of the  
16  
17 180 subgroup in this study, the value of the cut-off threshold still needs to be determined by further  
18  
19 181 prospective studies of multi-center and large samples.  
20  
21  
22  
23

24 182 Subgroup analysis according to the publication year confirmed that studies published after  
25  
26 183 2010 had only a slightly higher diagnostic accuracy than previous studies. With the  
27  
28 184 improvement of TTE technology and instruments in the past ten years, the diagnostic efficiency  
29  
30 185 of PH has not been significantly improved, which forces us to pay attention to other TTE  
31  
32 186 parameters.<sup>49 50</sup> Perhaps, this could be a new direction for future studies on PH diagnosis.  
33  
34  
35  
36

37 187 **Limitations:**

38  
39  
40 188 Several limitations are present in our study. Firstly, the systematic review and meta-analysis  
41  
42 189 is a secondary research method based on original research and the quality of the included study  
43  
44 190 affects the results. In addition, the possibility of missing relevant articles objectively exists, and  
45  
46 191 significant heterogeneity may limit the interpretation of the results. Secondly, the accuracy of  
47  
48 192 echocardiography relies heavily on the operator's ability, experience, and operational discipline.  
49  
50 193 In order to obtain more original studies, we did not consider this aspect as an exclusion criterion.  
51  
52 194 Thirdly, the studies included in this review involve several different types of PH, and some of  
53  
54 195 the included studies do not describe the basic disease and PH type in detail. It is clear that  
55  
56  
57  
58  
59  
60

1  
2  
3  
4 196 pulmonary lesions can affect the quality of TTE imaging, leading to underestimated results.  
5

6  
7 197 **Conclusion**  
8

9 198 TTE has clinical value in the diagnosis of PH thanks to its better sensitivity and moderate  
10  
11 199 specificity, but it cannot yet replace RHC considered as the gold standard. Shortening the time  
12  
13 200 interval between TTE and RHC and developing an appropriate threshold can improve the  
14  
15 201 accuracy of TTE. TTE may not be suitable to assess pulmonary arterial pressure in patients with  
16  
17 202 pulmonary disease. It may be necessary to combine multiple TTE parameters and conduct  
18  
19 203 multi-center, large-sample studies to further improve the accuracy of TTE in the diagnosis of  
20  
21 204 PH in future research.  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

### Author's contribution

The joint corresponding authors (JQ.L and B.S) are responsible for the design and implementation of the study. SD.L is responsible for the quality control of study selection. Y.H performed the quality control on the links of data extraction. KH.Y provided guidance in literature retrieval and data processing methodology and was responsible for the quality evaluation part. JR.N and PJ.Y performed the systematic review of the literature and extracted the data. JR.N conducted the meta-analyses, and two authors (JR.N, PJ.Y) substantially contributed to the interpretation of the data and wrote the article. All authors repeatedly revised the article. The corresponding authors (JQ.L and B.S) and JR.N take responsibility for the integrity of the analyses.

## Reference

1. Hoeper MM, Humbert M, Souza R, et al. A global view of pulmonary hypertension. *Lancet Respir Med* 2016;4:306-22.
2. Galie N, Rubin L, Hoeper M, et al. Treatment of patients with mildly symptomatic pulmonary arterial hypertension with bosentan (EARLY study): a double-blind, randomised controlled trial. *Lancet* 2008;371:2093-100.
3. Galie N, Humbert M, Vachiery JL, et al. 2015 ESC/ERS Guidelines for the diagnosis and treatment of pulmonary hypertension. *Kardiol Pol* 2015;73:1127-206.
4. Hoeper MM, Lee SH, Voswinckel R, et al. Complications of right heart catheterization procedures in patients with pulmonary hypertension in experienced centers. *J Am Coll Cardiol* 2006;48:2546-52.
5. Tian J, Zhang J, Ge L, et al. The methodological and reporting quality of systematic reviews from China and the USA are similar. *J Clin Epidemiol* 2017;85:50-58.
6. Xiu-xia L, Ya Z, Yao-long C, et al. The reporting characteristics and methodological quality of Cochrane reviews about health policy research. *Health Policy* 2015;119:503-10.
7. Janda S, Shahidi N, Gin K, et al. Diagnostic accuracy of echocardiography for pulmonary hypertension: a systematic review and meta-analysis. *Heart* 2011;97:612-22.
8. Taleb M, Khuder S, Tinkel J, et al. The diagnostic accuracy of Doppler echocardiography in assessment of pulmonary artery systolic pressure: a meta-analysis. *Echocardiography* 2013;30:258-65.

- 1  
2  
3  
4 9. Zhang RF, Zhou L, Ma GF, et al. Diagnostic value of transthoracic Doppler  
5  
6 echocardiography in pulmonary hypertension: a meta-analysis. *Am J Hypertens*  
7  
8 2010;23:1261-64.  
9
- 10  
11 10. Ahmed M, Elshinawy O, Agmy G, et al. Tricuspid regurgitation velocity versus right  
12  
13 ventricular systolic pressure in the echocardiographic evaluation of pulmonary  
14  
15 hypertension. *Egypt J Chest Dis Tuberc* 2019;68:203-08.  
16  
17
- 18  
19 11. Balci MK, Ari E, Vayvada M, et al. Assessment of Pulmonary Hypertension in Lung  
20  
21 Transplantation Candidates: Correlation of Doppler Echocardiography With Right  
22  
23 Heart Catheterization. *Transplant Proc* 2016;48:2797-802.  
24  
25
- 26  
27 12. Schneider M, Pistrutto AM, Gerges C, et al. Multi-view approach for the diagnosis of  
28  
29 pulmonary hypertension using transthoracic echocardiography. *Int J Cardiovasc*  
30  
31 *Imaging* 2018;34:695-700.  
32  
33
- 34  
35 13. Shujaat A, Bajwa AA, Al-Saffar F, et al. Diagnostic accuracy of echocardiography  
36  
37 combined with chest CT in pulmonary hypertension. *Clin Respir J* 2018;12:948-52.  
38  
39
- 40  
41 14. Moher D, Liberati A, Tetzlaff J, et al. Preferred reporting items for systematic reviews  
42  
43 and meta-analyses: the PRISMA statement. *BMJ* 2009;339:b2535.  
44  
45
- 46  
47 15. Shamseer L, Moher D, Clarke M, et al. Preferred reporting items for systematic review  
48  
49 and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. *BMJ*  
50  
51 2015;350:g7647.  
52  
53
- 54  
55 16. Ge L, Tian JH, Li YN, et al. Association between prospective registration and overall  
56  
57 reporting and methodological quality of systematic reviews: a meta-epidemiological  
58  
59 study. *J Clin Epidemiol* 2018;93:45-55.  
60

17. Wang X, Chen Y, Yao L, et al. Reporting of declarations and conflicts of interest in WHO guidelines can be further improved. *J Clin Epidemiol* 2018;98:1-8.
18. Whiting P, Rutjes AW, Dinnes J, et al. Development and validation of methods for assessing the quality of diagnostic accuracy studies. *Health Technol Assess* 2004;8:iii, 1-234.
19. Whiting P, Rutjes AW, Reitsma JB, et al. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Med Res Methodol* 2003;3:25.
20. Lee J, Kim KW, Choi SH, et al. Systematic Review and Meta-Analysis of Studies Evaluating Diagnostic Test Accuracy: A Practical Review for Clinical Researchers- Part II. Statistical Methods of Meta-Analysis. *Korean J Radiol* 2015;16:1188-96.
21. Deeks JJ, Macaskill P, Irwig L. The performance of tests of publication bias and other sample size effects in systematic reviews of diagnostic test accuracy was assessed. *J Clin Epidemiol* 2005;58:882-93.
22. Reitsma JB, Glas AS, Rutjes AW, et al. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. *J Clin Epidemiol* 2005;58:982-90.
23. Arcasoy SM, Christie JD, Ferrari VA, et al. Echocardiographic assessment of pulmonary hypertension in patients with advanced lung disease. *Am J Respir Crit Care Med* 2003;167:735-40.
24. Denton CP, Cailes JB, Phillips GD, et al. Comparison of Doppler echocardiography and right heart catheterization to assess pulmonary hypertension in systemic sclerosis.



- 1  
2  
3  
4 Br J Rheumatol 1997;36:239-43.  
5  
6  
7 25. Fisher MR, Criner GJ, Fishman AP, et al. Estimating pulmonary artery pressures by  
8  
9 echocardiography in patients with emphysema. Eur Respir J 2007;30:914-21.  
10  
11  
12 26. Greiner S, Jud A, Aurich M, et al. Reliability of noninvasive assessment of systolic  
13  
14 pulmonary artery pressure by Doppler echocardiography compared to right heart  
15  
16 catheterization: analysis in a large patient population. J Am Heart Assoc 2014;3:1-8.  
17  
18  
19 27. Habash F, Gurram P, Almomani A, et al. Correlation between echocardiographic  
20  
21 pulmonary artery pressure estimates and right heart catheterization measurement in  
22  
23 liver transplant candidates. J Cardiovasc Imaging 2018;26:75-84.  
24  
25  
26  
27 28. Hsu VM, Moreyra AE, Wilson AC, et al. Assessment of pulmonary arterial  
28  
29 hypertension in patients with systemic sclerosis: comparison of noninvasive tests with  
30  
31 results of right-heart catheterization. J Rheumatol 2008;35:458-65.  
32  
33  
34  
35 29. Hua R, Sun YW, Wu ZY, et al. Role of 2-dimensional Doppler echo-cardiography in  
36  
37 screening portopulmonary hypertension in portal hypertension patients. Hepatobiliary  
38  
39 Pancreat. Hepatobiliary Pancreat Dis Int 2009;8:157-61.  
40  
41  
42  
43 30. Keir GJ, Wort SJ, Kokosi M, et al. Pulmonary hypertension in interstitial lung disease:  
44  
45 Limitations of echocardiography compared to cardiac catheterization. Respirology  
46  
47 2018;23:687-94.  
48  
49  
50  
51 31. Kim WR, Krowka MJ, Plevak DJ, et al. Accuracy of doppler echocardiography in the  
52  
53 assessment of pulmonary hypertension in liver transplant candidates. Liver Transpl  
54  
55 2000;6:453-58.  
56  
57  
58  
59 32. Laaban JP, Diebold B, Zelinski R, et al. Noninvasive Estimation of Systolic  
60

- 1  
2  
3  
4 Pulmonary-Artery Pressure Using Doppler Echocardiography in Patients with Chronic  
5  
6 Obstructive Pulmonary-Disease. *Chest* 1989;96:1258-62.  
7  
8  
9 33. Lafitte S, Pillois X, Reant P, et al. Estimation of pulmonary pressures and diagnosis of  
10  
11 pulmonary hypertension by doppler echocardiography: A retrospective comparison of  
12  
13 routine echocardiography and invasive hemodynamics. *J Am Soc Echocardiogr*  
14  
15 2013;26:457-63.  
16  
17  
18 34. Lange TJ, Baumgartner S, Arzt M, et al. Qualitative echocardiography parameters for  
19  
20 prediction of pulmonary hypertension. *Int J Clin Pract* 2013;67:5-12.  
21  
22  
23 35. Lanzarini L, Fontana A, Campana C, et al. Two simple echo-doppler measurements  
24  
25 can accurately identify pulmonary hypertension in the large majority of patients with  
26  
27 chronic heart failure. *J Heart Lung Transplant* 2005;24:745-54.  
28  
29  
30 36. Matsuyama W, Ohkubo R, Michizono K, et al. Usefulness of transcutaneous Doppler  
31  
32 jugular venous echo to predict pulmonary hypertension in COPD patients. *Eur Respir*  
33  
34 *J* 2001;17:1128-31.  
35  
36  
37 37. Mogollon Jimenez MV, Escorcesca Ortega AM, Cabeza Letran ML, et al. Correlation  
38  
39 of Echocardiographic and Hemodynamic Parameters in Pulmonary Hypertension  
40  
41 Assessment Prior to Heart Transplantation. *Transplant Proc* 2008;40:3023-24.  
42  
43  
44 38. Mukerjee D, St George D, Knight C, et al. Echocardiography and pulmonary function  
45  
46 as screening tests for pulmonary arterial hypertension in systemic sclerosis.  
47  
48 *Rheumatology (Oxford)* 2004;43:461-6.  
49  
50  
51 39. Nagel C, Henn P, Ehlken N, et al. Stress Doppler echocardiography for early detection  
52  
53 of systemic sclerosis-associated pulmonary arterial hypertension. *Arthritis Res Ther*  
54  
55  
56  
57  
58  
59  
60

- 2015;17:165.
40. Nathan SD, Shlobin OA, Barnett SD, et al. Right ventricular systolic pressure by echocardiography as a predictor of pulmonary hypertension in idiopathic pulmonary fibrosis. *Respir Med* 2008;102:1305-10.
41. Parsaee M, Amin A, Nematollahi MR, et al. Comparison of transthoracic echocardiography and right heart catheterization for assessing pulmonary arterial pressure in patients with congenital or valvular heart defects. *Iran Heart J* 2012;12:54-61.
42. Penning S, Robinson KD, Major CA, et al. A comparison of echocardiography and pulmonary artery catheterization for evaluation of pulmonary artery pressures in pregnant patients with suspected pulmonary hypertension. *Am J Obstet Gynecol* 2001;184:1568-70.
43. Raevens S, Colle I, Reyntjens K, et al. Echocardiography for the detection of portopulmonary hypertension in liver transplant candidates: An analysis of cutoff values. *Liver Transpl* 2013;19:602-10.
44. Rajaram S, Swift AJ, Capener D, et al. Comparison of the diagnostic utility of cardiac magnetic resonance imaging, computed tomography, and echocardiography in assessment of suspected pulmonary arterial hypertension in patients with connective tissue disease. *J Rheumatol* 2012;39:1265-74.
45. Sohrabi B, Kazemi B, Mehryar A, et al. Correlation between pulmonary artery pressure measured by echocardiography and right heart catheterization in patients with rheumatic mitral valve stenosis (A prospective study). *Echocardiography* 2016;33:7-

- 1  
2  
3  
4 13.  
5  
6  
7 46. Galderisi M, Cosyns B, Edvardsen T, et al. Standardization of adult transthoracic  
8  
9 echocardiography reporting in agreement with recent chamber quantification, diastolic  
10  
11 function, and heart valve disease recommendations: an expert consensus document of  
12  
13 the European Association of Cardiovascular Imaging. *Eur Heart J Cardiovasc Imaging*  
14  
15 2017;18:1301-10.  
16  
17  
18  
19 47. Simonneau G, Montani D, Celermajer DS, et al. Haemodynamic definitions and  
20  
21 updated clinical classification of pulmonary hypertension. *Eur Respir J* 2019;53.  
22  
23  
24 48. Gibbs JSR, Torbicki A. Proposed new pulmonary hypertension definition: is 4 mm(Hg)  
25  
26 worth re-writing medical textbooks? *Eur Respir J* 2019;53.  
27  
28  
29 49. Misra D, Kendes A, Sulica R, et al. Exercise-induced pulmonary hypertension by stress  
30  
31 echocardiography: Prevalence and correlation with right heart hemodynamics. *Int J*  
32  
33 *Cardiol* 2017;228:518-22.  
34  
35  
36  
37 50. Vitarelli A, Mangieri E, Terzano C, et al. Three-dimensional echocardiography and  
38  
39 2D-3D speckle-tracking imaging in chronic pulmonary hypertension: diagnostic  
40  
41 accuracy in detecting hemodynamic signs of right ventricular (RV) failure. *J Am Heart*  
42  
43 *Assoc* 2015;4:e001584.  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

**Table 1 Characteristics of each study included in this meta-analysis.**

Study	Year	Country	Design	N	Disease Composition of the Population	Mean Age (Years)	Male (%)	Time Interval	TTE Threshold (mmHg)	RHC Threshold (mmHg)	TTE Method
<b>Ahmed Keir</b>	2019	USA	Retrospective	136	Multiple diseases	59±20	35	<3mo	SPAP≥40	MPAP≥25	4TRVmax <sup>2</sup> +RAP(IVC)
<b>Habash-1</b>	2018	Australia	Prospective	265	Interstitial lung disease	60.8±16.5	46	–	TRPG>46	MPAP≥25	4TRVmax <sup>2</sup>
<b>Habash-2</b>	2018	USA	Retrospective	31	Liver transplantation candidates	57±11	42	36.8±13.4d	SPAP>47	MPAP≥25	4TRVmax <sup>2</sup> +RAP(IVC)
<b>Schneider</b>	2018	Austria	Prospective	65	Cardiac and lung diseases	67.2	43	<48h	TRPG>32	MPAP≥25	4TRVmax <sup>2</sup>
<b>Balci</b>	2016	Turkey	Prospective	103	Lung transplantation candidates	47.6±10.4	66	<72h	SPAP>35	MPAP≥25	4TRVmax <sup>2</sup> +RAP (NR)
<b>Shujaat</b>	2016	USA	Retrospective	87	Multiple diseases	54.3±15.9	29	13d*	SPAP>40	MPAP>25	4TRVmax <sup>2</sup> +RAP (NR)
<b>Sohrabi</b>	2016	Iran	Prospective	300	Rheumatic mitral stenosis	59.9	31	<24h	SPAP≥35	MPAP≥25	4TRVmax <sup>2</sup> +RAP(IVC)
<b>Nagel</b>	2015	Germany	Prospective	76	Systemic sclerosis	58±14	16	–	SPAP>40	MPAP≥25	4TRVmax <sup>2</sup> +RAP(IVC)
<b>Greiner</b>	2014	Germany	Retrospective	1695	Cardiac disease	63±15	67	<5d	SPAP≥36	MPAP≥25	4TRVmax <sup>2</sup> +RAP(IVC)
<b>Lafitte</b>	2013	France	Retrospective	114	Cardiac and lung disease	64.8±15.9	52	<48h	SPAP≥38	MPAP>25	4TRVmax <sup>2</sup> +RAP(IVC)
<b>Lange</b>	2013	Germany	Retrospective	231	Multiple diseases	62±13	43	5±4d	SPAP>50	MPAP≥25	4TRVmax <sup>2</sup> +RAP (5)
<b>Raevens</b>	2013	Belgium	Retrospective	152	Liver transplantation candidates	58±11	66	–	SPAP>38	MPAP≥25	4TRVmax <sup>2</sup> +RAP (NR)
<b>Parsaee</b>	2012	Iran	Prospective	103	Cardiac diseases	41.0±15.8	44	<4h	SPAP≥35	MPAP>25	4TRVmax <sup>2</sup> +RAP(IVC)
<b>Rajaram</b>	2012	UK	Retrospective	81	Connective tissue disease	62±14	15	<48h	TRPG≥40	MPAP≥25	4TRVmax <sup>2</sup>
<b>Hua</b>	2009	China	Prospective	105	Liver transplantation candidates	49.5±11.8	63	4.2±2.0d	SPAP≥30	MPAP≥25	4TRVmax <sup>2</sup> +RAP(IVC)
<b>Nathan</b>	2008	USA	Retrospective	60	Idiopathic pulmonary fibrosis	62.9±8.6	55	32±78d	SPAP≥40	MPAP>25	4TRVmax <sup>2</sup> +RAP(IVC)
<b>Hsu</b>	2008	USA	Prospective	49	Systemic Sclerosis	55	18	<4h	SPAP>47	MPAP≥25	4TRVmax <sup>2</sup> +RAP (10)
<b>Mogollon</b>	2008	Spain	Retrospective	67	Heart transplantation candidates	–	–	–	SPAP>40	MPAP>35	4TRVmax <sup>2</sup> +RAP(IVC)
<b>Fisher</b>	2007	USA	Retrospective	63	Emphysema patients	65.6±6.6	60	23d	SPAP>40	MPAP≥25	4TRVmax <sup>2</sup> +RAP(IVC)
<b>Lanzarini</b>	2005	Italy	Prospective	57	Heart failure	52±11	74	<24h	SPAP≥32	SPAP≥35	4TRVmax <sup>2</sup> +RAP(IVC)

<b>Mukerjee</b>	2004	UK	Prospective	137	Systemic sclerosis	63	–	<3 mo	TRPG>40	MPAP≥25	4TRVmax <sup>2</sup>
<b>Arcasoy</b>	2003	USA	Prospective	166	COPD 68%, ILD 28%, PVD 4%	51	43	<72h	SPAP≥45	SPAP≥45	4TRVmax <sup>2</sup> +RAP(IVC)
<b>Penning</b>	2001	USA	Retrospective	27	Pregnant women with cardiac diseases	28.6	0	25.8d	SPAP≥40	SPAP≥35	4TRVmax <sup>2</sup> +RAP(IVC)
<b>Matsuyama</b>	2001	Japan	Prospective	35	COPD	66	94	–	SPAP≥40	MPAP>25	4TRVmax <sup>2</sup> +RAP(IVC)
<b>Kim</b>	2000	USA	Prospective	74	Liver transplantation candidates	54	50	59d	SPAP>50	MPAP≥35	4TRVmax <sup>2</sup> +RAP(JVP)
<b>Denton</b>	1997	UK	Prospective	20	COPD	48.6±11.7	30	1.8±2.3mo	SPAP≥30	SPAP≥30	4TRVmax <sup>2</sup> +RAP(JVP)
<b>Laaban</b>	1989	France	Prospective	27	COPD	63±9	78	<2d	SPAP≥35	SPAP≥35	4TRVmax <sup>2</sup> +RAP(5)

USA, United States of America; UK, United Kingdom of Great Britain and Northern Ireland; TTE, Transthoracic echocardiography; RHC, right heart catheterization; SPAP, systolic pulmonary artery pressure; MPAP, mean pulmonary artery pressure; TRPG, tricuspid regurgitation pressure gradient; TRV, tricuspid regurgitation velocity; RAP, right atrial pressure; IVC, Inferior vena cava; JVP, jugular vein pressure; COPD, chronic obstructive pulmonary disease; ILD, Interstitial lung disease; PVD, peripheral vascular disease; NR, not reported.

\* The median time (other terms are mean time)

**Table 2 Subgroup analysis**

Group	N	<i>I</i> <sup>2</sup> (95%CI)	AUC (95%CI)	Sensitivity (95%CI)	Specificity (95%CI)	PLR (95%CI)	NLR (95%CI)	DOR (95%CI)
<b>All Studies</b>	28	98(97–99)	0.88(0.85–0.90)	0.85(0.81–0.90)	0.74(0.64–0.81)	3.2(2.3–4.4)	20(0.15–0.26)	16(10–27)
<b>Time Interval</b>								
≤24h	4	95(90–99)	0.94(0.92–0.96)	0.88(0.73–0.95)	0.90(0.53–0.99)	8.9(1.5–54.5)	13(0.06–0.29)	68(13–348)
≤48h*	7	95(90–99)	0.94(0.91–0.95)	0.88(0.81–0.93)	0.89(0.71–0.96)	7.8(2.8–21.3)	13(0.09–0.21)	59(23–148)
≤72h*	9	94(89–99)	0.91(0.89–0.93)	0.87(0.82–0.91)	0.83(0.65–0.93)	5.2(2.4–11.2)	15(0.11–0.21)	34(14–82)
≤1 week	13	93(87–99)	0.91(0.88–0.93)	0.87(0.84–0.90)	0.80(0.68–0.88)	4.3(2.7–6.9)	16(0.12–0.21)	27(15–48)
>1 week	10	97(95–99)	0.82(0.78–0.85)	0.85(0.73–0.92)	0.60(0.40–0.77)	2.1(1.3–3.4)	25(0.14–0.45)	9(4–21)
Unclear	5	82(63–100)	0.85(0.81–0.88)	0.79(0.63–0.99)	0.76(0.61–0.87)	3.4(1.9–5.9)	27(0.15–0.51)	12(5–33)
<b>Population Disease</b>								
Cardiac diseases	6	94(89–99)	0.90(0.87–0.92)	0.90(0.86–0.93)	0.67(0.29–0.91)	2.7(0.9–8.1)	15(0.08–0.30)	18(3–95)
Lung diseases	8	90(81–100)	0.73(0.69–0.77)	0.81(0.70–0.88)	0.61(0.53–0.69)	2.1(1.8–2.4)	32(0.21–0.48)	7(4–10)
Multiple diseases <sup>#</sup>	6	93(87–99)	0.90(0.87–0.92)	0.89(0.84–0.92)	0.70(0.40–0.89)	3.0(1.3–7.1)	16(0.11–0.23)	19(6–60)
Unclear <sup>&amp;</sup>	8	88(77–100)	0.88(0.85–0.90)	0.80(0.64–0.90)	0.85(0.80–0.89)	5.3(4.0–7.0)	23(0.12–0.45)	23(10–51)
<b>Published Year</b>								
≥2010	15	97(95–99)	0.89(0.86–0.91)	0.87(0.81–0.91)	0.74(0.62–0.83)	3.3(2.3–4.9)	18(0.13–0.25)	19(11–13)
<2010	13	96(93–99)	0.86(0.83–0.89)	0.84(0.74–0.90)	0.73(0.56–0.85)	3.1(1.8–5.3)	22(0.14–0.37)	14(6–33)
<b>Study Design</b>								
Prospective	15	97(95–99)	0.90(0.87–0.92)	0.86(0.77–0.91)	0.79(0.69–0.87)	4.2(2.7–6.4)	18(0.11–0.28)	23(12–45)
Retrospective	13	96(92–99)	0.86(0.83–0.89)	0.86(0.80–0.90)	0.65(0.49–0.78)	2.5(1.6–3.7)	22(0.15–0.32)	11(6–22)
<b>TTE Threshold</b>								

25

SPAP $\geq$ 40 mmHg	8	96(93–99)	0.80(0.76–0.83)	0.84(0.75–0.91)	0.52(0.31–0.71)	1.7(1.2–2.5)	0.30(0.21–0.44)	6(3–11)
SPAP $\geq$ 35 mmHg	4	76(47–100)	0.92(0.89–0.94)	0.92(0.88–0.94)	0.65(0.43–0.83)	2.6(1.4–4.9)	0.13(0.08–0.22)	16(9–28)
TRPG	4	0(0–100)	0.85(0.82–0.88)	0.75(0.58–0.86)	0.81(0.70–0.89)	4.0(2.2–7.3)	0.31(0.17–0.57)	13(4–40)

AUC, Area under the curve; PLR, positive likelihood ratio; NLR, negative likelihood ratio; DOR, diagnostic odds ratio; TTE, transthoracic echocardiography; SPAP, systolic pulmonary artery pressure; TRPG, tricuspid regurgitation pressure gradient.

\*Studies with time intervals less than or equal to 24 hours were included in this subgroup.

\* Studies with time intervals less than or equal to 24 hours and 48 hours were included in this subgroup.

# Studies including a variety of diseases, including heart disease and lung disease.

&Diseases were not specifically identified in the studies (transplant candidates).



BMJ Open: first published as 10.1136/bmjopen-2019-033084 on 22 December 2019. Downloaded from <http://bmjopen.bmj.com/> on April 17, 2024 by guest. Protected by copyright.

**Table 3 Sensitivity analysis for diagnostic accuracy meta-analysis.**

Study characteristic	N	<i>I</i> <sup>2</sup> (95%CI)	AUC (95%CI)	Sensitivity (95%CI)	Specificity (95%CI)	PLR (95%CI)	NLR (95%CI)	DOR (95%CI)
All included studies	28	98(97–99)	0.88(0.85–0.90)	0.85(0.81–0.90)	0.74(0.64–0.81)	3.2(2.3–4.4)	0.20(0.15–0.26)	16(10–27)
Excluding study of Penning*	27	98(97–99)	0.88(0.85–0.91)	0.86(0.81–0.89)	0.75(0.66–0.82)	3.4(2.5–4.6)	0.19(0.14–0.26)	18(11–28)
RHC threshold MPAP≥25 mmHg	21	98(97–99)	0.87(0.84–0.90)	0.83(0.77–0.88)	0.76(0.67–0.83)	3.5(2.5–4.8)	0.22(0.16–0.30)	16(10–26)
RAP method(IVC) <sup>‡</sup>	17	96(93–99)	0.89(0.86–0.91)	0.86(0.82–0.90)	0.73(0.59–0.84)	3.2(2.0–5.1)	0.19(0.13–0.27)	17(8–35)
Excluding high TTE threshold*	21	97(95–99)	0.90(0.87–0.92)	0.88(0.85–0.91)	0.72(0.59–0.82)	3.2(2.1–4.8)	0.16(0.12–0.22)	20(11–36)

AUC, Area under the curve; PLR, positive likelihood ratio; NLR, negative likelihood ratio; DOR, diagnostic odds ratio; MPAP, mean pulmonary artery pressure; TTE, transthoracic echocardiography; RHC, right heart catheterization; RAP, right atrial pressure; IVC, inferior vena cava.

\* The study of Penning was excluded because only pregnant women with cardiac disease were included as subjects.

‡The RAP was calculated through the diameter and collapse rate of IVC during spontaneous respiration. Habashi's study was divided into two independent parts, thus the results section showed 16 studies, but 17 sets of data were analyzed.

\* High TTE threshold was defined as SPAP greater than 45 mmHg or tricuspid regurgitation pressure gradient (TRPG) greater than 40 mmHg.

## Figure legends/captions

**Figure 1** Flowchart for identification of the studies.

Habash's study was divided into two independent parts because of the differences between the case group (Habash-1) and the control group (Habash-2). A total of 27 studies were included, but 28 sets of data were analyzed.

**Figure 2** Risk of bias and applicability concerns summary: review authors' judgements regarding each domain for each included study (n=28).

**Figure 3** Summary receiver operating characteristic (SROC) graph with 95% confidence region and 95% prediction region for TTE in the diagnosis of pulmonary hypertension (n=28).

**Figure 4** Forest plot of the sensitivity and specificity of each individual study, summary sensitivity and specificity and  $I^2$  statistic for heterogeneity (n=28).

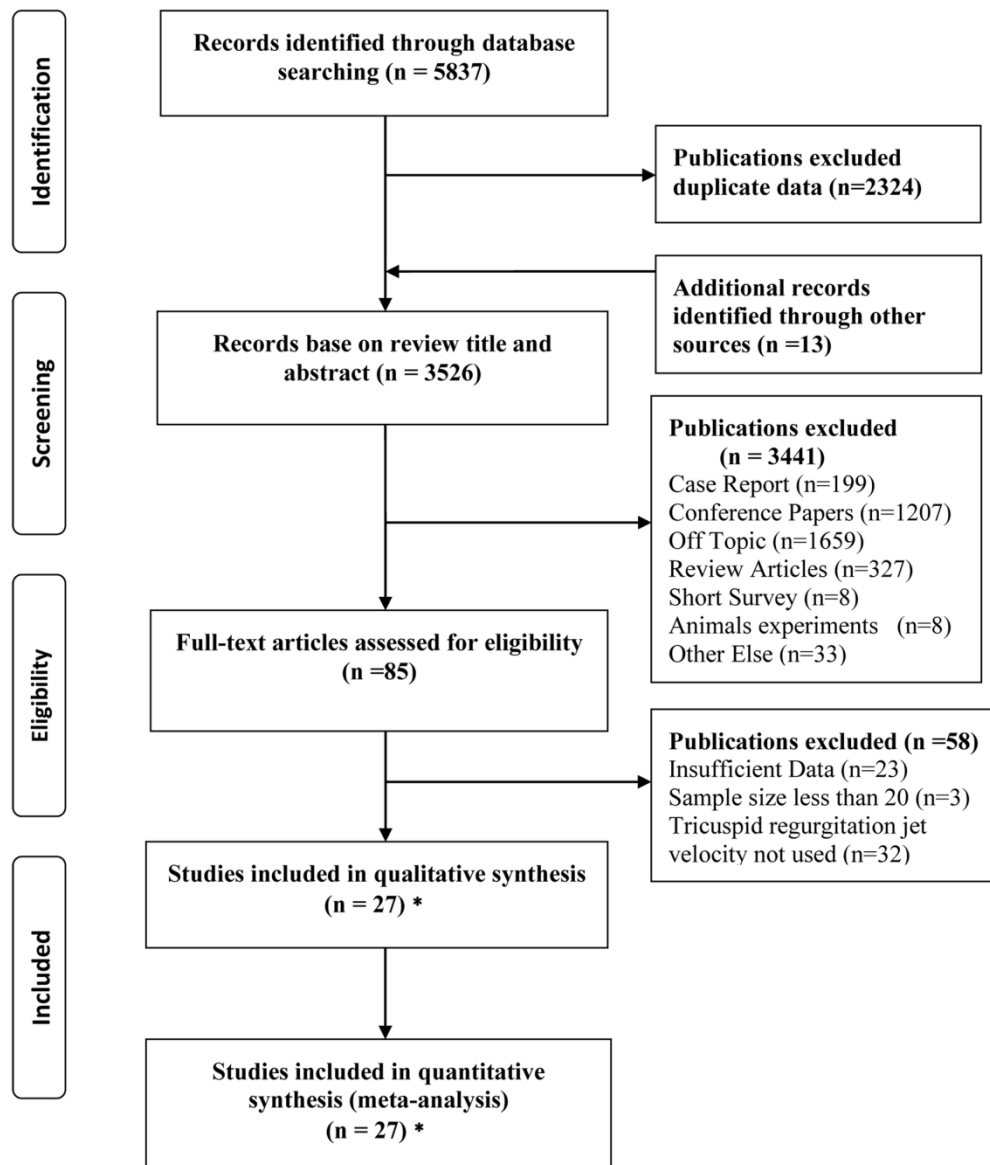


Figure 1 Flowchart for identification of the studies.

·Habash's study was divided into two independent parts because of the differences between the case group (Habash-1) and the control group (Habash-2). A total of 27 studies were included, but 28 sets of data were analyzed.

157x188mm (300 x 300 DPI)

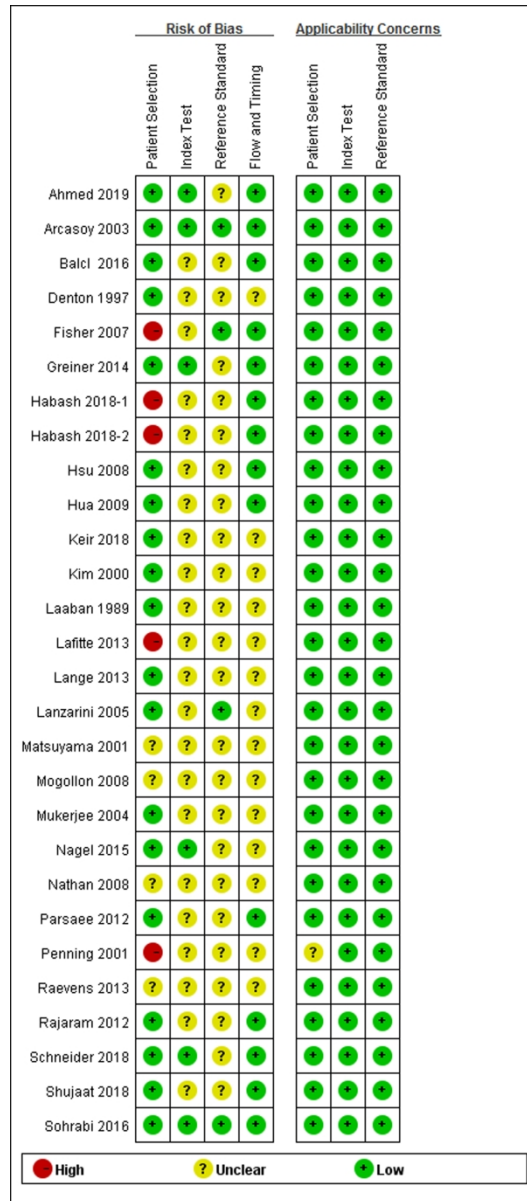


Figure 2 Risk of bias and applicability concerns summary: review authors' judgements regarding each domain for each included study (n=28).

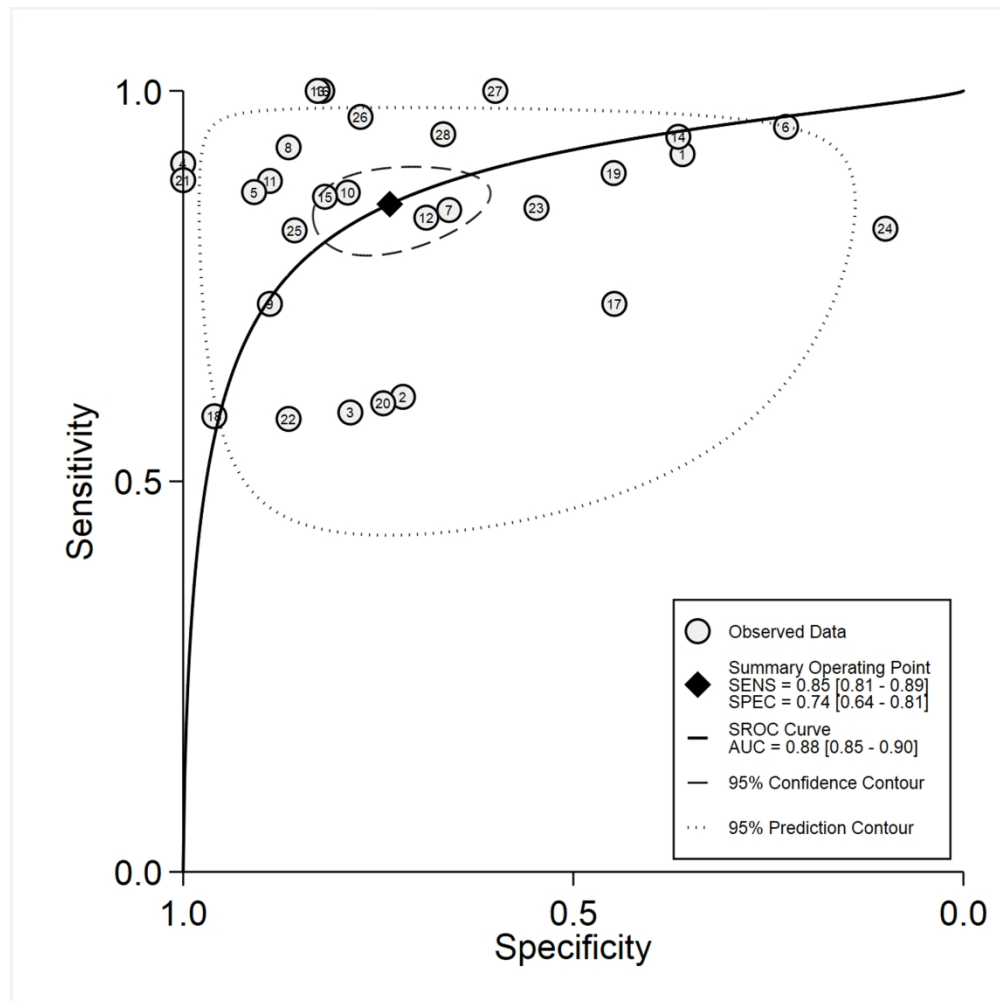


Figure 3 Summary receiver operating characteristic (SROC) graph with 95% confidence region and 95% prediction region for TTE in the diagnosis of pulmonary hypertension (n=28).

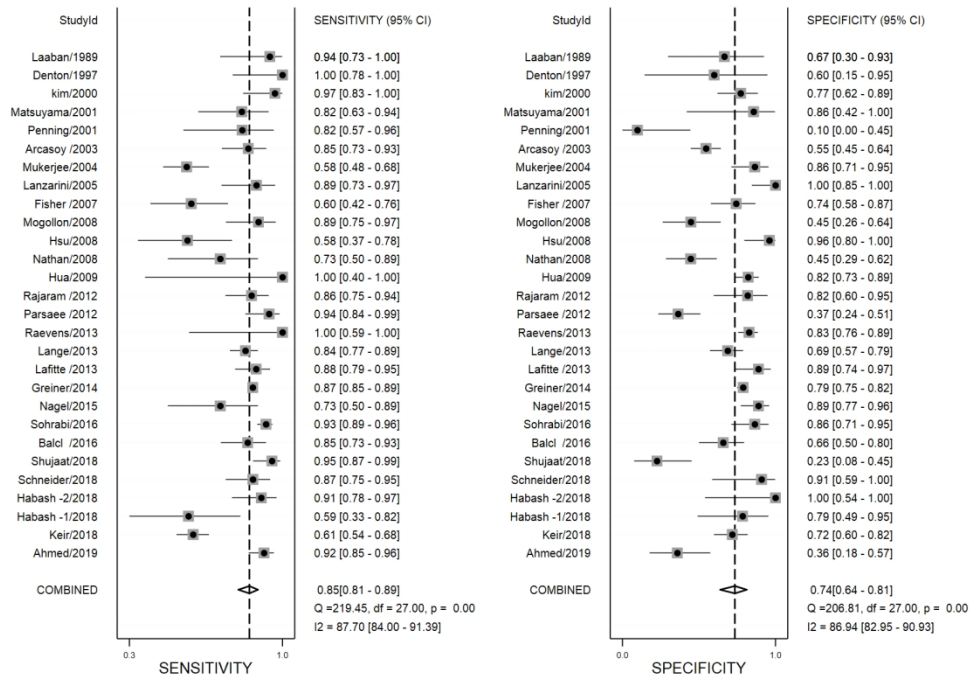


Figure 4 Forest plot of the sensitivity and specificity of each individual study, summary sensitivity and specificity and I2 statistic for heterogeneity (n=28).



# PRISMA-DTA Checklist

Section/topic	#	PRISMA-DTA Checklist Item	Reported on page #
<b>TITLE / ABSTRACT</b>			
Title	1	Identify the report as a systematic review (+/- meta-analysis) of diagnostic test accuracy (DTA) studies.	1
Abstract	2	Abstract: See PRISMA-DTA for abstracts.	3
<b>INTRODUCTION</b>			
Rationale	3	Describe the rationale for the review in the context of what is already known.	5
Clinical role of index test	D1	State the scientific and clinical background, including the intended use and clinical role of the index test, and if applicable, the rationale for minimally acceptable test accuracy (or minimum difference in accuracy for comparative design).	5
Objectives	4	Provide an explicit statement of question(s) being addressed in terms of participants, index test(s), and target condition(s).	5
<b>METHODS</b>			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	5
Eligibility criteria	6	Specify study characteristics (participants, setting, index test(s), reference standard(s), target condition(s), and study design) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	6
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	6
Search	8	Present full search strategies for all electronic databases and other sources searched, including any limits used, such that they could be repeated.	Supplementary Data
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	Figure 1
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	6
Definitions for data extraction	11	Provide definitions used in data extraction and classifications of target condition(s), index test(s), reference standard(s) and other characteristics (e.g. study design, clinical setting).	6-7
Risk of bias and applicability	12	Describe methods used for assessing risk of bias in individual studies and concerns regarding the applicability to the review question.	7
Diagnostic accuracy measures	13	State the principal diagnostic accuracy measure(s) reported (e.g. sensitivity, specificity) and state the unit of assessment (e.g. per-patient, per-lesion).	7
Synthesis of results	14	Describe methods of handling data, combining results of studies and describing variability between studies. This could include, but is not limited to: a) handling of multiple definitions of target condition. b) handling of multiple thresholds of test positivity. c) handling multiple index test readers. d) handling of indeterminate test results. e)	7

open-2019-033084 on 22 December 2019. Downloaded from <http://bmjopen.bmj.com/> on April 7, 2024 by guest. Protected by copyright.



# PRISMA-DTA Checklist

grouping and comparing tests, f) handling of different reference standards

Page 1 of 2

Section/topic	#	PRISMA-DTA Checklist Item	Reported on page #
Meta-analysis	D2	Report the statistical methods used for meta-analyses, if performed.	7
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	8
<b>RESULTS</b>			
Study selection	17	Provide numbers of studies screened, assessed for eligibility, included in the review (and included in meta-analysis, if applicable) with reasons for exclusions at each stage, ideally with a flow diagram.	Figure 1
Study characteristics	18	For each included study provide citations and present key characteristics including: a) participant characteristics (presentation, prior testing), b) clinical setting, c) study design, d) target condition definition, e) index test, f) reference standard, g) sample size, h) funding sources	23 Table 1
Risk of bias and applicability	19	Present evaluation of risk of bias and concerns regarding applicability for each study.	Figure 2
Results of individual studies	20	For each analysis in each study (e.g. unique combination of index test, reference standard, and positivity threshold) report 2x2 data (TP, FP, FN, TN) with estimates of diagnostic accuracy and confidence intervals, ideally with a forest or receiver operator characteristic (ROC) plot.	Figure 3
Synthesis of results	21	Describe test accuracy, including variability; if meta-analysis was done, include results and confidence intervals.	Figure 4
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression; analysis of index test: failure rates, proportion of inconclusive results, adverse events).	24-26 Table 2,3
<b>DISCUSSION</b>			
Summary of evidence	24	Summarize the main findings including the strength of evidence.	11
Limitations	25	Discuss limitations from included studies (e.g. risk of bias and concerns regarding applicability) and from the review process (e.g. incomplete retrieval of identified research).	13
Conclusions	26	Provide a general interpretation of the results in the context of other evidence. Discuss implications for future research and clinical practice (e.g. the intended use and clinical role of the index test).	12-13
<b>FUNDING</b>			
Funding	27	For the systematic review, describe the sources of funding and other support and the role of the funders.	2

Adapted From: McInnes MDF, Moher D, Thoms BD, McGrath TA, Bossuyt PM, The PRISMA-DTA Group (2018). Preferred Reporting Items for a Systematic Review and Meta-analysis of Diagnostic Test Accuracy Studies: The PRISMA-DTA Statement. JAMA. 2018 Jan 23;319(4):388-396. doi: 10.1001/jama.2017.19163.

For more information, visit: [www.prisma-statement.org](http://www.prisma-statement.org).