Comparative accuracy of non-invasive imaging versus right heart catheterization for the diagnosis of pulmonary hypertension: A systematic review and meta-analysis.

Waqas Ullah  
*Thomas Jefferson University*

Artem Minalyan  
*Abington Jefferson Health*

Sameer Saleem  
*University of Kentucky*

Follow this and additional works at: [https://jdc.jefferson.edu/abingtonfp](https://jdc.jefferson.edu/abingtonfp)

Recommended Citation

Ullah, Waqas; Minalyan, Artem; Saleem, Sameer; Nadeem, Nayab; Abdullah, Hafez M; Abdalla, Abdennoury; Chan, Vincent; Saeed, Rehan; Khan, Maria; Collins, Sara; Mukhtar, Maryam; Grover, Harshwant; Sattar, Yasar; Panchal, Ankur; Narayana Gowda, Smitha; Khwaja, Uneza; Lashari, Bilal; and Fischman, David L, "Comparative accuracy of non-invasive imaging versus right heart catheterization for the diagnosis of pulmonary hypertension: A systematic review and meta-analysis." (2020). *Abington Jefferson Health Papers*. Paper 31.  
[https://jdc.jefferson.edu/abingtonfp/31](https://jdc.jefferson.edu/abingtonfp/31)
Authors
Waqas Ullah, Artem Minalyan, Sameer Saleem, Nayab Nadeem, Hafez M Abdullah, Abdelmohaymin Abdalla, Vincent Chan, Rehan Saeed, Maria Khan, Sara Collins, Maryam Mukhtar, Harshwant Grover, Yasar Sattar, Ankur Panchal, Smitha Narayana Gowda, Uneza Khwaja, Bilal Lashari, and David L Fischman

This article is available at Jefferson Digital Commons: https://jdc.jefferson.edu/abingtonfp/31
Comparative accuracy of non-invasive imaging versus right heart catheterization for the diagnosis of pulmonary hypertension: A systematic review and meta-analysis


Abington Jefferson Health, PA, United States
University of Kentucky, Bowling Green, KY, United States
University of South Dakota Sanford School of Medicine, SD, United States
Ochsner Louisiana State University, Shreveport, United States
Rawalpindi Institute of Cardiology, Pakistan
Icahn School of Medicine, NY, United States
University of Pittsburgh Medical Center, Pittsburgh, PA, United States
Danbury Hospital, CT, United States
Thomas Jefferson University, PA, United States

Article info
Article history:
Received 23 April 2020
Received in revised form 30 May 2020
Accepted 4 June 2020

Abstract
Background: Right heart catheterization (RHC) is the gold-standard in the diagnosis of pulmonary hypertension (PH) but at the cost of procedure-related complications. We sought to determine the comparative accuracy of RHC versus non-invasive imaging techniques such as computed tomography (CT), magnetic resonance imaging (MRI), and transthoracic echocardiography (TTE).

Methods: Pulmonary hypertension was defined as a mean pulmonary artery pressure (mPAP) of >20 mmHg. Multiple databases were queried for relevant articles. Raw data were pooled using a bivariate model to calculate the measures of diagnostic accuracy and to estimate Hierarchical Summary Receiver Operating Characteristic (HSROC) on Stata 13.

Results: A total of 51 studies with a total patient population of 3947 were selected. The pooled sensitivity and specificity of MRI for diagnosing PH was 0.92 (95% confidence interval (CI) 0.88–0.96) and 0.86 (95% CI, 0.77–0.95), respectively. The net sensitivities for CT scan and TTE were 0.79 (95% CI 0.72–0.89) and 0.85 (95% CI 0.83–0.91), respectively. The overall specificity was 0.82 (0.76–0.92) for the CT scan and 0.71 (95% CI 0.61–0.84) for TTE. The diagnostic odds ratio (DOR) for MRI was 124 (95% CI 36–433) compared to 30 (95% CI 11–78) and 24 (95% CI 11–38) for CT scan and TTE, respectively. Chi-squared test showed moderate heterogeneity on the test for equality of sensitivities and specificities.

Conclusions: MRI has the highest sensitivity and specificity compared to CT and TTE. MRI can potentially serve as a surrogate technique to RHC for the diagnosis of PH.

1. Introduction

RHC is a gold standard for the diagnosis of PH by measuring mean pulmonary arterial systolic pressure (mPASP) [1]. Although RHC has the highest diagnostic accuracy, it comes at the cost of invasive procedure, catheter-related complications, and occasion-
Despite the recent advances, the comparative accuracy of less invasive methods (TTE, CT, and MRI) against RHC remains an area of uncertainty.

2. Methods:

2.1. Search strategy

A literature search for relevant articles was performed from inception to April 15, 2020, using PubMed, Ovid, Embase, clinicaltrials.gov, and Cochrane databases. There was no language or time restriction placed. The search strategies included various combinations of keywords and medical subject headings (MeSH) to generate four subsets of citations: magnetic resonance imaging (MRI), computed tomography (CT), transthoracic echocardiography (TTE) and pulmonary hypertension (PH). The terms from all subsets were combined in 1:1 combination using Boolean operators, and results from all possible combinations were downloaded into an EndNote library. Based on our research question, articles from the reference lists relevant to the clinical question were also screened by an independent author.

2.2. Selection criteria and study collection

All observational (retrospective/prospective) studies and randomized control trials (RCT) were evaluated. Articles comparing the diagnostic accuracy of MRI, CT, or TEE with RHC were included in the study. Inclusion criteria for study articles were as follows: (1) original articles published in English; (2) it included subjects who had mPAP, mPAD or PA:A ratio measurement for detection of PH; (3) surrogate markers for PH such as RV strain or tricuspid regurgitation was assessed; (4) patients were in stable condition; and (5) articles reported sufficient raw data to calculate sensitivity and specificity. (6) The most common criteria of TTE used to define PH included pulmonary artery systolic and mean pressure, tricuspid regurgitation, right ventricular systolic pressure, right ventricular diameter, and right ventricular free wall strain. CT and MRI also used interventricular septum curvature, ratio of interventricu-
lar septum curvature and left ventricular free wall, pulmonary artery pulsatility, pulmonary artery distensibility, ventricular mass index, pulmonary artery stiffness, and dilation of the main pulmonary artery as diagnostic criterias. Papers with insufficient data, review articles, case reports, editorials, and conference papers were excluded. Pulmonary hypertension was defined as a mean pulmonary artery pressure (mPAP) of >20 mmHg.

2.3. Data extractions

Five authors independently screened titles and abstracts of all the articles for relevance, and the sixth author selected articles which met the inclusion criteria. Full text articles that were potentially relevant to the study was screened by the seventh author to confirm eligibility. Baseline characteristics of the included population were reviewed, and data were collected for quality assessment to ensure study cohorts were statistically comparable. Data was finally extracted into the combined Excel sheet. Disagreements were resolved by mutual consensus and after a detailed group discussion.

2.4. Data and quality analysis

Raw data regarding the true and false positives and true and false negatives of each included study was obtained, and the combined measures of test accuracy [i.e., sensitivity, specificity, and diagnostic odds ratio (DOR)] for all non-invasive imaging modalities were calculated, keeping RHC a standard. A bivariate model was obtained from data fitting. The parameters from the bivariate model were transformed into Hierarchical Summary Receiver Operating Characteristic (HSROC) estimates. The area under the summary HSROC curve (AUC) was calculated to reveal the probability of accurately ranked diagnostic test values for a random pair of subjects (one with disease and one without disease). Chi-squared (x2) tests were conducted to evaluate the heterogeneity of sensitivities and specificities among all studies. Subgroup analysis based on the diagnostic threshold cutoff points of the index test was done to explore the potential heterogeneity factors using meta-regression analysis. Quality of the included studies was assessed for the potential bias and applicability concerns using the Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) standard questionnaire. The statistical analysis was performed using the diagnostic accuracy statistical model on the Stata and R mad package.

2.5. Quality of included studies

Based on the QUADAS-2 tool, the overall risk of bias and applicability concerns based on patient selection, index and reference standard was low. (Fig. 1) In the study by Cotton et al., the patient selection bias was of unclear concern. The risk of bias associated
with flow and timing was of unclear concern due to inadequate reporting across most studies. (Fig. 1) The optimal timing between the index test (TTE, CT or MRI) and the reference standard (RHC) was uncertain across most articles. The impact of inadequate reporting on the interpretation of their respective results is uncertain.

3. Results

3.1. Search results and study selection

The initial search on included databases revealed 15,741 articles. After the removal of irrelevant and duplicate items, only 269 articles were deemed relevant for full-text review. We further excluded 219 articles based on our selection criteria; only 51 studies were included for final analysis. The Prisma flow diagram is shown in Fig. 2.

A total of 51 studies comprising 3947 patients, were included in the study. The mean age of the included patients was 56 years, with an average male population of 46%. The detailed characteristics of the included studies are summarized in supplementary table 1.

3.2. Pooled sensitivity and specificity Results:

Compared to RHC, the pooled sensitivity and specificity of MRI for diagnosing PH was 0.92 (95% confidence interval (CI) 0.88–0.96) and 0.86 (95% CI, 0.77–0.95). (Figs. 3, 4) The sensitivity and
specificity of CT scan was 0.79 (95% CI, 0.72–0.89) and 0.82 (95% CI, 0.76–0.92) (Figs. 5, 6). TTE showed a sensitivity and specificity of 0.85 (95% CI, 0.83–0.91) and 0.71 (95% CI, 0.61–0.84) (Figs. 7, 8).

3.3. Subgroup Analysis:

A subgroup analysis based on the matched inclusion criteria and diagnostic parameters of PH was also performed. The subgroup analysis mirrored the overall findings of the pooled analysis with a few exceptions. In contrast to the pooled results of TTE, the sensitivity and specificity values were 0.92 and 0.45, when RV free wall strain was used as a diagnostic criteria. On the other hand, the sensitivity and specificity of different parameters used for defining PH varied significantly. For example, the sensitivity and specificity values were 0.92 and 0.86, respectively. Regardless of the measurement used, MRI is found to be valuable in terms of prognosis and follow-up after treatment [14–16].

3.4. Chi-squared testing results

The heterogeneity among the outcomes of the included studies was moderate through all comparisons. The Chi-squared ($\chi^2$) test for the sensitivity and specificity of MRI was 13 and 38, with a p<0.1. The $\chi^2$-test for the sensitivity and specificity of CT was 131 and 73 with a p<< 0.1 while it was 47 and 162 (p< 0.1), respectively.

3.5. Likelihood and diagnostic odds ratio results

The positive likelihood ratios (PLR) for MRI, CT and TTE were 8.7 ± 3.5 (95% CI 3.93–19.27), 6.0 ± 0.8 (95% CI 3.3–10.6) and 3.4 ± 0.8 (95% CI 2.23–5.31) respectively. The negative likelihood ratios (NLR) for MRI, CT and TTE was 0.069 ± 0.2 (95% CI 0.035–0.13), 0.19 ± 0.01 (0.11–0.33) and 0.15 ± 0.03 (95% CI 0.11–0.22) respectively. The diagnostic odds ratio (DOR) for MRI was highest at 124 (95% CI, 36–433) compared to 30 (95% CI, 11–78) and 24 (95% CI, 11–38) for CT and TTE, respectively.

4. Discussion

To the best of our knowledge, this is the first meta-analysis comparing multiple noninvasive modalities with RHC for the diagnosis of pulmonary hypertension. This is also the most updated meta-analysis with the largest number of studies and the largest sample size in each noninvasive diagnostic test category. Our results revealed that all three noninvasive imaging modalities might have a predictive value for the diagnosis of pulmonary hypertension (PH). On pooled analysis, MRI demonstrated the greatest diagnostic accuracy as signified by a large area under the curve on Receiver Operating Characteristic (ROC). The sensitivity and specificity of TTE and CT scan were also comparable to right heart catheterization (RHC), however, was relatively lower compared to MRI. Chi-squared testing for the evaluation of heterogeneity resulted in moderate heterogeneity, inferring that the results demonstrated likely reflect the true ability of other imaging modalities in the prediction of PH as compared to RHC. Though all three imaging modalities demonstrated a positive and significant relation to RHC, the likelihood ratios (LR) and diagnostic odds ratio (DOR) substantially favored MRI than CT and TTE in comparison to RHC.

PH results in dilatation of the right-sided heart chambers and subsequent tricuspid regurgitation [7]. These two phenomena play a critical role in the estimation of PH via noninvasive tests [8]. A mean pulmonary arterial pressure (PAP > 20 mmHg) is considered a diagnostic of PH when measured at rest in the supine position. However, a variety of parameters and cutoffs were found to be suggestive of PH, explaining the heterogeneity among the included studies [9]. The suggestive findings for PH on MRI included ventricular mass index (VMI) and right atrial (RA) pulsatility. VMI is calculated by dividing the right ventricular mass by that of the left ventricle mass (VMI = RVM/LVM) [10]. The reported sensitivity and specificity of VMI is 84% and 82%, respectively [11]. A study by Seba et al. found VMI assessment of PH by MRI has a higher accuracy as compared to pressure assessment of TTE [12]. A meta-analysis by Wang et al. also used VMI as a measure and found to have high pooled sensitivity (84%) and specificity (82%) [11]. Our meta-analysis included 6 studies to assess the diagnostic accuracy of MRI using both RA pulsatility (below 40%) and VMI (>0.45). Other suggestive measures included pulmonary artery (PA) to ascending aorta (AO) ratio (PA/AO). This ratio is an indirect measure of interventricular septal flattening during systole and correlates with PASP [13]. Our meta-analysis reports higher results than wang et al. and yielded a pooled sensitivity and specificity of MRI as 0.92 and 0.86, respectively. Regardless of the measurement used, MRI is found to be valuable in terms of prognosis and follow up after treatment [14–16].

By contrast, TTE is the most convenient and initial noninvasive screening tool. The including measurements of TTE for PH diagnosis and assessment include pulmonary artery regurgitant jet method, measurement of the tricuspid annular plane systolic excursion (TAPSE), two-dimensional strain tissue Doppler echocardiography, the speckle tracking method, acceleration time across the pulmonic valve and the tricuspid regurgitant jet velocity (TRV) [17–21]. TRV is used to calculate the ePASP using RAP, which can be clinically estimated from jugular venous pressure (JVP) using a fixed value of 10 mm Hg, and the diameter and collapse of the inferior vena cava (IVC) during spontaneous respiration [18,22–24]. ESC and ERS recommend using peak TRV alongside other echocardiographic findings suggestive of PH such as right ventricle/left ventricle basal diameter ratio > 1.0, right ventricular outflow Doppler acceleration time < 105 ms and/or mid-systolic notching & right atrial end-systole area > 18 cm² [25]. Our review...
showed that the study by Condliffe et al. used TRV to calculate the tricuspid gradient (TG), while Collin et al. and Rajaram et. al. used TR gradient (cut off of > 27 mmHg and > 40) for diagnosing PH [26–28]. Surrender et al., used sPAP as a measure of assessment of PH with a cutoff range ≥ 30 mmHg [4]. The mPASP was calculated based on the maximum velocity of the tricuspid regurgitant jet method (TRJ + RAP) (estimated using JVP/IVC collapse). Zhang et al. included 6 studies, the measure used for the diagnosis of PH was RVSP and PASP, demonstrating intermediate sensitivity results compared to RHC [29]. A prior systematic review published in 2010, assessed 29 studies, comparing TTE for screening and diagnosis of PH. PASP measurement had a correlation of about 0.7 with RHC, with pooled sensitivity and specificity of 0.83 and 0.72, respectively [4]. Another meta-analysis examining two variables on CT, PA diameter and PA:A ratio was published by Shen et al. in 2014. The reported sensitivities and specificities for the former was found to be 0.79 and 0.83, while for the latter, it was 0.74 and 0.81, respectively [5]. Of note, PA diameter was also found to correlate with prognosis in patients with chronic respiratory illness, e.g., bronchiectasis [6].

**Fig. 7.** Forest plot showing the individual and pooled sensitivity and specificity of TTE.
On the contrary, our analysis included the largest population from a total of 27 studies comparing TTE and RHC. Most included studies used TRJ and RAP to calculate mPASP. Of the included studies, Opotowsky et al. used multiple parameters including RV dilatation, prediction rule (ranges from $C_{0}^{2}$ to $+2$) and mid systolic notch or acceleration time $<80$ msec [30]. Single studies used right atrial strain and RV free wall strain. Some studies were performed in specific patient populations, such as patients with diffuse pulmonary fibrosis, patients with systemic sclerosis, or liver transplant candidates, which could result in a lack of generalizability of results; however, this was a small minority of studies included in our meta-analysis. Overall, our results showed a pooled sensitivity of 85% and specificity of 71% for TTE compared to RHC. Additionally, it is important to note that both TTE and CMR can be useful to gauge response to therapy on regular follow-ups by demonstrating reversal of right ventricular dilatation and change in pulmonary vascular resistance.

The third non-invasive imaging modality used was CT scan, which, compared to MRI, is feasible to obtain and can also give information about the etiology of PH. The measure of diagnosis on CT is main pulmonary artery diameter (MPAD) with a cutoff $>29$ mm with a reported diagnostic sensitivity of 87%, specificity of 89%, and PPV of 97% [5,31]. The specificity of PH can reach up to 100% with MPAD $>29$ in combination with segmental artery-to-bronchus diameter ratio of 1:1 in 3–4 lobes [32]. Other measures of diagnosis of PH on CT can be the main pulmonary artery to ascending aorta diameter ratio of $\geq 1$ [5,33,34]. CT scan also helps to differentiate group-1 PH (secondary to lung disease and/or hypoxia) from group-3 PH (Idiopathic, drug/toxin-induced, connective tissue disease, and autoimmune cases) [15]. Shen et al. reported a sensitivity of 79% and specificity of 83% based on the pulmonary diameter and a sensitivity and specificity of 74% and 81% for TTE, respectively [5]. By contrast, our results included 27 studies comparing CT scan with RHC for diagnostic accuracy of PH. Our results were in concordance with Shen et al. study, with a sensitivity of 79% and specificity of 82% for both pulmonary artery diameter and PA:A ratio. We think this is a valuable conclusion as CT scans are readily available and done more frequently than ever, mostly for reasons other than PH [35]. In addition, CT scans can help in the diagnosis of comitantant lung parenchymal diseases associated with higher pulmonary vascular resistance (like bronchiectasis and interstitial lung disease). Furthermore, other non-numerical surrogate indicators of cor pulmonale can aid in the diagnosis of PH if seen by CT images (e.g.: leftward septal deviation and thickening of right ventricular free wall). [14]

Overall, we believe that both TTE and CT have been falling out of favor in serially following patients diagnosed with PH due to operator dependency in the former and risk of radiation exposure in the latter. MRI, however, has been shown to be exceptionally reliable with good inter-study reproducibility in both healthy subjects and patients with heart failure.[36]

## 5. Limitations

Though the data supports MRI being the most sensitive and specific, further studies would need to be conducted to evaluate the cost relationship between the three imaging modalities. Issues such as the availability of the imaging modalities within different health systems around the world, as well as having the appropriate staff to read the imaging modality are required. A widespread variability in the measures and cutoffs of different parameters across the included studies limited our ability to measure its predictive effect. It is important to note that due to variable diagnostic criterias used by the included studies, our results might fail to identify a selected patient population who might have a normal or near-normal PAP. Some articles used the RV remodeling as a surrogate marker for PH, which could be misleading. Additionally, we could not take into account the pre-test probability of PH and could not include studies not reporting raw data of diagnostic accuracy.

### Table 1

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Imaging</th>
<th>No. of studies</th>
<th>Mean sensitivity</th>
<th>Mean specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pulmonary artery systolic pressure (PASP)</td>
<td>TTE</td>
<td>12</td>
<td>0.78</td>
<td>0.68</td>
</tr>
<tr>
<td>Right ventricular systolic pressure (RVSP)</td>
<td>TTE</td>
<td>6</td>
<td>0.88</td>
<td>0.62</td>
</tr>
<tr>
<td>Tricuspid regurgitant jet (TRJ)</td>
<td>TTE</td>
<td>6</td>
<td>0.81</td>
<td>0.84</td>
</tr>
<tr>
<td>RV free wall strain</td>
<td>TTE</td>
<td>1</td>
<td>0.92</td>
<td>0.45</td>
</tr>
<tr>
<td>Right ventricle diameter (RVD)</td>
<td>TTE</td>
<td>2</td>
<td>0.79</td>
<td>0.77</td>
</tr>
<tr>
<td>Pulmonary artery diameter (PAd)</td>
<td>CT</td>
<td>15</td>
<td>0.71</td>
<td>0.78</td>
</tr>
<tr>
<td>RV diameter</td>
<td>CT</td>
<td>2</td>
<td>0.84</td>
<td>0.90</td>
</tr>
<tr>
<td>RV wall thickness</td>
<td>CT</td>
<td>3</td>
<td>0.74</td>
<td>0.81</td>
</tr>
<tr>
<td>PA diameter/AA diameter</td>
<td>CT</td>
<td>5</td>
<td>0.65</td>
<td>0.79</td>
</tr>
</tbody>
</table>

![Fig. 8. ROC demonstrating a higher sensitivity and specificity for TTE compared to RHC.](image)
6. Conclusions

MRI has the highest diagnostic accuracy among the noninvasive diagnostic modalities for the detection of pulmonary hypertension. Echocardiography is the most feasible test but has lower sensitivity and specificity.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We thank Dr. Margot Boigon, Dr. David Smith and Dr. Richard Eisenstaedt for providing research opportunities and resources in the institute.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ijcha.2020.100568.

References: