Predicting Group II pulmonary hypertension: diagnostic accuracy of the H2FPEF and OPTICS scores in Scotland

Harrison Stubbs,1,2 Alexander MacLellan,1,3 Michael McGettrick1,3 Bhautesh Jani2, Melanie Brewis,1,3 Colin Church1,3 Martin Johnson1,3

ABSTRACT

Objective Group II pulmonary hypertension (PH) can be challenging to distinguish from Group I PH without proceeding to right heart catheterisation (RHC). The diagnostic accuracy of the H2FPEF and OPTICS scores was investigated in Scotland.

Methods Patients were included in the study if they were referred to the Scottish Pulmonary Vascular Unit between 2016 and 2020 and subsequently diagnosed with Group II PH or Group I PH which was either idiopathic, heritable or pulmonary veno-occlusive disease. The established cut offs for the H2FPEF and for the OPTICS scores were applied retrospectively to predict the presence of Group II PH. The diagnosis from the scores were compared with the MDT consensus diagnosis following RHC.

Results 107 patients with Group I PH and 86 patients with Group II PH were included. Retrospective application of the OPTICS score demonstrated that pretest scoring would detect 28% of cases with Group II PH yet at the cost of misdiagnosing 4% of patients with Group I as Group II PH (specificity 0.96). The H2FPEF score had a far greater sensitivity (0.70) yet reduced specificity (0.91), leading to misdiagnosis of 9% of Group I PH cases.

Conclusion While the specificity of these scores was high, the lack of perfect specificity limits their utility as it results in missed patients with Group I PH. As a consequence, they cannot replace RHC as the means of diagnosing the aetiology of PH in their current form. The scores may still be used to support clinical judgement or to indicate the advisability for further provocative testing at RHC.

BACKGROUND

Group II pulmonary hypertension (PH) is vital to distinguish from Group I PH due to significant differences in management and prognosis.1 Group II PH, attributable to left heart disease, is defined by an elevated pulmonary artery wedge pressure ≥15mmHg during right heart catheterisation (RHC), in conjunction with pulmonary hypertension (mean pulmonary artery pressure (mPAP)≥25mmHg).1,2 Prior to invasive testing, Group II PH can be diagnosed when there is clear evidence of left heart disease on echocardiogram and other testing modalities.1 However, in the absence of clear features, or in cases where the extent of left heart disease appears insufficient to explain the elevated pulmonary artery pressure, RHC may be required in order to clarify the aetiology.1,4 While confirming Group II PH on RHC may refine the diagnosis, it rarely changes the subsequent left heart failure management.

Patients with heart failure with preserved ejection fraction (HFpEF) are a heterogeneous group, who are more elderly and comorbid, and in whom clear echocardiographic features of left heart disease can be lacking.5,6 Scoring systems have been developed to aid clinicians in determining the
prognostic information among patients with HFpEF. It has been reported to provide a specificity between 69% and 91% of HFpEF if a total score of equal to or greater than 6 is calculated and provides prognostic information among patients with HFpEF. Although not specifically designed for this purpose, it has been suggested as a diagnostic tool to predict Group II PH. In a prospective cohort in the Netherlands, the H2FPEF score was found to predict Group II PH with a sensitivity of 48%, yet also predicted HFpEF in a significant number of patients who were found to have Group I PH at RHC (positive predictive value 88%). The H2FPEF score weights atrial fibrillation (AF) highly and while AF is a predictor of HFpEF, it coexists in patients with Group I PH due to right atrial dilation or as a non-contributing comorbidity, potentially leading to overestimation of Group II PH in this cohort. The OPTICS score was created for evaluating the pretest probability of Group II PH in new referrals to tertiary PH centres. The sensitivity of diagnosing Group II PH has been reported as 100% if a score of ≥104 is calculated. This score weights previous valvular surgery highly, while AF is less highly weighted.

The Scottish population has high rates of AF, obesity and diabetes which are likely to alter the accuracy of both scores. Consequently, we wished to investigate the diagnostic accuracy of the H2FPEF and OPTICS scores for predicting Group II PH among a cohort of patients referred to the tertiary pulmonary vascular unit for Scotland, the Scottish Pulmonary Vascular Unit (SPVU).

METHODS

The Standards for Reporting of Diagnostic Accuracy Studies (STARD) guidelines were used to structure the study. A retrospective analysis was undertaken of all patients who were referred to SPVU between 2016 and 2020. Patients investigated at SPVU follow a standard diagnostic protocol of thoracic and cardiac imaging, exercise capacity testing and RHC. A final diagnosis is based on multidisciplinary consensus between pulmonary vascular physicians and cardiopulmonary imaging specialists. The study was approved by the East of Scotland research ethics service (Ref 21/ES/0078).

Patients were included if they were accepted for invasive investigation, underwent RHC and were subsequently diagnosed with Group II PH, or with Group I PH that included the subtypes idiopathic pulmonary arterial hypertension (IPAH), heritable pulmonary arterial hypertension (HPAH) and pulmonary veno-occlusive disease (PVOD), as defined by established criteria. Patients with other suspected aetiologies for precapillary PH (such as evidence of chronic thromboembolic disease (CTED) or underlying connective tissue disease) were excluded in order to reflect true-to-life circumstances where the suspicion of a Group I cause is great enough in these cases to proceed with investigation, even when left heart disease risk factors are present.

Individual patient records were screened for the constituent components of the H2FPEF and OPTICS score. The established cut-offs of ≥6 for the H2FPEF and ≥104 for the OPTICS scores were used to denote whether the score indicated the presence of Group II PH. This was compared with the multidisciplinary team (MDT) consensus diagnosis of Group II PH. Group II PH was used to define cases of both isolated postcapillary and combined precapillary and postcapillary PH as either subgroup are not candidates for pulmonary vasodilator therapy outside clinical trials. False-positive cases, where patients with Group I PH were predicted to have Group II PH by either score, were analysed in further detail. Calibration was performed by systematically making incremental adjustments to the cut offs and weighting of components for both scores, in an attempt to improve the sensitivity while not adversely impacting specificity.

Statistical analysis

Patient demographics were presented as means±SD unless otherwise stated. Unpaired t and χ² tests were used to calculate differences between patient groups for continuous and categorical variables respectively. A p value of <0.05 was considered as statistically significant. All data were analysed with GraphPad Prism (V.9.3.0 for Windows, GraphPad Software, San Diego, California, USA). For both the H2FPEF and the OPTICS scores, sensitivity, specificity, positive predictive value, negative predictive value, receiver operating curves and area under the curve (AUC) were calculated.

RESULTS

Seven hundred and eighty one diagnostic RHC were performed in the study period, with 107 patients with IPAH, HPAH or PVOD and 86 patients with Group II PH included in the analysis. Patient demographics are presented in Table 1. When compared with patients with Group I PH, patients with Group II PH were more elderly, had a higher body mass index (BMI), were more likely to have high cardiovascular risk comorbidities, were more likely to present with a raised E/e’ ratio and LA dilation and had a lower mean pulmonary artery pressure and pulmonary vascular resistance at RHC.

Table 2 demonstrates the sensitivity and specificity for both scores from this cohort. Retrospective application of the OPTICS score demonstrates that pretest scoring would allow detection of between one in four and one in three cases of Group II PH (sensitivity 0.28) yet at the cost of misdiagnosing 4% of referred Group I cases (specificity 0.96). The H2FPEF score had a far greater sensitivity (0.70) yet at the cost of reduced specificity (0.91), implying 9% of Group I cases would be misdiagnosed. ROC analysis demonstrated a similar area under the curve for both scores (see figure 1).
Table 3 demonstrates the characteristics of false-positive cases for both scoring systems. Of 107 cases of Group I PH, 4 (3.7%) cases were mislabelled as Group II PH by the OPTICS score, all of whom were diagnosed with IPAH. Individual analysis of these patients reveals a high prevalence of cardiovascular risk factors, a mean OPTICS score of 122 and a mean age of 78 years. Of these four patients, one patient did not symptomatically improve with treatment and three patients died within 8 months of diagnosis. Ten cases (9.3%) were false positive by the H2FPEF score; in this cohort, the mean age was 74 years and the mean H2FPEF score was 7.1. Of these, four patients had symptomatic and objective improvement with treatment, five patients had no benefit and one patient was lost to follow-up; two patients died within a year of diagnosis. There was an overlap of two patients who were mislabelled by both scores.

Table 2  Validation metrics for the H2FPEF and OPTICS scores at predicting Group II pulmonary hypertension

<table>
<thead>
<tr>
<th></th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>PPV</th>
<th>NPV</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>OPTICS</td>
<td>0.28</td>
<td>0.96</td>
<td>0.86</td>
<td>0.62</td>
<td>0.66</td>
</tr>
<tr>
<td>H2FPEF</td>
<td>0.70</td>
<td>0.91</td>
<td>0.86</td>
<td>0.79</td>
<td>0.81</td>
</tr>
</tbody>
</table>

NPV, negative predictive value; PPV, positive predictive value.
scores. The OPTICS score was unable to be calibrated without significantly increasing the chance of false positives. Applying the scores consecutively, in either order, did not improve sensitivity or specificity. However, by changing two variables of the H2FPEF score (BMI cut-off >35 as opposed to 30 and reducing the weighting of AF from 3 to 2), the specificity improved to 0.98 and the sensitivity fell to 0.41. This method identified two false-positive cases, neither of whom derived benefit from treatment.

**DISCUSSION**

Retrospective application of the OPTICS score demonstrated that pretest scoring would detect 28% of cases with Group II PH yet at the cost of misdiagnosing 4% of patients with Group I PH as Group II PH (specificity 0.96). The H2FPEF score had a far greater sensitivity (0.70) yet reduced specificity (0.91), leading to misdiagnosis of 9% of Group I PH cases. This study is a further demonstration that the OPTICS score is able to correctly predict Group II PH in only a minority of patients, although maintaining a low false-positive rate. The H2FPEF score performed similarly, with a greater sensitivity yet crucially a lower specificity and hence a higher risk of false positives. While the specificity of these scores was high, the lack of 100% specificity might result in PH centres failing to proceed with invasive investigations in a subgroup of patients with Group I PH if the score is applied to screen new referrals.

The results from this study are similar to those from Jansen et al, in which the OPTICS score was found to have a sensitivity of 0.22 and specificity of 1.00, while the H2FPEF score’s sensitivity was 0.48 and specificity 0.92. Our approach differed from that of Jansen, in that patients with clear other attributable causes of PH, such as evidence of CTED and connective-tissue disease, were not included in this study, as it is likely that regardless of left heart disease risk factors these patients are likely to be accepted for further investigation. This may explain that while the validation results from this cohort are similar to the results of Jansen et al, the OPTICS score had a lower specificity.
The accuracy of any test is dependent on the studied population, and the high prevalence of cardiovascular comorbidities in Scotland, as compared with the Netherlands, is likely to have had an impact in decreasing the sensitivity and specificity of the OPTICS score, as both patients with Group I and Group II PH had a high prevalence of left-heart disease risk factors. As a consequence, the ultimate use for such scoring systems may be limited to prompting clinicians to consider provocative testing during RHC in order to unmask left heart disease, rather than a diagnostic tool in their own right.19 20

While ethnicity was not collected as part of this study, it can be presumed that this was generally a homogeneous white European population. The haemodynamic definition of PH may be changing, with a revised mPAP cut-off >20 mmHg being considered. This will affect the performance of these scores and a reassessment would be required. This study was additionally limited by lack of a prospective cohort to validate our proposed revision of the H2FPEF score, which may be a future consideration for research.

In conclusion, this study quantifies the probability of misdiagnosing patients with Group II PH, prior to invasive investigations, based on the OPTICS and H2FPEF scores. Overall, the utility of both these scores in their current versions is diminished by the small but measurable risk of missing true Group I PH, which would deprive patients of effective treatment. Therefore, these scores cannot replace RHC for determining the aetiology of PH. The scores may, however, play a role in helping to support clinical decisions or to suggest the need for provocative testing at RHC. Further research should aim to validate these scores further with a focus on calibration to perfect the specificity without affecting sensitivity.

**Contributors** HS: Data collection, statistical analysis, data interpretation, writing of the manuscript with support from AM, MM, JB, MB, CC and MJ. AM: Data collection. MM: Data collection. JB: Data interpretation, review and editing of the manuscript draft. MB: Data interpretation, review and editing of the manuscript draft. CC: Data interpretation, review and editing of the manuscript draft. MJ: Study conceptualisation, supervision, drafting of the manuscript, approval of the final draft for submission and is responsible for overall content as the guarantor.

**Funding** The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

**Competing interests** The authors whose names are listed immediately below certify that they have no affiliations with or involvement in any organisation or entity with any financial interest or non-financial interest in the subject matter or materials discussed in this manuscript: Harrison Stubbs, Alexander MacLellan, Michael McGettrick, Bhautesh Jani, Melanie Brewis, Colin Church and Martin Johnson. Outwith this study, Martin Johnson has received financial support for attending conferences and for partaking on an advisory board. Harrison Stubbs has received financial support for attending conferences. Neither of these relate to the matter or materials discussed within this manuscript. These are discussed in detail in the attached ICMJE form.

**Patient consent for publication** Not applicable.

**Ethics approval** The study was approved by the East of Scotland research ethics service (Ref 21/ES/0078). This was a retrospective study using existing clinical data that was anonymised by the research team at the time of data entry. The research team was compromised wholly of the clinical care team.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Data availability statement** No data are available.

**Open access** This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.

**ORCID iDs**

Harrison Stubbs http://orcid.org/0000-0003-1786-342X
Alexander MacLellan http://orcid.org/0000-0003-3436-9925
Michael McGettrick http://orcid.org/0000-0003-0070-6452
Bhautesh Jani http://orcid.org/0000-0001-7348-514X
Melanie Brewis http://orcid.org/0000-0002-4468-6978
Colin Church http://orcid.org/0000-0002-4446-0100

---

*Figure 2* Comparison of pyramid charts for the H2FPEF and OPTICS scores. Blue lines represent the number of patients with precapillary PH and red lines the number of patients with postcapillary PH, as stratified according to score. The dotted line represents the cut-off of ≥104 for the OPTICS score and ≥6 for the H2FPEF score.
REFERENCES