Prognostic Value and Reproducibility of Simplified Noninvasive Pulmonary Artery Pressure-Flow Relationships in Exercise Echocardiography

Peer-reviewed author version

BEKHUIS, Youri; MILANI, Mauricio; FALTER, Maarten; L'Hoyes, W; HOEDEMAKERS, Sarah; Soens, L; MOURA FERREIRA, Sara; DHONT, Sebastiaan; PAUWELS, Rik; DE PAEPE, Jarne; DELPIRE, Boris; Jacobs , Annemie; De Schutter, Stephanie; Spencer, L; Rowe, S; Lagerche, A; STASSEN, Jan; Gevaert, AB; Debonnaire, P; van de Bruaene, A; JASAITYTE, Ruta; BERTRAND, Philippe; HERBOTS, Lieven; VERBRUGGE, Frederik; CLAESSEN, Guido & VERWERFT, Jan (2025) Prognostic Value and Reproducibility of Simplified Noninvasive Pulmonary Artery Pressure-Flow Relationships in Exercise Echocardiography. In: Journal of Cardiac Failure, 31 (11) , p. 1707 -1715.

DOI: 10.1016/j.cardfail.2025.08.013
Handle: http://hdl.handle.net/1942/49323

Prognostic Value and Reproducibility of Simplified Non-Invasive Pulmonary Artery Pressure-Flow Relationships in Exercise Echocardiography

Data availability

The data underlying this article are available in the article. The anonymized data will be shared on reasonable request to the corresponding author

WORDS: 2612/3500 (excluding references, figure legends and tables)

ABBREVIATIONS

exPHT	Exercise Pulmonary Hypertension
mPAP	Mean Pulmonary Artery Pressure
CO	Cardiac Output
CPET	Cardiopulmonary Exercise Test
CPETecho	Cardiopulmonary Exercise Test combined with Exercise Echocardiography
ROC	Receiver Operating Characteristic
AUC	Area Under the Curve
ICC	Intraclass Correlation Coefficient

ABSTRACT

Background

The non-invasive mean pulmonary artery pressure (mPAP) to cardiac output (CO) relationship has been associated with cardiovascular outcomes across various populations. This study compared different methodologies for assessing this non-invasive relationship during exercise, with a dual focus on reproducibility and prognostic value. Specifically, we sought to determine whether simplified one-point mPAP/CO measurements offer equal reproducibility without compromising prognostic performance in patients with unexplained exertional dyspnea undergoing exercise echocardiography.

Methods and Results

This was a secondary analysis of a prospective, multicenter cohort study including 1,619 patients referred for unexplained dyspnea to six dedicated clinics in Belgium between January 2016 and March 2023. All patients underwent symptom-limited cardiopulmonary exercise testing with concurrent exercise echocardiography. The mPAP-CO relationship was assessed using four approaches: (1) three-point linear regression, (2) two-point slope (rest to peak), (3) one-point mPAP/CO ratio at low-level exercise, and (4) one-point mPAP/CO ratio at peak exercise. The primary outcome was a composite of all-cause mortality or heart failure hospitalization. Median follow-up was 28.2 months (IQR: 15.9-45.1). Prognostic performance was evaluated by comparing discrimination across methods using receiver operating characteristic (ROC) curves. Reproducibility was assessed by intra-observer and inter-observer agreement.

The one-point at peak exercise showed the highest discriminative performance (AUC 0.715), with modest but significant improvement over the two-point (AUC 0.685, $p = 0.019$) and three-point methods (AUC 0.686, $p = 0.028$). The one-point at low-level exercise showed similar accuracy (AUC 0.713). Reproducibility was highest for the one-point ratios, with narrower intra- and inter-observer limits of agreement ($p<0.001$) compared to slope-based approaches.

Conclusions

Simple, one-point mPAP/CO ratios measured during exercise offer prognostic performance comparable to slope-based methods, with higher reproducibility. This supports adoption as a practical and robust approach for evaluating exercise pulmonary hypertension.

Keywords:

Unexplained dyspnea

Exercise echocardiography

Exercise pulmonary hypertension

Risk Assessment

Clinical Relevance to Patient Care:

- Simplified mPAP/CO ratios offer a practical, non-invasive tool to assess exercise pulmonary hypertension
- Low-level exercise ratios provide a reliable alternative when peak measurements are not feasible
- High reproducibility supports consistent use across clinicians and institutions

INTRODUCTION

Exertional dyspnea is a prevalent yet diagnostically challenging symptom encountered across multiple medical specialties.[1–3] Despite the absence of abnormal resting findings, the symptom is frequently associated with exercise-induced cardiac or pulmonary pathology and has a well-documented prognostic significance.[4,5] Since resting evaluations are commonly inconclusive, exercise-based assessments have become central to diagnostic work-up in this setting and dedicated dyspnea clinics have emerged worldwide. [6,7]

Within this diagnostic landscape, exercise pulmonary hypertension (exPHT) represents a clinically important entity, affecting approximately 41% of patients with dyspnea, occurring as either a primary cause or secondary to left heart disease, pulmonary disease, or systemic conditions.[4] Historically, exPHT was defined as a mean pulmonary artery pressure (mPAP) > 30 mmHg during exercise. However, this threshold was later abandoned, as healthy individuals can exceed it at high cardiac outputs (CO).[8] Accordingly, interpretation of pressure-flow relationships, rather than absolute pressure alone, has become essential to distinguish physiological from pathological responses to exercise.[9,10] Current pulmonary hypertension (PHT) guidelines define exPHT invasively as an abnormal mPAP-to-CO slope >3.0 mmHg/L/min during exercise.[11]

ExPHT independently predicts adverse outcomes in patients with exertional dyspnea.[4,12,13] Although invasive hemodynamic measurements remain the gold standard, previous studies have demonstrated the prognostic value of non-invasively assessed exPHT using exercise echocardiography across various cardiac conditions on top of resting evaluations. [12,14,15]

Beyond prognostic risk stratification, pulmonary pressure–flow responses during exercise may also reflect underlying pathophysiological mechanisms. Recent work has linked abnormal mPAP/CO slopes to impaired ventricular-arterial interactions and systemic factors such as iron deficiency, irrespective of LVEF phenotype.[16,17]

Despite its clinical importance, there is currently no standardized non-invasive method to quantify the mPAP-CO relationship during exercise. The reproducibility and prognostic implications of different non-invasive approaches remain largely unexplored. To date, no large study has directly compared multiple non-invasive methods for assessing mPAP/CO in a well-characterized clinical population.

In this study, we aimed to compare four commonly used non-invasive methods for calculating the mPAP-CO relationship in patients with unexplained exertional dyspnea undergoing exercise echocardiography. We hypothesized that a simplified, one-point mPAP/CO ratio during exercise would improve reproducibility without compromising prognostic value.

METHODS

Study design and population

This was a secondary analysis of a multicenter cohort study conducted at six multidisciplinary dyspnea clinics in Belgium. It included patients with unexplained dyspnea referred between January 2016 and March 2023 for cardiopulmonary exercise testing with concurrent exercise echocardiography (CPETecho). Clinical procedures were standardized across all six centers, as previously described.[7,18] Data was systematically collected during routine care. Patients with a clear resting diagnosis were excluded (**Figure 1**). Only patients with complete mPAP and CO measurements at rest, low-level, and peak exercise were included in the analysis.

The primary composite outcome was the time to first occurrence of all-cause mortality or heart failure hospitalization. Outcome was assessed via retrospective chart review of all available medical records. Heart failure hospitalization was defined as an unplanned admission requiring intravenous diuretics, inotropes, or vasodilators. Patients who were lost to follow-up within the first year and did not experience an event were excluded to ensure complete outcome classification.

The study protocol received approval from local ethics committees. Due to the retrospective nature of the analysis, the requirement for written informed consent was waived. This study complies with the principles of the Declaration of Helsinki and Good Clinical Practice guidelines.

Assessment of exercise pulmonary pressures and cardiac output

All patients underwent a maximal symptom-limited CPETecho on a semi-supine bicycle positioned at a 45° angle, using individualized ramp protocols (5-20 W/min) targeting a total duration of 10-15 minutes. Pulmonary pressures and CO were assessed at three standardized stages: rest, low-intensity exercise (dedicated hold stage), and peak exercise, as previously described.[18]

The low-intensity stage was defined as the period after surpassing the first ventilatory threshold but before reaching a heart rate (HR) of 100 bpm. Peak exercise measurements were obtained during a brief hold stage immediately before exhaustion, typically defined as a respiratory exchange ratio (RER) > 1.05, or earlier in case of symptom-limitation.

Stroke volume (SV) was calculated as the product of the time-velocity integral (VTI) at the left ventricular outflow tract and its cross-sectional area. CO was derived as SV x HR. Systolic pulmonary artery pressure (sPAP) was estimated from the maximal tricuspid regurgitation velocity (TRV) using continuous wave Doppler:

$$\mathbf{sPAP = 4 \times TRV^2}$$

Mean pulmonary artery pressure (mPAP) was calculated from sPAP using the Chemla formula (without an estimate of right atrial pressure):

$$\mathbf{mPAP = 0.61 \times sPAP + 2\ mmHg.}$$[19]

To enhance the TR envelope, an agitated colloid bolus (1-3 mL, tailored to TR quality) was routinely administered at rest and during exercise.[20] This method has been previously validated against invasive exercise hemodynamics.[21] All measurements were averaged over at least three cardiac cycles and recorded in a standardized data sheet. Image acquisition was

performed using a Vivid E9 or E95 ultrasound system (GE Vingmed Ultrasound AS), and offline analysis was performed using *EchoPAC* (v112, GE Healthcare).

Evaluation of the mPAP-CO relationship

The pressure-flow relationship between the pulmonary vasculature and CO was evaluated using four different non-invasive methods:

1. **Three-point slope** (linear regression across rest, low and peak),
2. **Two-point slope** (rest to peak),
3. **One-point mPAP/CO ratio** at low exercise,
4. **One-point mPAP/CO ratio** at peak exercise.

The three-point slope was derived from linear regression with mPAP as the dependent variable (y) and CO as the independent variable (x). Calculation was implemented via an embedded Excel® formula within a report-generating database. The two-point slope was calculated as the change in mPAP divided by the corresponding change in CO between rest and peak exercise.

Inter- and intra-observer variability analysis

To evaluate reproducibility, 40 CPETecho exams from X Hospital were randomly selected. Three experienced observers from different centers (X.X – X, Y.Y – Y, Z.Z – Z) independently measured mPAP and CO. All observers were blinded to prior measurements and to post-hoc slope calculations. To assess intra-observer variability, each observer repeated their measurements after a three-week interval.

Statistical Analysis

Prognostic accuracy of the four mPAP/CO methods was assessed using receiver operating characteristic (ROC) curves. Areas under the curve (AUCs) were compared using DeLong's test. Sensitivity, specificity, and optimal cut-off values were identified using the Youden index.

Reproducibility was evaluated using intraclass correlation coefficients (ICC) with 95% confidence intervals (CIs) for inter- and intra-observer agreement. Fleiss' kappa was used for agreement among all raters, and Cohen's kappa for intra-rater agreement, considering the identification of exPHT (mPAP/CO > 3.0 mmHg/L/min). Agreement between paired measurements was further assessed using Bland-Altman plots, reporting bias, and limits of agreement (±1.96 SD). Method comparisons were evaluated by analyzing differences in mean bias using one-way ANOVA and unpaired t-tests, while differences in variability among methods were assessed using Brown-Forsythe, Bartlett's, and F-tests.

Cox proportional hazards regression models were used to evaluate the association between the four methods and the composite endpoint of all-cause death or hospitalization for heart failure. Univariable models were first performed, followed by multivariable models adjusting for age and VO2peak (as a continuous variable) (model B) and predicted VO2peak (categorical – with 80% as the cut-off for reduced exercise capacity) and RER (categorical – with 1.05 as the cut-off for maximal effort) (model C). Results are reported as hazard ratios (HR) with corresponding 95% confidence intervals (CI). Logistic regression was performed, including each method, resting mPAP, and their interaction term.

All statistical analyses were conducted using *SPSS Statistics* (v29, IBM, Chicago, IL) and *MedCalc Software*. Figures were generated with *GraphPad Prism* (v10, GraphPad Software, San Diego, CA). P-values < 0.05 were considered statistically significant.

RESULTS

Study population

Eleven percent (521 patients) of the initial cohort were excluded due to non-feasibility or incomplete pulmonary pressure or CO measurements at one or more exercise levels (**Figure 1**). The final study cohort consisted of 1,619 patients with unexplained dyspnea. Baseline characteristics are summarized in **Table 1**. The mean age was 62 ± 15 years and 54% were women. The median NT-proBNP was 120 [52 - 209] ng/L.

Over a median follow-up of 28 months (IQR 16-45), 137 patients (8.5%) experienced the primary composite outcome, which included 66 deaths (4.1%) and 84 heart failure hospitalizations (5.2%).

Prognostic value of mPAP/CO methods

Among the four methods evaluated, the one-point mPAP/CO ratio at peak exercise demonstrated the highest prognostic performance, with an AUC of 0.715 (95% CI: 0.692–0.737). This was modest but significantly higher than the 2-point slope from rest to peak (AUC: 0.685; 95%CI 0.662–0.708; p=0.028) and the 3-point regression (AUC: 0.686, 95% CI: 0.663–0.709; p=0.028) (**Table 2, Figure 2**).

The one-point mPAP/CO ratio at low-level exercise showed comparable prognostic performance (AUC: 0.713; 95% CI: 0.690–0.735) to that observed at peak exercise. The cut-off values ranged from 3.4 to 3.6 mmHg/L/min. Sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and optimal cut-off values for each method are summarized in **Table 2**.

Reproducibility of mPAP/CO Methods

One-point mPAP/CO ratio assessments at low and peak exercise demonstrated excellent intra-observer reliability, with ICCs of 0.903 (95% CI: 0.863–0.932) and 0.923 (95% CI: 0.890–0.946), respectively (**Table 3** and **Supplementary Figure 1**). The 2-point slope and 3-point regression methods showed good intra-observer reliability, with ICCs of 0.874 (95% CI: 0.822–0.911) and 0.857 (95% CI: 0.799–0.899).

Inter-observer reliability was also good across all methods, with ICCs ranging from 0.802 to 0.880 (**Table 4** and **Supplementary Figure 1**). The highest Fleiss' kappa was observed for the peak one-point mPAP/CO ratio method (0.825, 95% CI: 0.610–1.000) (**Table 4** and **Supplementary Figure 2**).

Bland-Altman plots for intra-observer agreement (**Figure 3A**) demonstrated similar bias between methods (p=0.43), but revealed significantly narrower limits of agreement for the one-point approaches compared to the 2-point and 3-point slope methods (p<0.001), indicating superior repeatability. Similar patterns were observed in inter-observer analysis

(**Figure 3B**), where slope-based methods exhibited broader limits of agreement (p<0.001), reflecting reduced reproducibility.

Clinical Profile by one-point mPAP/CO ratio

Patients with a one-point mPAP/CO ratio at peak of >3.6 mmHg/L/min were older and more frequently had hypertension, diabetes, and atrial fibrillation. These patients also had larger left atrial volumes, higher NT-proBNP concentrations, and more advanced diastolic dysfunction at rest (**Table 1**).

Sensitivity analyses

Cox proportional hazards regression was performed as a sensitivity analysis to account for the time-to-event nature of the data and to assess the prognostic value of the four different methods (**Supplementary Table 1)**. All the methods were predictors of the composite endpoint in a univariable analysis (all p <0.001).

The one-point approach at low and peak exercise remained significant in a multivariable model A, which adjusted for age and cardiorespiratory fitness (CRF) as a continuous variable (p = 0.003 and 0.04, respectively), whereas the 2-point and 3-point methods were not (p = 0.14 and 0.06, respectively) (**Supplementary Table 1**). When CRF was modeled categorically (VO_2peak < 80% vs. ≥ 80% of predicted) in model B, mPAP/CO remained a significant predictor across all methods (all p < 0.001). In model C, which incorporated exercise intensity (RER < 1.05 vs. ≥ 1.05), mPAP/CO continued to demonstrate independent prognostic value across all models (p<0.001).

To assess whether a non-linear pressure-flow relationship influenced the prognostic value of the mPAP/CO, the association between CO and mPAP was evaluated across exercise stages. A linear relationship was observed between mPAP and the relative percentage of peak CO (**Supplementary Figure 3**). Resting mPAP had a small but significant effect on the prognostic performance of the mPAP-CO relationship, but this was similar across mPAP/CO methods (**Supplementary Table 2**).

DISCUSSION

Accurate assessment of exercise pulmonary hypertension (exPHT) is crucial for risk stratification and management of patients with unexplained dyspnea. In this large multicenter cohort study, we systematically compared the prognostic performance and reproducibility of four non-invasive methods for evaluating the mPAP-CO relationship during exercise. Our findings support two key conclusions: first, one-point mPAP/CO ratios, obtained either at low-level or peak exercise, provide prognostic performance comparable to more complex slope-based approaches; second, these simplified methods demonstrate superior reproducibility, with higher inter- and intra-rater reliability and narrower limits of agreement. Taken together, these findings support the use of one-point mPAP/CO ratio measurements as a practical, robust, and prognostically relevant strategy for evaluating exPHT in patients with unexplained dyspnea.

Recent studies have underscored the utility of non-invasive exercise testing for assessing pulmonary vascular response in patients with unexplained dyspnea and a range of cardiorespiratory disorders, complementing traditional resting evaluations that include pulmonary pressure assessment.[12,14,15] Non-invasive mPAP/CO slope estimates have been validated against invasive hemodynamics and shown to be feasible in up to 87% of cases when using agitated contrast.[20,21] However, no consensus exists on a standardized non-invasive approach, and calculation methods vary widely across studies.

This variability is clinically relevant: in an invasive study by Godinas et al., different criteria for defining exPHT led to only moderate concordance, with disagreement up to 22% of cases.[22] Such heterogeneity underscores the need for methodological consistency and supports the potential advantages of simplified, one-point approaches.

In our study, although the differences in AUC values between methods were modest (**Figure 2**), one-point measures demonstrated consistently better reproducibility (**Figure 3, Supplementary Figures 1-2**). This advantage may reflect the vulnerability of slope-based methods to compounding error, as they rely on multiple measurements that each carry some degree of imprecision.[23,24] In contrast, one-point estimates are inherently less susceptible to cumulative variability. This is consistent with the data in **Supplementary Figure 4**, which did not identify any specific component of the mPAP-CO relationship, such as stroke volume, CO, or mPAP, as a primary contributor to the observed differences between methods.

Sensitivity analysis confirmed that the one-point mPAP/CO ratio at low and peak exercise remained independently associated with outcome, even after adjustment for age and cardiorespiratory fitness (**Supplementary Table 1**). All methods were independently associated with outcome at different levels of physical effort and fitness.

Notably, neither the curvilinear pressure–flow relationship previously described in other studies[25,26] (**Supplementary Figure 3**), nor a “higher take-off” pattern associated with resting mPAP (**Supplementary Table 2**), appeared to explain the observed differences between methods. These findings support the notion that technical and methodological factors, rather than physiological mechanisms, may underlie the superior performance of one-point measures.

Stratification by mPAP/CO ratio revealed a more adverse clinical profile in patients with a steeper pressure–flow response, consistent with an HFpEF phenotype (**Table 1**). We recently demonstrated that this prognostic association was independent of and incremental to NT-proBNP, HFpEF probability scores, diastolic stress testing, and absolute pulmonary pressures, supporting the notion that these individuals may represent an early or exertional HFpEF phenotype not captured by current non-invasive diagnostic criteria.[15] Notably, patients with an elevated exercise mPAP/CO ratio but negative HFpEF scores showed no evidence of pulmonary vascular disease: they had normal ventilatory efficiency and reserve, and no hypoxemia during exercise, further suggesting a predominantly cardiac origin of their abnormal hemodynamic response.

Clinical implications

Our findings have direct implications for practice. Simplified one-point mPAP/CO ratio ratios offer a feasible and reproducible tool for evaluating total pulmonary resistance during exercise. By reducing methodological complexity and improving reproducibility, this approach facilitates broader implementation in both routine clinical follow-up and the evaluation of treatment effects in research.

Importantly, when peak exercise measures are not feasible, due to early termination, technical limitations, or patient discomfort, low-level one-point measurements provide comparable prognostic information and serve as a reliable alternative.

Higher inter- and intra-rater agreement further supports the consistency of one-point approaches across readers and institutions. For longitudinal research or clinical trials, their reproducibility makes them particularly suited for tracking change over time.

Study limitations

The study has several limitations. First, outcome data was collected retrospectively through chart review, which may have resulted in incomplete event capture. However, we used a nationwide digital platform that consolidates medical records from all Belgian hospitals, minimizing the risk of missed outcomes.

Second, patients with incomplete hemodynamic data at all three stages were excluded, potentially introducing selection bias. Nonetheless, this allowed for consistent within-subject comparisons across methods and improved data integrity.

Third, the ROC analysis did not account for the time-to-event nature of the data. However, a time-to-event analysis using Cox proportional hazards regression was performed in our primary investigation of this cohort[15], and a similar sensitivity analysis in the current study yielded consistent results, supporting the robustness of the prognostic comparisons across methods.

Fourth, there was an unbalanced distribution of patients across centers, with more than 80% recruited at a single site. This precluded robust center-adjusted analysis and highlights the need for external validation of the cut-off values in more balanced multicenter cohorts.

Lastly, all hemodynamic parameters were derived non-invasively. Although invasive measures remain the gold standard for diagnosing exPHT, the echocardiographic protocols used here have been previously validated against invasive exercise hemodynamics.[21] Moreover, the primary aim of this study was not to assess diagnostic accuracy, but to evaluate prognostic performance. All measurements were performed by experienced personnel using standardized protocols, and high reproducibility across observers supports the validity of our findings.

CONCLUSION

In this large, multicenter cohort of patients with unexplained exertional dyspnea, simple one-point mPAP/CO ratios during exercise provided prognostic value comparable to more complex two- or three-point slope-based methods, while offering significantly greater reproducibility. These findings support the use of the mPAP/CO ratio as a practical, reproducible, and

physiologically meaningful approach for the non-invasive evaluation of exercise pulmonary hypertension. Superior reproducibility is crucial for longitudinal follow-up and assessment of therapeutic interventions.

REFERENCES

1. Ramalho SHR, Santos M, Claggett B, et al. Association of Undifferentiated Dyspnea in Late Life With Cardiovascular and Noncardiovascular Dysfunction: A Cross-sectional Analysis From the ARIC Study. *JAMA Netw Open*. 2019;2(6):e195321-e195321. doi:10.1001/JAMANETWORKOPEN.2019.5321

2. Parshall MB, Schwartzstein RM, Adams L, et al. An official American Thoracic Society statement: update on the mechanisms, assessment, and management of dyspnea. *Am J Respir Crit Care Med*. 2012;185(4):435-452. doi:10.1164/RCCM.201111-2042ST

3. Frostad A, Søyseth V, Andersen A, Gulsvik A. Respiratory symptoms as predictors of all-cause mortality in an urban community: a 30-year follow-up. *J Intern Med*. 2006;259(5):520-529. doi:10.1111/J.1365-2796.2006.01631.X

4. Ho JE, Zern EK, Lau ES, et al. Exercise Pulmonary Hypertension Predicts Clinical Outcomes in Patients With Dyspnea on Effort. *J Am Coll Cardiol*. 2020;75(1):17-26. doi:10.1016/J.JACC.2019.10.048

5. Santos M, Kitzman DW, Matsushita K, Loehr L, Sueta CA, Shah AM. Prognostic Importance of Dyspnea for Cardiovascular Outcomes and Mortality in Persons without Prevalent Cardiopulmonary Disease: The Atherosclerosis Risk in Communities Study. *PLoS One*. 2016;11(10). doi:10.1371/JOURNAL.PONE.0165111

6. Ryan JJ, Waxman AB. The dyspnea clinic. *Circulation*. 2018;137(19):1994-1996. doi:10.1161/CIRCULATIONAHA.117.029769/ASSET/31B70E03-2095-4A86-8813-5B482BD3B93A/ASSETS/IMAGES/LARGE/1994FIG01.JPG

7. Verwerft J, Bertrand PB, Claessen G, Herbots L, Verbrugge FH. Cardiopulmonary Exercise Testing With Simultaneous Echocardiography: Blueprints of a Dyspnea Clinic for Suspected HFpEF. *JACC Heart Fail*. 2023;11(2):243-249. doi:10.1016/J.JCHF.2022.11.004

8. Kovacs G, Berghold A, Scheidl S, Olschewski H. Pulmonary arterial pressure during rest and exercise in healthy subjects: a systematic review. *Eur Respir J*. 2009;34(4):888-894. doi:10.1183/09031936.00145608

9. La Gerche A, MacIsaac AI, Burns AT, et al. Pulmonary transit of agitated contrast is associated with enhanced pulmonary vascular reserve and right ventricular function during exercise. *J Appl Physiol*. 2010;109(5):1307-1317. doi:10.1152/japplphysiol.00457.2010

10. Tolle JJ, Waxman AB, Van Horn TL, Pappagianopoulos PP, Systrom DM. Exercise-induced pulmonary arterial hypertension. *Circulation*. 2008;118(21):2183-2189. doi:10.1161/CIRCULATIONAHA.108.787101

11. Humbert M, Kovacs G, Hoeper MM, et al. 2022 ESC/ERS Guidelines for the diagnosis and treatment of pulmonary hypertension: Developed by the task force for the diagnosis and treatment of pulmonary hypertension of the European Society of Cardiology (ESC) and the European Respiratory Society (ERS). Endorsed by the International Society for Heart and Lung Transplantation (ISHLT) and the European Reference Network on rare respiratory diseases (ERN-LUNG). *Eur Heart J*. 2022;43(38):3618-3731. doi:10.1093/EURHEARTJ/EHAC237

12. Gargani L, Pugliese NR, De Biase N, et al. Exercise Stress Echocardiography of the Right Ventricle and Pulmonary Circulation. *J Am Coll Cardiol*. 2023;82(21):1973-1985. doi:10.1016/J.JACC.2023.09.807

13. Kovacs G, Humbert M, Avian A, et al. Prognostic relevance of Exercise Pulmonary Hypertension: Results of the multi-center PEX-NET Clinical Research Collaboration. *European Respiratory Journal*. 2024;64(6). doi:10.1183/13993003.00698-2024

14. Hoedemakers S, Pugliese NR, Stassen J, et al. mPAP/CO Slope and Oxygen Uptake Add Prognostic Value in Aortic Stenosis. *Circulation*. 2024;149(15):1172-1182. doi:10.1161/CIRCULATIONAHA.123.067130

15. Falter M, Bekhuis Y, L'Hoyes W, et al. Exercise Echocardiography for Risk Stratification in Unexplained Dyspnea: the Incremental Value of the mPAP/CO slope. *Journal of the American Society of Echocardiography*. 2025;0(0). doi:10.1016/j.echo.2025.06.007

16. Pugliese NR, Mazzola M, Madonna R, et al. Exercise-induced pulmonary hypertension in HFpEF and HFrEF: Different pathophysiologic mechanism behind similar functional impairment. *Vascul Pharmacol*. 2022;144. doi:10.1016/j.vph.2022.106978

17. De Biase N, Del Punta L, L'Hoyes W, et al. Associations of iron deficiency with cardiac function, congestion, exercise capacity and prognosis in heart failure. *Eur J Heart Fail*. 2025;27(5):889-900. doi:10.1002/EJHF.3534

18. Verwerft J, Soens L, Wynants J, et al. Heart failure with preserved ejection fraction: relevance of a dedicated dyspnoea clinic. *Eur Heart J*. 2023;44(17):1544-1556. doi:10.1093/EURHEARTJ/EHAD141

19. Chemla D, Castelain V, Humbert M, et al. New formula for predicting mean pulmonary artery pressure using systolic pulmonary artery pressure. *Chest*. 2004;126(4):1313-1317. doi:10.1378/CHEST.126.4.1313

20. Bekhuis Y, Verwerft J, Claessen G. Exercise Echocardiography for mPAP/CO Slope Estimation: A New Standard for Routine Clinical Care? *J Am Coll Cardiol*. 2024;83(14):e141-e142. doi:10.1016/J.JACC.2023.12.039

21. Claessen G, La Gerche A, Voigt JU, et al. Accuracy of Echocardiography to Evaluate Pulmonary Vascular and RV Function During Exercise. *JACC Cardiovasc Imaging*. 2016;9(5):532-543. doi:10.1016/j.jcmg.2015.06.018

22. Godinas L, Lau EM, Chemla D, et al. Diagnostic concordance of different criteria for exercise pulmonary hypertension in subjects with normal resting pulmonary artery pressure. *European Respiratory Journal*. 2016;48(1):254-257. doi:10.1183/13993003.01678-2015

23. Hutcheon JA, Chiolero A, Hanley JA. Random measurement error and regression dilution bias. *BMJ*. 2010;340(7761):1402-1406. doi:10.1136/BMJ.C2289

24. Groenwold RHH, Dekkers OM. Measurement error in clinical research, yes it matters. *Eur J Endocrinol*. 2020;183(3):E3-E5. doi:10.1530/EJE-20-0550,

25. Naeije R, Vanderpool R, Dhakal BP, et al. Exercise-induced pulmonary hypertension: physiological basis and methodological concerns. *Am J Respir Crit Care Med*. 2013;187(6):576-583. doi:10.1164/RCCM.201211-2090CI

26. Lewis GD, Bossone E, Naeije R, et al. Pulmonary vascular hemodynamic response to exercise in cardiopulmonary diseases. *Circulation*. 2013;128(13):1470-1479. doi:10.1161/CIRCULATIONAHA.112.000667

Figure captions

Figures (separate files)

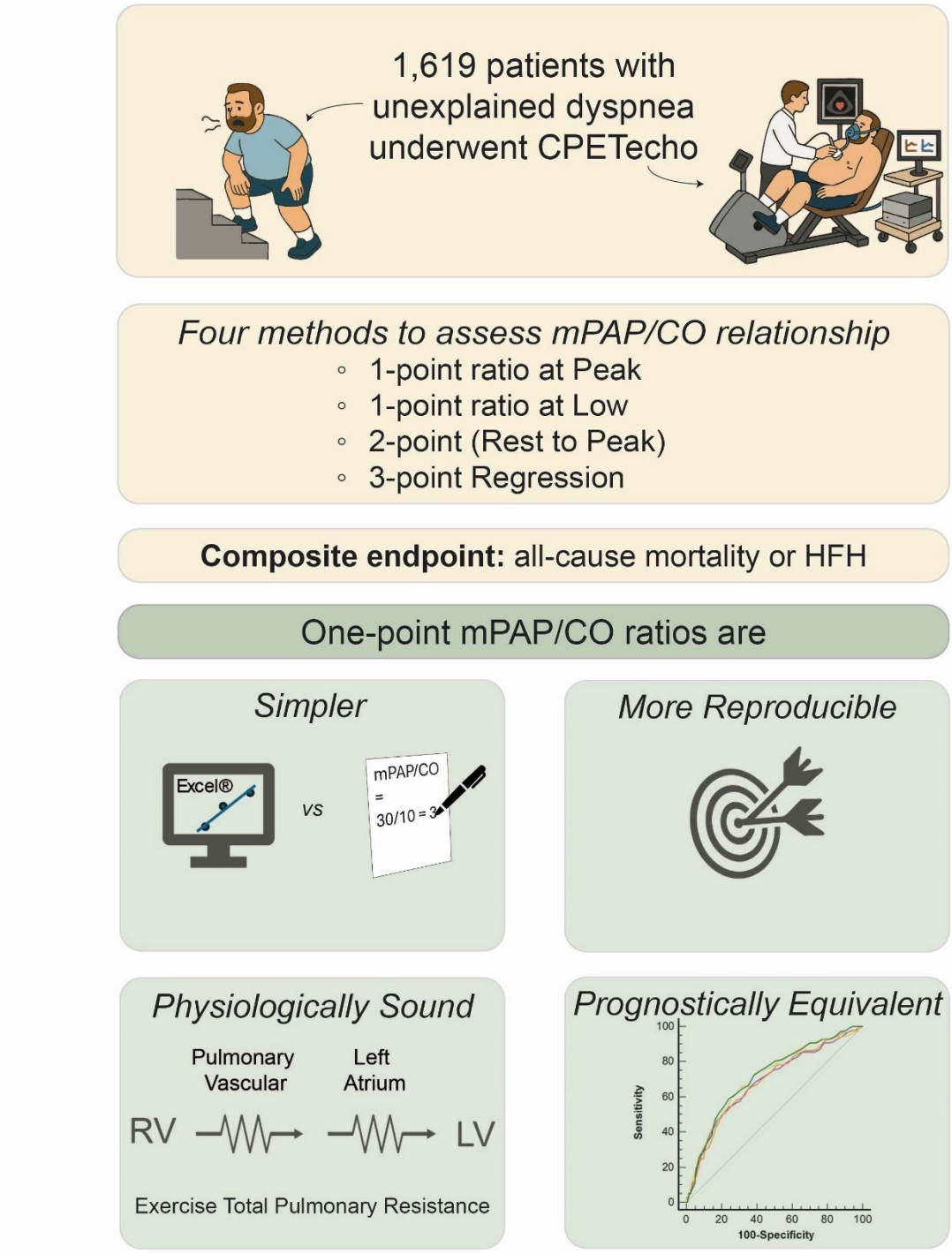
Central Figure (Visual Take Home Graphic): Study design and key findings

Figure 1. Flowchart of study population

Figure 2. Receiver operating characteristic (ROC) analysis comparing the prognostic accuracy of different methods for assessing the mPAP-CO relationship

Figure 3. Bland-Altman plots for intra- and inter-rater agreement

435 **Central illustration**



436

437

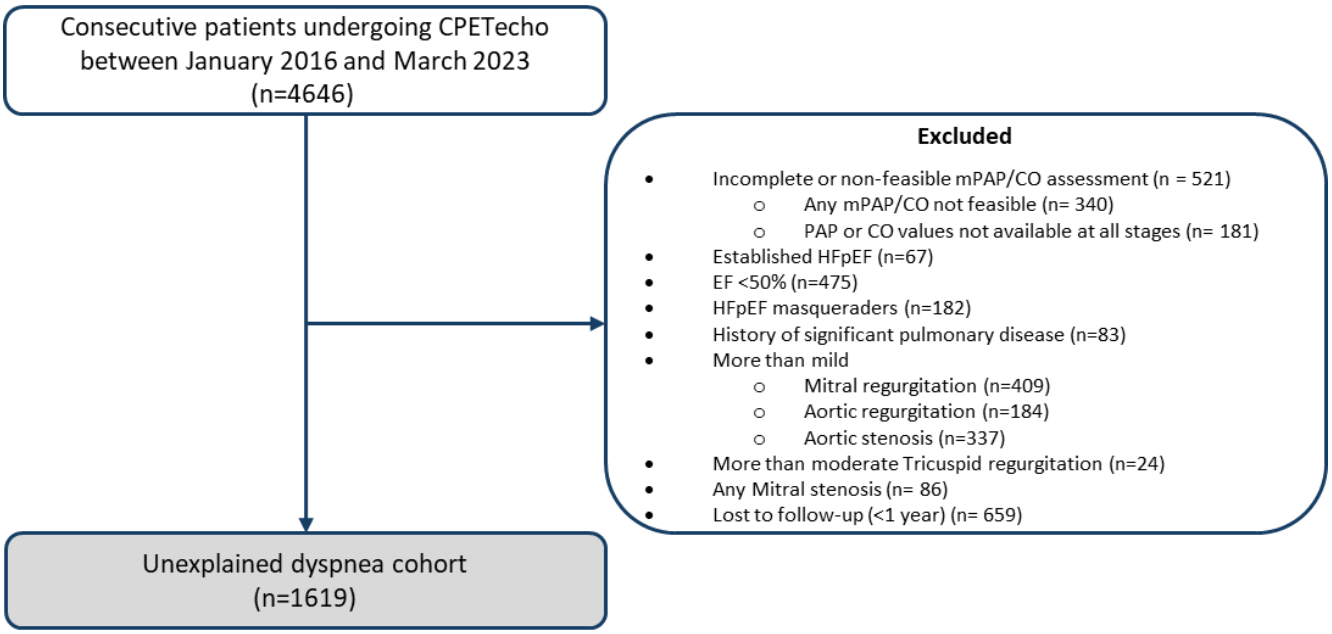


Figure 1: *CPETecho, Cardiopulmonary Exercise Test combined with echocardiography; CO, cardiac output; PAP, pulmonary artery pressure; EF, ejection fraction; HFpEF, Heart Failure with Preserved Ejection Fraction*

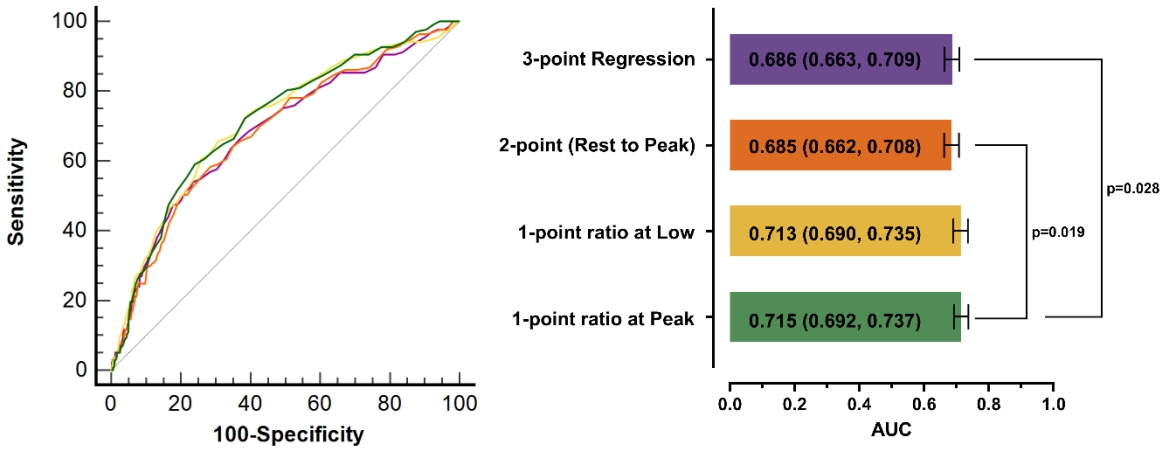


Figure 2: *Receiver operating characteristic (ROC) curves for the different mPAP/CO methods. Bars indicate the area under the curve (AUC) and 95% confidence intervals. P-values reflect the statistical significance of differences in AUC between methods.*

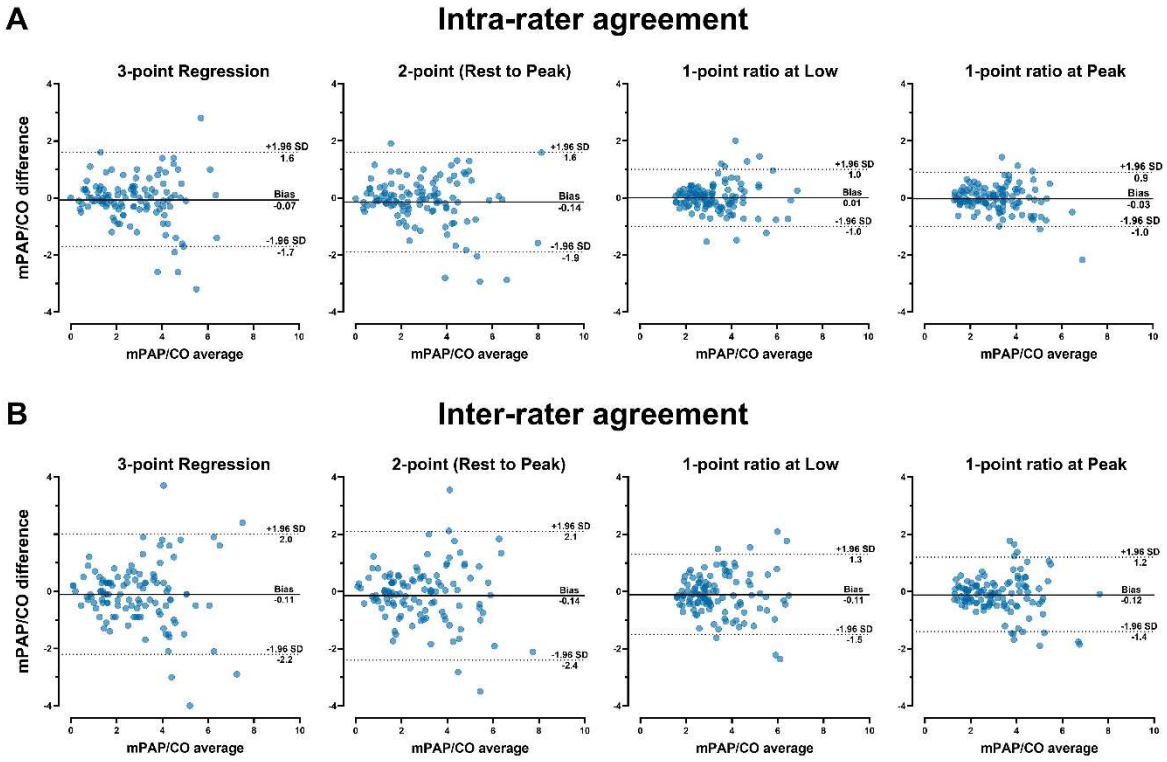


Figure 3: *Bland-Altman plots illustrating the mean difference (bias) and 95% limits of agreement (± 1.96 SD) for four different methods to assess the mPAP-CO relationship. Panel A shows intra-rater agreement; Panel B shows inter-rater agreement. Each point represents an individual paired comparison between two measurements by the same rater (Panel A) or between two different raters (Panel B).*

CO, cardiac output; mPAP, mean pulmonary artery pressure; SD, standard deviation

Table captions

Tables (separate files)

Table 1. Patient characteristics

Table 2. Diagnostic accuracy according of methods used to calculate the mPAP-CO relationship.

Table 3. Intra-rater reliability

Table 4. Inter-rater reliability

	Total cohort	mPAP/CO ratio ≤3.6	mPAP/CO ratio >3.6	p-value
Number of patients	1619	1159	460	
Demographics				
Age (years)	62±15	58±16	72±9	<0.001
Female, n (%)	872 (54)	578 (49.9%)	164 (35.7%)	<0.001
Systolic blood pressure (mmHg)	137±22	136±24	140±25	0.003
Diastolic blood pressure (mmHg)	79±14	80±14	77±16	<0.001
Body Mass Index ($kg.m^{-2}$)	27±5	27±5	27±5	0.030
Hypertension, n (%)	728 (45)	441 (38)	287 (62)	<0.001
Diabetes, n (%)	189 (12)	121 (10)	68 (15)	0.016
Atrial Fibrillation history, n (%)	296 (18)	145 (13)	151 (33)	<0.001
Biomarker				
NT-proBNP ($ng.L^{-1}$)	120 (52, 290)	84 (40,180)	290 (140,570)	<0.001
Echocardiography (Rest)				
LAVi ($mL.m^{-2}$)	21±14	20±13	24±17	<0.001
E/A	1.1±0.6	1.1±0.6	1.1±0.7	0.081
E/e'	11.2±4.6	10.0±3.5	14.3±5.9	<0.001
TRPG (mmHg)	22.7±6.1	21.1±4.8	26.7±7.1	<0.001
Outcome				
Deaths, n (%)	66 (4.1)	32 (2.8%)	34 (7.4%)	<0.001
Heart failure hospitalizations, n (%)	84 (5.2)	29 (2.5%)	55 (12.0%)	<0.001
Composite outcome, n (%)	137 (8.5)	56 (4.8%)	81 (17.6%)	<0.001

Table 1: *Baseline characteristics of the total cohort (n=1619), stratified by one-point mPAP/CO ratio at peak exercise (≤3.6 vs. >3.6 mmHg/L/min). Data presented as mean ± standard deviation, median (25^{th} – 75^{th} percentiles), or number (%). P-values reflect between-group comparisons using t-tests or chi-square tests as appropriate. NT-proBNP, N-terminal pro-B-type natriuretic peptide; LAVi, left atrium volume index; TRG, Tricuspid Regurgitation Pressure Gradient.*

	AUC (95%CI)	Optimal Cut-off	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)
3-point Regression	0.686 (0.663-0.709)	> 3.5	54.0	76.4	17.5	94.7
2-point (Rest to Peak)	0.685 (0.662-0.708)	> 3.4	58.4	71.5	16.0	94.9
1-point ratio at Low	0.713 (0.690-0.735)	> 3.5	65.7	69.3	16.6	95.6
1-point ratio at Peak	0.715 (0.692-0.737)	> 3.6	59.1	76.1	18.7	95.2

***Table 2**: Prognostic performance of the four methods used to assess the mPAP-CO relationship. Area under the curve (AUC) with 95% confidence intervals (CI), and optimal cut-off values based on the Youden Index are presented for each method. Sensitivity and specificity are reported according to the respective optimal cut-off values. Positive predictive value (PPV) and negative predictive value (NPV) were calculated assuming a prevalence of 8.5% (137 events among 1619 patients).*

CI, confidence interval; CO, cardiac output; mPAP, mean pulmonary artery pressure.

mPAP/CO method	First	Second	R (Pearson)	ICC	Cohen's Kappa*
3-point Regression	2.7 ±1.5	2.8 ±1.6	0.859	0.857 (0.799-0.899)	0.875
2-point (Rest to Peak)	2.9 ± 1.7	3.0 ± 1.9	0.884	0.874 (0.822-0.911)	0.841
1-point ratio at Low	3.3 ± 1.2	3.2 ± 1.2	0.902	0.903 (0.863-0.932)	0.772
1-point ratio at Peak	3.1 ± 1.2	3.1 ± 1.3	0.926	0.923 (0.890-0.946)	0.842

Table 3: *Intra-observer reliability for the four methods used to calculate the mPAP-CO relationship. Data is expressed as mean ± SD for the first and second measurements by the same rater. Pearson correlation coefficients (R), intraclass correlation coefficients (ICC) with 95% confidence intervals, and Cohen's kappa values are reported.*
** Categorization of mPAP/CO > 3.0 mmHg/L/min.*
CO, cardiac output; ICC, intraclass correlation coefficient; mPAP, mean pulmonary artery pressure.

mPAP/CO method	Rater 1	Rater 2	Rater 3	ICC	Fleiss' Kappa*
3-point Regression	2.6 ± 1.4	3.1 ± 2.1	2.8 ± 1.5	0.802 (0.682-0.885)	0.681 (0.497-0.864)
2-point (Rest to Peak)	2.8 ± 1.6	3.2 ± 2.1	3.0 ± 2.1	0.828 (0.728-0.899)	0.718 (0.534-0.901)
1-point ratio at Low	3.1 ± 1.2	3.8 ± 1.5	3.2 ± 1.2	0.851 (0.559-0.938)	0.684 (0.501-0.868)
1-point ratio at Peak	3.0 ± 1.2	3.5 ± 1.5	3.2 ± 1.3	0.880 (0.747-0.941)	0.825 (0.641-1.000)

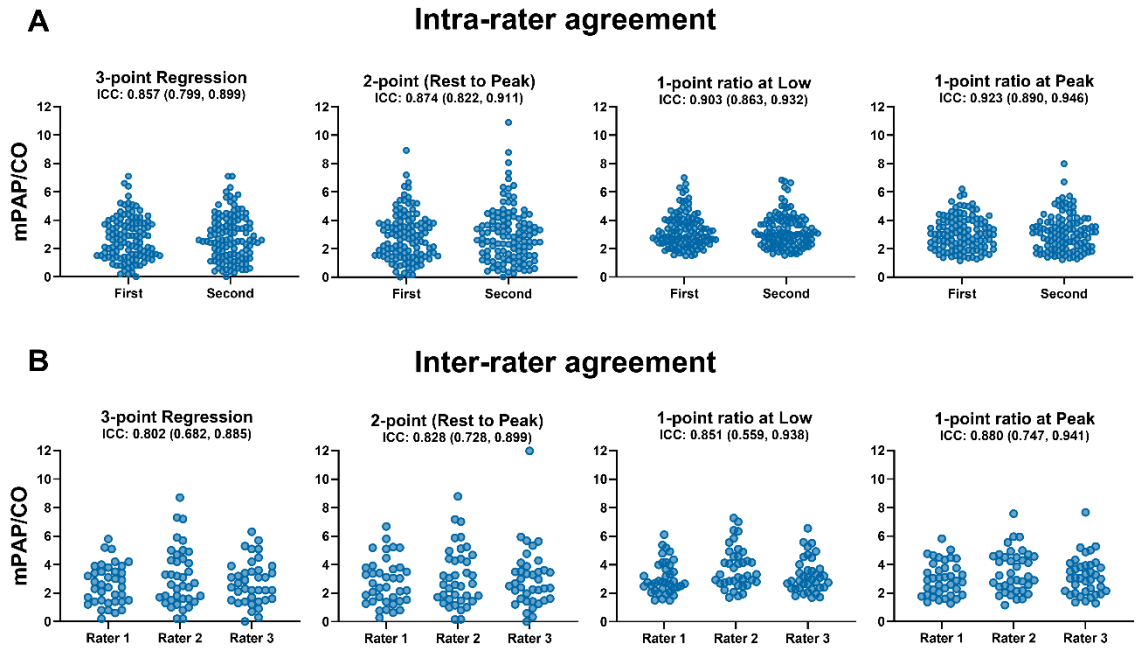
Table 4: *Inter-observer reliability for the four methods used to calculate the mPAP-CO relationship. Data is expressed as mean ± SD for each rater. Intraclass correlation coefficients (ICC) with 95% confidence intervals and Fleiss' kappa values are reported.*

** Categorization of mPAP/CO > 3.0 mmHg/L/min*

CO, cardiac output; ICC, intraclass correlation coefficient; mPAP, mean pulmonary artery pressure.

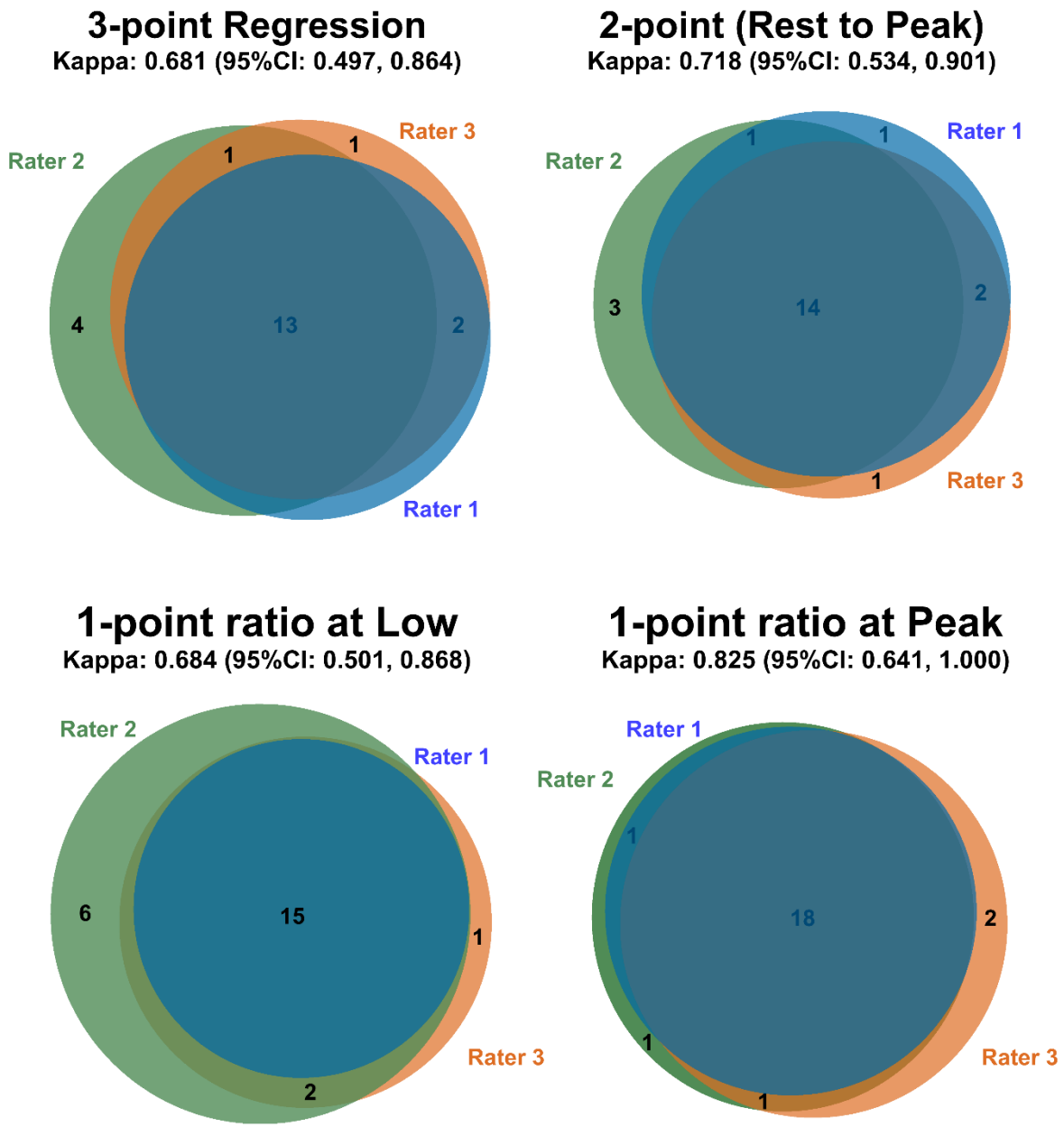
SUPPLEMENTARY MATERIALS

Supplementary Figure 1. Intra- and inter-rater variability by intraclass correlation coefficient (ICC)



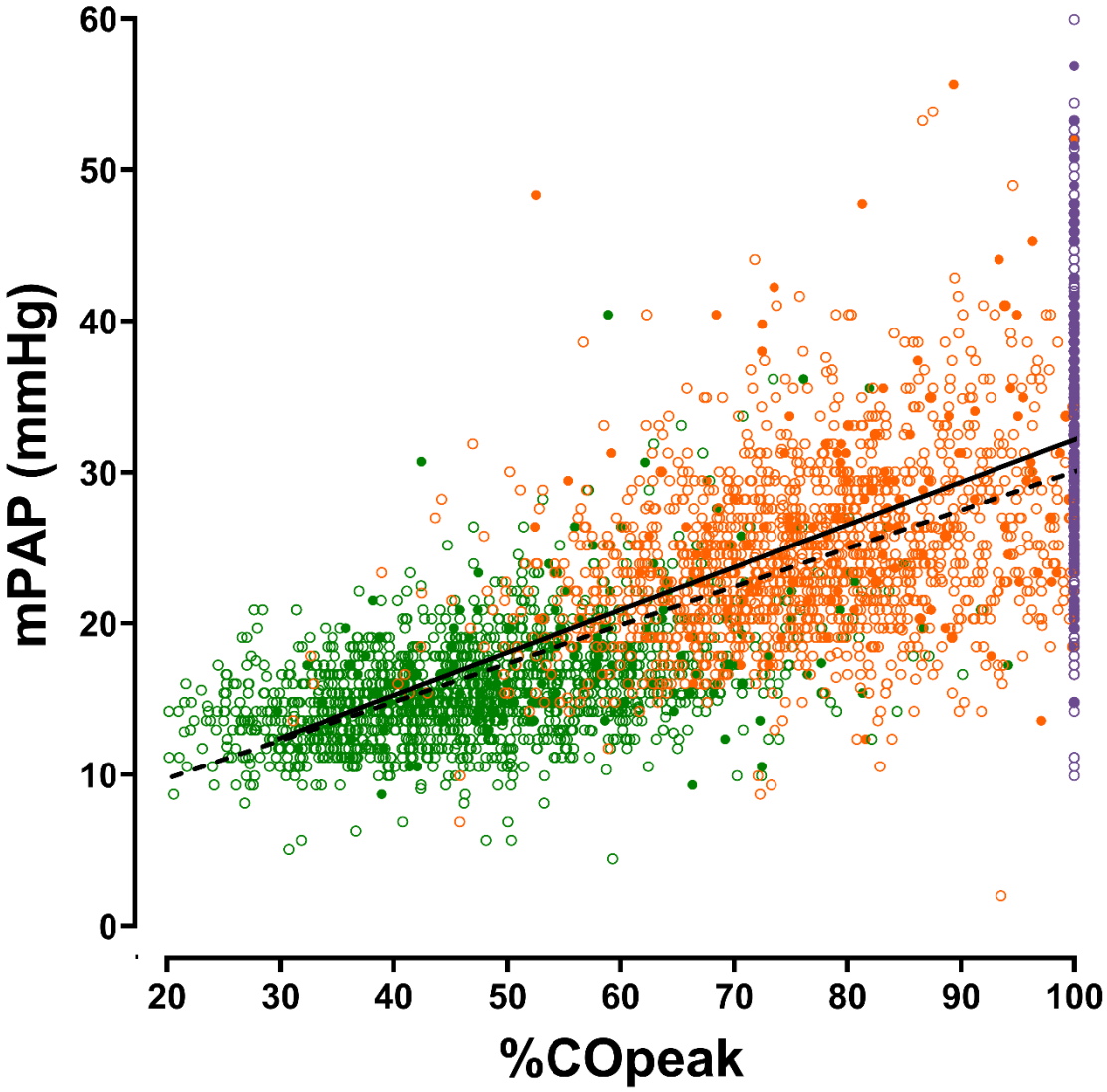
Supplementary Figure 1: *Data expressed as intraclass correlation coefficient (ICC) with its corresponding 95% confidence interval.*

Supplementary Figure 2: Venn diagram for inter-rater agreement in identifying exercise pulmonary hypertension.



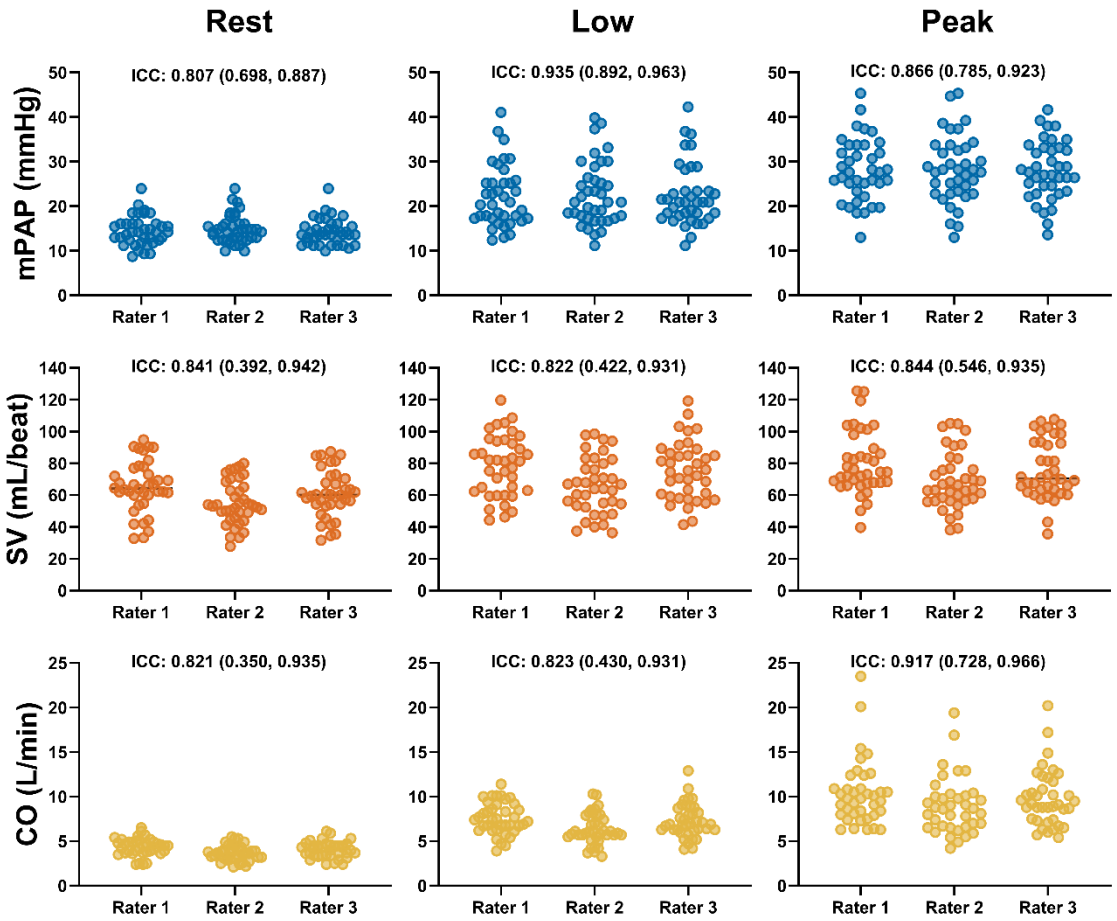
Supplementary Figure 2: *Inter-rater agreement across the four methods for evaluating the mPAP-CO relationship. Venn diagrams illustrate diagnostic concordance among three raters. Fleis' kappa values (with 95% CI) indicate the level of agreement for each method, with the highest reproducibility observed for one-point mPAP/CO ratio at peak exercise.*

Supplementary Figure 3: Hemodynamic response of mPAP to increasing Cardiac Output during Exercise



Supplementary Figure 3: *Individual measurements during exercise are plotted across different stages (green: rest; orange: low-level exercise; purple: peak exercise). Cardiac output (CO) is expressed as a percentage of peak cardiac output. A linear relationship between CO and mPAP is observed from rest to low-level exercise (solid line: individuals with events; dashed line: individuals without events).*
CO, cardiac output; mPAP, mean pulmonary artery pressure

Supplementary Figure 4: Interrater variability expressed as intraclass correlation coefficient (ICC) for the individual components of the mPAP-CO relationship.



Supplementary Figure 4: *Data expressed as intraclass correlation coefficient (ICC) with its corresponding 95% confidence interval.*

Evaluation of the individual components of the mPAP-CO relationship did not reveal any single factor that accounted for the observed differences in ICCs across methods. Slope-based approaches, which rely on multiple measurements each subject to some degree of imprecision, are more susceptible to compounding error. In contrast, one-point mPAP/CO ratios are inherently less affected by cumulative variability

Supplementary Table 1. Cox regression univariate and multivariate models for mPAP/CO methods.

mPAP/CO method	3-point Regression		2-point (Rest to Peak)		1-point ratio at Low		1-point ratio at Peak	
	Hazard ratio	Harrell's C-Index	Hazard ratio	Harrell's C-Index	Hazard ratio	Harrell's C-Index	Hazard ratio	Harrell's C-Index
Univariate models with mPAP/CO (n = 1619, 137 events)								
mPAP/CO method	1.18 (1.13-1.23) p < 0.001	0.696 (0.644-0.748)	1.15 (1.10-1.19) p < 0.001	0.694 (0.643-0.744)	1.52 (1.38-1.67) p < 0.001	0.715 (0.666-0.764)	1.45 (1.32-1.59) p < 0.001	0.720 (0.674-0.766)
Multivariate model A (mPAP/CO, age and cardiorespiratory fitness as a continuous variable; n = 1583, 134 events)								
mPAP/CO method	1.06 (1.00-1.12) p = 0.06	0.752 (0.712-0.793)	1.04 (0.99-1.10) p = 0.14	0.750 (0.709-0.790)	1.21 (1.07-1.38) p = 0.003	0.761 (0.719, 0.802)	1.14 (1.01-1.29) p = 0.04	0.755 (0.715-0.796)
Age, year	1.05 (1.04-1.07) p < 0.001		1.06 (1.04-1.07) p < 0.001		1.05 (1.03-1.07) p < 0.001		1.05 (1.03-1.07) p < 0.001	
VO_2peak, ml/kg/min	0.96 (0.94-0.99) p = 0.001		0.96 (0.94-0.98) p < 0.001		0.97 (0.95-0.99) p = 0.004		0.96 (0.94-0.99) p = 0.002	
Multivariate model B (mPAP/CO and cardiorespiratory fitness as a categorical variable; n = 1583, 134 events)								
mPAP/CO method	1.15 (1.10-1.20) p < 0.001	0.714 (0.666, 0.763)	1.12 (1.07-1.17) p < 0.001	0.712 (0.664-0.761)	1.44 (1.30-1.60) p < 0.001	0.729 (0.682-0.777)	1.37 (1.24-1.51) p < 0.001	0.724 (0.678-0.770)

VO_2peak (< 80% *versus* ≥ 80 of predicted)	2.33 (1.54-3.45) p < 0.001		2.33 (1.56-3.57) p < 0.001		2.13 (1.41-3.23) p < 0.001		2.13 (1.43-3.23) p < 0.001	
Multivariate model C (mPAP/CO, cardiorespiratory fitness as a categorical variable and intensity of effort; n = 1350, 97 events)								
mPAP/CO method	1.18 (1.11-1.27) p < 0.001	0.716 (0.661-0.771)	1.14 (1.07-1.21) p < 0.001	0.714 (0.660-0.768)	1.42 (1.25-1.60) p < 0.001	0.727 (0.670-0.785)	1.34 (1.19-1.51) p < 0.001	0.723 (0.668-0.778)
VO_2peak (< 80% *versus* ≥ 80 of predicted)	2.15 (1.40-3.31) p < 0.001		2.18 (1.41-3.36) p < 0.001		2.06 (1.34-3.18) p = 0.001		2.07 (1.34-3.20) p = 0.001	
RERpeak (< 1.05 *versus* ≥ 1.05)	1.75 (1.14-2.70) p = 0.011		1.78 (1.15-2.74) p = 0.001		1.64 (1.06-2.54) p = 0.026		1.63 (1.05-2.53) p = 0.029	

Supplementary Table 1: *Data expressed as Hazard Ratio or Harrell's C-Index, and 95% confidence interval. Univariable and multivariable Cox models are shown for the four mPAP/CO methods. Multivariable model A adjusts for age and cardiorespiratory fitness (VO_2peak, continuous); model B adjusts for VO_2peak as a categorical variable (<80% vs. ≥ 80% predicted); and model C additionally includes effort level (RERpeak <1.05 vs ≥ 1.05).*

In the univariable Cox regression, all four mPAP/CO methods were individually associated with clinical outcomes. In multivariable model A, only the one-point mPAP/CO ratio at low and peak remained significant, while the 3-point regression and 2-point (rest to peak) did not. Age and VO_2peak remained independent predictors. In multivariable model B (VO_2peak categorical), all models remained statistically significant. In multivariable model C (VO_2peak and intensity of effort), all models remained statistically significant.

Supplementary Table 2: Multivariable logistic regression models evaluating the interaction between resting mPAP and the three-point regression method (model A), two-point slope (model B), the one-point mPAP/CO ratio at low exercise (model C), the one-point mPAP/CO ratio at peak exercise (model D).

Variable	β	Sig. (p)	Exp(B) (95% CI)	Model
3-point Regression	0.586	<0.001	1.796 (1.397–2.310)	Model A
mPAP (rest)	0.223	<0.001	1.250 (1.158–1.349)	Model A
Interaction (mPAP rest × 3-point Regression)	-0.023	<0.001	0.978 (0.965–0.991)	Model A
2-point (Rest to Peak)	0.469	<0.001	1.599 (1.286–1.989)	Model B
mPAP (rest)	0.208	<0.001	1.231 (1.146–1.321)	Model B
Interaction (mPAP rest × 2-point (Rest to Peak))	-0.018	0.002	0.982 (0.971-0.994)	Model B
1-point mPAP/CO ratio at low	0.764	<0.001	2.147 (1.393–3.308)	Model C
mPAP (rest)	0.207	<0.001	1.230 (1.095–1.381)	Model C
Interaction (mPAP (rest) × 1-point mPAP/CO ratio at low)	-0.025	0.027	0.975 (0.954–0.997)	Model C
1-point mPAP/CO ratio at peak	0.911	<0.001	2.488 (1.593–3.884)	Model D
mPAP (rest)	0.23	<0.001	1.258 (1.116–1.418)	Model D
Interaction (mPAP (rest) × 1-point mPAP/CO ratio at peak)	-0.031	0.008	0.969 (0.947–0.992)	Model D

Supplementary Table 2: *Results are presented as beta coefficients (β), which estimate the change in the Log odds of the outcome for a one-unit increase in the predictor. Odds ratios (Exp(B)) with corresponding 95% confidence intervals (95% CI) are also reported.*

The one-point mPAP/CO ratio at peak exercise showed a stronger association with outcome compared to the three-point slope and the two-point slope (β = 0.911 vs. β = 0.469 vs. β = 0.586, respectively). When mPAP at rest was higher, the additional risk contributed by each method was slightly lower, as reflected by the negative interaction terms (model A: β = -0.023, model B: -0.018, model C: -0.025, model D: -0.031). However, the magnitude of the interaction effect was low and comparable between methods, suggesting that resting mPAP only modestly attenuates the prognostic value of the different methods.

CO, cardiac output; mPAP, mean pulmonary artery pressure

Supplementary Table 3: Correlation and collinearity diagnostic for mPAP/CO methods

Variable	Correlation matrix (mPAP at rest)	Variance Inflation Factors (VIF)
3-point Regression	0.34	1.13
2-point (Rest to Peak)	0.32	1.12
1-point ratio at Low	0.55	1.44
1-point ratio at Peak	0.54	1.42

Supplementary Table 3: This table shows the pairwise correlation between mPAP and the four mPAP/CO methods, as well as their corresponding Variance Inflation Factors (VIF). Correlations were modest (r = 0.32–0.55), and all VIF values were low (1.11–1.44), indicating no evidence of problematic multicollinearity.

CO, cardiac output; mPAP, mean pulmonary artery pressure