Development of an open-source tool for risk assessment in pulmonary endarterectomy

Short title: A novel risk assessment tool for PEA

Authors:

James Liley PhD*[1], Katherine Bunclark MBChB*[2], Michael Newnham PhD[3], John Cannon PhD[2], Karen Sheares PhD[2], Dolores Taboada MD[2], Choo Ng FRCS[2], Nicholas Screaton FRCR[2], David Jenkins FRCS[2], Joanna Pepke-Zaba PhD[2], Mark Toshner MD [2,4]

*joint first authors

Affiliations:

[1] Durham Biostatistics Unit, Durham University, UK

[2] Pulmonary Vascular Diseases Unit, Royal Papworth Hospital, Cambridge, UK

[3] Institute of Applied Health Research, University of Birmingham, UK

[4] VPD Heart & Lung Institute, University of Cambridge, UK

Corresponding author:

Dr Katherine Bunclark

Pulmonary Vascular Diseases Unit, Royal Papworth Hospital, Cambridge Biomedical Campus

CB2 0AY

Email: katherine.bunclark@nhs.net **Tel:** +44 (0)1223 639697

Total word count: 3511 words

Abstract

Background

Risk prediction tools are routinely utilised in cardiothoracic surgery but have not been developed for pulmonary endarterectomy (PEA). There is no data on whether patients undergoing PEA may benefit from a tailored risk modelling approach. We develop and validate a clinically-usable tool to predict PEA 90-day mortality (90DM) with the secondary aim of informing factors that may influence five-year mortality (5YM) and improvement in patient-reported outcomes (PROchange) using common clinical assessment parameters. Derived model predictions were compared to those of the currently most widely implemented cardiothoracic surgery risk tool, EuroSCORE II.

Methods

Consecutive patients undergoing PEA for chronic thromboembolic pulmonary hypertension (CTEPH) between 2007 and 2018 (n = 1334) were included in a discovery dataset. Outcome predictors included an intentionally broad array of variables, incorporating demographic, functional and physiological measures. Three statistical models (linear regression, penalised linear regression and random forest) were considered per outcome, each calibrated, fitted and assessed using cross-validation, ensuring internal consistency. The best predictive models were incorporated into an open-source PEA risk tool and validated using a separate prospective PEA cohort from 2019 to 2021 (n = 443) at the same institution.

Results

Random forest models had the greatest predictive accuracy for all three outcomes. Novel risk models had excellent discriminatory ability for outcome 90DM (AUROC 0.82) outperforming that of EuroSCORE II (AUROC 0.65). CTEPH related factors were important for outcome 90DM but 5YM was driven by non-CTEPH factors, dominated by generic cardiovascular risk. We were unable to accurately predict a positive improvement in PRO status (AUROC 0.47).

Conclusions

Operative mortality from PEA can be predicted pre-operatively to a potentially clinically useful degree. Our validated models enable individualised risk stratification at clinician point-of-care to better inform shared decision making.

Abstract word count: 284/350

Introduction

Chronic Thromboembolic Pulmonary Hypertension (CTEPH) is an infrequent but important complication of acute pulmonary embolism which, if left untreated, results in progressive right ventricular failure and death [1]. Pulmonary Endarterectomy (PEA) is potentially curative and is the guideline recommended treatment in those with surgically accessible disease [2 – 5].

Prognostic outcomes from PEA continue to improve, reflective of evolving procedural experience and surgical expertise. Contemporaneous data indicates in-hospital mortality rates are now as low as < 5% with one-year survival > 90% in high-volume units [4, 5]. Nonetheless, there remains an unavoidable operative risk associated with PEA, even in the best centres. In-light-of emerging non-surgical treatment options in CTEPH, such as Balloon Pulmonary Angioplasty, the ability to identify those at greatest operative risk has the potential to inform both patient selection and choice. Whilst risk stratification tools have been in use routinely in cardiothoracic surgery for some time, with the most widely adopted system, European System for Cardiac Operative Risk Evaluation II (EuroSCORE II, [6]) validated for predicting inpatient mortality across a number of major cardiothoracic procedures [7, 8], none have been tested and validated for use in PEA.

Whilst PEA carries with it generic operative risks of any major cardiac surgical procedure, it is also unique in its effects on the pulmonary vasculature. With this in mind, several authors have attempted to identify particular factors specific to PEA that predict mortality outcomes. Identified pre-operative independent predictors have included: New York Heart Association (NYHA) functional class [9, 10], mean pulmonary artery pressure [11], pulmonary vascular resistance [9 - 12] and age [11]. Whilst these studies have provided useful information on post-operative mortality risk across PEA cohorts there remains limited data to define risk at the level of the individual.

Though operative mortality is a reliable and clinically important outcome following PEA, those outcomes deemed meaningful to patients often extend beyond traditional measures such as survival [13]. Whilst there has been an increasing trend in assessing Patient-Reported Outcomes (PROs) in cardiothoracic surgery [14] and PEA [15], whether post-operative PROs can be predicted with any degree of accuracy prior to PEA has yet to be explored. Conceptually, the ability to produce valid and reliable predictor models of post-operative PROs has however been proven in other surgeries [16].

Given the aforementioned, our primary aim was to develop and validate a multivariate risk tool specific to PEA for the prediction of operative mortality and compare its predictive ability to the currently best available risk scoring system in cardiothoracic surgery, EuroSCORE II. As secondary aims, we harness our modelling approach and extensive dataset to inform, for the first time, factors which influence long-term outcomes following PEA, including those deemed important to the patient.

Methods

Patient selection

Consecutive individuals undergoing PEA for CTEPH at the United Kingdom National PEA unit (Royal Papworth Hospital) from August 2007 – December 2018 were included in a discovery cohort. Patient selection is summarised in Figure 1. CTEPH was confirmed as per current international diagnostic criteria at the time of invasively derived pulmonary haemodynamics (unless otherwise contraindicated) using multi-modal imaging [2, 17, 18]. Eligibility for PEA was assessed by an experienced multi-disciplinary team comprised of Pulmonary Hypertension (PH) physicians, CTEPH radiologists and PEA surgeons. PEA technique was as previously described [19]. All patients were reviewed by one of the 7 United

Kingdom adult specialist PH centres prior to PEA referral. Following surgery all patients were reviewed at 3 to 6 months post-operatively at Royal Papworth Hospital and for at least 5 years by referring specialist PH centre. This study was approved by Royal Papworth Hospital research governance committee (project reference S02560).

Predictor variables

Predictor variables were derived from prospectively entered data stored on local electronic clinical systems. An intentionally broad array of predictors were included for the purposes of hypothesis-free driven modelling, a comprehensive list of which are detailed in the Table S1. Predictor variables incorporated demographic information, relevant co-morbid conditions, current medications (cardiovascular medications, pulmonary vasodilators and anticoagulants) and genetic risk scores for common diseases (atrial fibrillation, type 2 diabetes, coronary artery disease [20]). Results of routine investigations from diagnostic baseline (within 6-months pre-PEA) and at 3 to 6-month post-PEA review were included, comprised of; echocardiographic and pulmonary function measures, right heart catheter derived pulmonary haemodynamics, six-minute walk distance (6MWD), blood tests performed as part of standard clinical care, and NYHA functional class (see Table S1 for full details). PROs were assessed at pre- and post-PEA time-points using the Cambridge Pulmonary Hypertension Outcome Review (CAMPHOR) which comprises of three negatively weighted scales measuring symptoms, activity levels and quality of life [21]. Symptom and Quality of Life scales are both scored out of 25 and Activities out of 30. Intra-operative variables including concomitant surgical procedures, cardiopulmonary arrest and deep hypothermic circulatory arrest time, peri-operative complications and length of hospital stay were also recorded.

For the purposes of risk modelling, pre-operative variables were divided into three predictor sets; NONINV, PREOP and EUROSCORE. Although not part of our risk modelling, two post-operative variable sets, DISCHARGE and ALL, were included in analysis for completion.

Pre-operative variable set NONINV included results of anthropometric assessment, co-comorbid conditions and non-invasive investigation results prior to PEA whilst variable set PREOP, included all variables in NONINV plus pre-operative invasive pulmonary haemodynamics. The third pre-operative predictor set EUROSCORE comprised the output of EuroSCORE II risk prediction assessment from our dataset variables. The post-operative variable dataset DISCHARGE included all PREOP variables plus intra-operative and post-operative variables available at the time of PEA hospital discharge whilst variable set ALL contained all variables in DISCHARGE plus non-invasive and invasive investigation results from first post-operative follow-up within one year of PEA.

Outcome measures

Three outcome measures were considered in model development: 90-day post-PEA mortality (90DM), 5-year mortality (5YM) and CAMPHOR total score change following surgery (PROchange). Individuals who died before follow-up were excluded from the analysis of PROchange. Death for the discovery cohort was censored at 30th June 2022.

Statistical analysis

A detailed description of the statistical approach is detailed in the Supplementary Materials. Analysis and risk modelling was performed in R version 4.1.2 [22]. Missingness for each variable was assessed with unknown values imputed as the mean of the predictor value across those for whom it was observed. Univariate associations between each variable and outcome

were performed using; t-tests (90DM), Cox Proportional Hazard models (5YM) and Pearson correlation (PROchange). Multivariate analysis considered three models (generalised linear, lasso regression and random forests) for each of the 15 predictor set/outcome combinations (NONIV/PREOP/DISCHARGE/ALL/EUROSCORE x 90DM/5YM/PROchange), using random survival forests [23] to predict survival times (outcome 5YM). The generalised linear model differed dependent on outcome assessed; logistic regression for 90DM (binary measure), Cox Proportional Hazards for 5YM (survival-time) and linear regression for outcome PROchange (integer measure). Models were compared according to Area Under the Receiver-Operator Characteristic (AUROC) curve for 90DM, concordance for 5YM and Spearman correlation for PROchange, averaged over cross-validation folds. Standard errors and 95% confidence intervals were estimated asymptotically for 90DM and empirically using bootstrap samples for 5YM and PROchange. For each ROC curve, we established the threshold which optimised the Youden index (that is, the predictor which maximised the sum of specificity and sensitivity) and reported the sensitivity, specificity and overall accuracy as this threshold.

Model validation

Internal validation of our predictive model for 90DM was performed using a separate prospective cohort of consecutive PEA patients at the same institution between 2019 and 2021 (n = 443). Validation cohort inclusion criteria was as for the discovery dataset (Figure 1). Sample size calculations estimated a minimum of 145 individuals to be required in the validation cohort to achieve 90% power to reject the null hypothesis that our predictive model for 90DM performed no better than randomly at 5% significance; derived from simulation using a pseudo-replication cohort of n individuals (with replacement) from the discovery cohort. Variables included in validation modelling were a subset of those ascertained from discovery risk models. Death date was censored at 19th July 2024.

Results

Discovery cohort characteristics

A total of 1334 eligible individuals underwent PEA during the study period. There were 93 deaths within 90 days of PEA, 46% of whom died before post-PEA hospital discharge. There were a further 103 deaths within 5-years of PEA. Cohort survival at 1, 3 and 5-years was 91.6%, 84.6%, 73.7% respectively. There were 966 individuals with paired pre- and post-PEA CAMPHOR scores. Of the 279 patients with pre-operative but not post-operative CAMPHOR scores, 84 individuals were confirmed to have died before first post-operative follow-up and 195 had no follow-up data for other reasons.

Discovery cohort characteristics are summarised in Table 1. Median ± IQR age at PEA was 61 ± 21 years and 54% were male. Prior to PEA, mean PAP was 45 ± 15 mmHg, PVR 674 ± 484 dynes and cardiac index 2.1 ± 0.8 l/min/m^2. 86% of individuals were in NYHA class III or IV at pre-operative baseline and 6MWD was 309 ± 206m. PEA resulted in significant improvements in haemodynamic, functional and PRO measures ($p \leq 0.01$, all; Table 1).

EuroSCORE predictive modelling

The EuroSCORE II inpatient mortality risk model was evaluated for its ability to predict outcomes 90DM, 5YM and PROchange using the predictor variable set EUROSCORE. Random forest models were the most accurate predictor models for all three outcome measures. For 90DM AUROC was 0.65 (95% CI 0.59, 0.71; Figure 2A). Although not validated for post-surgical long-term survival or PRO improvement, EUROSCORE concordance for 5YM (AUROC 0.67 [95% CI 0.63, 0.70]) and positive PROchange (AUROC 0.50 [95% CI 0.46,

0.54]) were similar or worse to that for 90DM. Sensitivity, specificity and overall accuracy of models are reported in Table S5.

Early post-operative mortality (90DM)

On univariate analysis the strongest associations between pre-operative variables and 90DM were for 6MWD ($p = 3.7 \times 10^{-8}$), age at PEA ($p = 5.1 \times 10^{-6}$), CAMPHOR Activity score ($p = 6.2 \times 10^{-4}$) and NYHA functional class ($p = 9.8 \times 10^{-4}$; Table S6). AUROC values were derived for each predictor set (NONINV, PREOP, DISCHARGE) against outcome 90DM using general linear regression, lasso regression and random forest models. The strongest predictor model was random forest for all three predictor sets (Tables S2 – S4). Random forest prediction accuracies for 90DM were higher for all three novel predictor sets (NONINV, PREOP, DISCHARGE compared to EUROSCORE ($p < 0.001$). When comparing random forest predictions of 90DM across variable sets, predictions were strongest for DISCHARGE (AUROC 0.87 [95% CI 0.83, 0.91]) although pre-operatively NONINV (AUROC 0.82 [95% CI 0.78, 0.86]) and PREOP (AUROC 0.81 [95% CI 0.76, 0.85]) still predicted 90DM with a good degree of accuracy (Table 2, Figure 2B). Pre-operatively, predictor set PREOP had the highest total accuracy at 86% (sensitivity 60%, specificity 88%; Table S5) for 90DM. Age at PEA, 6MWD, cardiac output, cardiac index, and PVR were pre-operative variables of highest importance from predictor set PREOP (Table 3).

Long-term mortality (5YM)

The strongest univariate associations between pre-operative variables and 5YM were; age at PEA ($p = 5.0 \times 10^{-14}$), 6MWD ($p = 1.9 \times 10^{-12}$), , CAMPHOR Activity score ($p = 5.5 \times 10^{-8}$) and NYHA class ($p = 4.6 \times 10^{-7}$; Table S6). Predictive accuracy for 5YM was greatest using random forest modelling although concordances were slightly lower than that achieved for

outcome 90DM (Tables 1, S2 and S3). The strongest predictor set for 5YM was ALL (concordance 0.85 (95% CI 0.84, 0.87) but variable sets derived from pre-operative variables were still able to predict 5YM with relative accuracy: NONIV: AUROC 0.75 (95% CI 0.73, 0.78), PREOP: AUROC 0.74 (95% CI 0.72, 0.77; Table 2, Figure 2C, Table S3, Table S5). Age at PEA and 6MWD were variables of high importance in 5YM random forests, as for 90DM, alongside cardiovascular risk factors such as a history of tobacco smoking or ischaemic heart disease and left atrial dilatation (Table 3).

Change in CAMPHOR score (PROchange)

The strongest univariate associations between individual pre-operative variables and PROchange (other than baseline CAMPHOR score) were right ventricular ejection fraction ($p = 3.1 \times 10^{-7}$), right ventricular dilatation ($p = 3.8 \times 10^{-5}$), tricuspid annular plane systolic excursion (TAPSE; $p = 1.4 \times 10^{-4}$) and PVR ($p = 2.4 \times 10^{-4}$; Table S6). PROchange was predicted moderately accurately from predictor sets (Table 2, Table S4). Random forest Spearman rank correlations between predicted and observed outcomes averaged over cross-validation folds were; NONINV 0.47 (95% CI 0.44, 0.50), PREOP 0.46 (95% CI 0.43, 0.49), DISCHARGE 0.48 (95% CI 0.46, 0.51) and ALL 0.61 (95% CI 0.59, 0.63; Table 2; Figure 2D, Table 2, Table S4, Table S5). By far the biggest contribution to outcome PROchange from pre-operative variables as assessed by variable importance was baseline CAMPHOR score (Table 3).

PEA risk model validation

The validation cohort comprised of a total of 443 prospective CTEPH PEA cases between 2019 and 2021. Cohort characteristics were similar to the discovery dataset (Table S7). There were 19 deaths within 90 days of PEA and 66 deaths by censoring date. Our derived risk model had

reasonable discriminatory ability for 90DM using the validation dataset 90DM (AUROC 0.71; [95% CI 0.57, 0.84]) and significantly different to random ($p = 0.004$). For 5YM, discrimination was moderate (AUROC 0.65; [95% CI 0.59, 0.71]). Predictors for both 90DM and 5YM were significantly better-than-random ($p < 0.005$ in both cases).

Implementation

Pre-determined variables of importance were incorporated into open-source risk tool for predicting early mortality and long-term survival. Inputting variables produces a Kaplan-Meier curve for post-PEA survival as depicted in Figure 3A enabling the visual representation of average cohort survival (black solid line) against that of the best/worse 20% of the cohort (black dotted line) and that specific to the patient of interest (red line).

The visual output from the prediction of change in PRO score is represented by the worked example in Figure 3B. Violin plots show the distribution of CAMPHOR at pre-operative baseline and 6-months post-PEA whilst grey lines show CAMPHOR score changes with surgery for the PEA population. The red line depicts the projected total CAMPHOR score change for the individual in question.

Discussion

The prognostic, haemodynamic and functional benefits of PEA in CTEPH are well-described [3 – 5], yet there is an unavoidable risk associated with PEA. Limited data exist as to factors affecting PEA outcome and the ability to ascribe risk at the level of the individual remains subjective, based on the clinical experience of the PEA surgeon and/or CTEPH MDT. In this largest evaluation of a PEA cohort (total n = 1509) to-date, we identify important and novel

risk factors for PEA morbidity and mortality using a random forest modelling approach and incorporate these variables into a prospectively validated open-source risk tool.

Although tools exist that can predict operative risk in cardiac surgery with a high degree of accuracy, none have been tested in PEA. We provide the first evaluation of the most widely-implemented cardiac surgery risk tool, EuroSCORE II in a PEA cohort. Whilst EuroSCORE II is validated for in-hospital mortality, 91% of our deaths within 90 days of PEA occurred before hospital discharge. EuroSCORE II predictions of post-PEA 90-day mortality were modest (90DM AUROC 0.65 [95% CI 0.59, 0.71]), suggesting additional factors outside of EuroSCORE variables may be implicated in PEA outcomes.

In our novel risk modelling, 90-day PEA mortality random forest predictions were more accurate than those from linear or linear penalised models and were significantly more accurate than predictions from EuroSCORE II ($p < 0.001$). A high degree of concordance was seen for pre-operative random forest predictor sets utilising both non-invasive (NONINV AUROC 0.82 (0.78, 0.86) and haemodynamic measures (PREOP AUROC 0.81 (0.76, 0.85)) in predicting 90-day mortality. Sensitivity and specificity was highest for the predictor set PREOP (total accuracy 86%).

Conventionally, higher pre-operative PVR has been considered the strongest predictor of post-PEA mortality. [9 - 12]. In our PREOP predictor set, haemodynamic measures of CTEPH severity also predominated, although a greater influence was seen from cardiac index than PVR (Table 3). The most important variables for 90-day mortality, however, for both NINV and PREOP variable sets were age and 6MWD. Historically, diffusion capacity for carbon monoxide (TLCO) has also been deemed an important predictor of CTEPH outcomes but did

not feature within our top 10 variables of importance for operative mortality. This may reflect selection bias in that those with low TLCO are generally excluded from surgery.

As a secondary aim, our modelling approach was also utilised to establish whether, and which, pre-operative factors may influence those PEA outcomes also deemed important to patients; long-term survival and health-related quality of life (PROs). For 5-year survival random forest models were more accurate than linear or linear penalised models and able to predict long-term mortality with excellent discriminative ability. Concordance was similar for models incorporating non-invasive variables (NONIV AUROC 0.81 (0.78 - 0.84)) vs. haemodynamic measures (PREOP AUROC 0.81 (0.77 - 0.82)). Unlike, 90-day mortality however, where measures of CTEPH severity directly influenced outcome, long-term survival following PEA was largely governed by non-CTEPH factors, driven by generic cardiovascular risk (Table 3). This reflects prior findings of Cannon et al 2016 [5] from our own cohort, where the most common causes of death outside of the immediate post-PEA period were non-CTEPH related, namely pneumonia and malignancy.

Left atrial area was an unexpected variable of importance in both early and late PEA mortality. Left atrial volume index has previously been shown to be an independent predictor of PEA mortality when cardiac magnetic resonance data is taken in isolation [24] but this is the first demonstration of its importance in combination with a broad array of clinical, functional and haemodynamic variables. Left atrial size is unlikely a surrogate of ventricular systolic dysfunction given that severe left ventricular dysfunction is a contraindication to PEA surgery and is not solely associated with atrial fibrillation. We hypothesise that left atrial dilatation may represent underlying left ventricular diastolic dysfunction, which has consistently been shown

to adversely affect outcomes in cardiac surgery [25] and is an independent predictor of survival post-PEA [26]. This however requires further study for confirmation.

We aim our mortality models to be used as a point-of-care tool in the clinical setting to better inform patients of their likely operative risk and inform patient choice. We acknowledge that our tool requires external validation for clinical implementation and actively encourage others centres in the validation of our tool. We provide an online implementation of our risk model at https://ajl-apps.shinyapps.io/pea_risk/

Our modelling was not able to predict health-related quality of life outcomes following PEA with accuracy. Newnham *et al* 2020 [15], have previously reported weak correlations between PRO score change and both haemodynamics and NYHA functional class following PEA in our own cohort. Crucially, our study reinforces the notion that observed improvements in objective measures of CTEPH, such as haemodynamics and functional class with PEA surgery, may not readily translate into those improvements deemed important to patients. Predicting patient-perceived improvement following PEA is therefore nuanced and it is important to counsel patients on this prior to deciding upon surgery.

Internal consistency and performance of our models was carefully assessed, with separate datasets for training and testing (through cross-validation). Furthermore, the best model type for each variable set/outcome pair was determined in a separate procedure to evaluate its performance, so our assertions should not be affected by regression to the mean. Our dataset has many missing values, and we expect that similar rates of missingness will be present in patients whose risks we wish to predict in clinic. Since the main aim of our current work is prediction of risk rather than accurately estimating effects of risk factors, we opt to use a simple

mean-value imputation method [27]. For new patients, missing values of a predictor should be replaced by the mean value of that predictor in our training dataset. Like all risk models our model will be subject to calibration drift as patient population and surgical expertise evolves. Although fairly complex, random forests have the advantage that predictor importance can be straightforwardly assessed, and that the overall model architecture resembles a voting majority of a mixture of experts, as is routine in general medical decision making. We chose not to consider further machine learning methods for prediction given our limited training data and capacity to optimise hyperparameters.

A critical mode in CTEPH diagnosis and workup for PEA is radiological imaging [2], which is not included in our risk models. The reason for this is partly pragmatic: there is no currently accepted grading for CTEPH, using any imaging modality, which would facilitate inclusion of imaging data as a predictor, and image-mining approaches are beyond the scope of this work. Since all patients in our database were, by definition, assessed as technically suitable and medically fit for PEA, we are not in a position to design a predictor tool for use in determining operability.

Conclusion

PEA mortality can be predicted pre-operatively to a potentially clinically useful degree and is driven by CTEPH factors in the early post-operative period, and non-CTEPH factors in the long-term. Our validated models enable individualised risk stratification at clinician point-of-care to better inform shared decision making between the clinical team and patient.

Code and data availability

Our full analysis pipeline is publicly available online at https://github.com/jamesliley/PEA_risk.

Acknowledgements:

With thanks to the National UK CTEPH Collaboration:

Gerry Coghlan MD (Pulmonary Hypertension service, Royal Free Hospital, London, UK), Luke Howard PhD (Pulmonary Hypertension service, Hammersmith Hospital, London, UK), Martin Johnson MD (Scottish Pulmonary Vascular Unit, Golden Jubilee National Hospital, Glasgow, UK), David Kiely MD (Sheffield PVDU, Royal Hallamshire Hospital, Sheffield, UK), James Lordan PhD (Pulmonary Vascular Unit, Freeman Hospital, Newcastle Upon Tyne, UK) and John S Wort PhD (Pulmonary Hypertension Service, Royal Brompton Hospital, London, UK).

Sources of funding: None to declare

Disclaimer

JL, KB, MN, JPZ, NS, NG have nothing to declare. JC has received support for attending meetings and/or travel from J&J Innovative Medicine. DJ has received payment for adjudication for the Maciteph clinical trial, has received payment for lectures, presentations and educational events from J&J Innovative Medicine and is a scientific board member of the International CTEPH Association and the World Symposium for Pulmonary Hypertension. DT has received payment or honoraria for lectures, presentations and educational events from J&J and support for attending meetings and/or travel from J&J Innovative Medicine, MSD and Ferrer. KS sits on the chronic thromboembolic pulmonary hypertension advisory board for Actelion/J&J Innovative Medicine. MT has received consulting fees and sits on the advisory board for Morphogen IX and J&J Innovative Medicine, has received support for attending meetings and/or travel from GSK and J&J Innovative Medicine and has participated on a data safety monitoring of advisory board for ComCov and FluCov clinical trials.

References

1. Kim NH, Delcroix M, Jais X, Madani MM, Matsubara H, Mayer E, Ogo T, Tapson VF, Gofrani H-A, Jenkins DP, et al. Chronic thromboembolic pulmonary hypertension. *Eur J Respir* 2019; 53:1801915
2. Humbert M, Kovacs G, Hoeper MM, Badagliacca R, Berger RMF, Brida M, et al. 2022 ESC/ERS Guidelines for the diagnosis and treatment of pulmonary hypertension. *ERJ* 2022; 2200879; DOI: 10.1183/13993003.00879-2022
3. Delcroix M, Lang I, Pepke-Zaba J, Jansa P, D'Armini AM, Snijder R, Bresser P, Torbicki A, Mellemkjaer S, Lewczuk J, et al. Long-Term Outcome of Patients With Chronic Thromboembolic Pulmonary Hypertension (CTEPH): Results From an International Prospective Registry. *Circulation* 2016;133:859-871.
4. Jenkins DP, Madani M, Fadel E, D'Armini AM, Mayer E. Pulmonary endarterectomy in the management of chronic thromboembolic pulmonary hypertension. *Eur Respir Rev* 2017. 26:160111.
5. Cannon JE, Su L, Kiely DG, Page K, Toshner M, Swietlik E, Treacy C, Ponnaberanam A, Condliffe R, Sheares K, et al. Dynamic Risk Stratification of Patient Long-Term Outcome After Pulmonary Endarterectomy. *Circulation* 2016;133:1761-1771
6. Nashef SA, Roques F, Sharples LD, Nilsson J, Smith C, Goldstone AR, Lockowandt U. EuroScore II. *Eur J Cardiothorac Surg* 2012; 41: 734 – 745
7. Biancari F, Vasques F, Mikkola R, Martin M, Lahtinen J, Heikkinen. Validation of EuroScore II in patients undergoing coronary artery bypass surgery. *Ann Thorac Surg* 2012; 93: 1930 – 1935.
8. Hogervorst EK, Rosseel PMJ, Van de Watering LMG, Brand A, Bentala M, Van der Meer BMJ, Van der Bom JG. Prospective validation of the EuroSCORE II risk model

427 in a single Dutch cardiac surgery centre. *Netherlands Heart Journal* 2018; 26: 540 –
428 551.
429 9. Corsico AG, D'Armini AM, Cerveri I, Klersy C, Ansaldo E, Niniano R, Gatto E,
430 Monterosso C, Morsolini M, Nicolardi S, et al. Long-term outcome after pulmonary
431 endarterectomy. *Am J Respir* 2008; 178: 419–424.
432 10. Saouti N, Morshuis WJ, Heijimen RH, Snijder RJ. Long-term outcome after
433 pulmonary endarterectomy for chronic thromboembolic pulmonary hypertension: a
434 single institution experience. *Eur J Cardiothorac Surg* 2009; 35: 947 – 952.
435 11. Ishida K, Masusa M, Tanabe N, Matsumiya G, Tatsumi K, Nakajima N. Long term
436 outcome after pulmonary endarterectomy for chronic thromboembolic pulmonary
437 hypertension. *J Thorac Caridiovasc* 2012; 144: 321 – 326.
438 12. Tromeur C, Jais X, Mercier O, Couturaud F, Montani D, Savale L, Jevnikar M,
439 Weatherald J, Sitbon O, Parent F, et al. Factors predicting outcome after pulmonary
440 endarterectomy. *PloS One* 2018; 21: e0198198.
441 13. Birkmeyer, JD., Dimick, JB and Birkmeyer, NJ. Measuring the quality of surgical
442 care: structure, process or outcomes? *J Am Coll Surg* 2004; 198: 626 – 632
443 14. Subramanian M, Kozower BD, Brown LM, Khullar OV, Fernandez FG. Patient-
444 reported outcomes in cardiothoracic surgery. *Ann Cardiothorac Surg* 2019; 107: 294
445 – 301
446 15. Newnham M, Bunclark K, Abraham N, Ali S, Amaral-Almeida L, Cannon JE,
447 Doughty N, Ng C, Ponnaberanam A, Sheares K et al. 2020. CAMPHOR score:
448 patient-reported outcomes are improved by pulmonary endarterecomy in chronic
449 thromboembolic pulmonary hypertension. *Eur Respir J* 2020; 56:1902096
450 16. Dietz N, Sharma M, Alhourani A, Ugiliweneza B, Wang D, Nuno MA, Drazin D,
451 Boakye. Variability in the utility of predictive models in predicting patient-reported

outcomes following spine surgery for degenerative conditions: a systematic review. Neurosurg Focus 2018; 45 (5): E10

17. Task Force for Diagnosis and Treatment of Pulmonary Hypertension of European Society of Cardiology (ESC); European Respiratory Society (ERS); International Society of Heart and Lung Transplantation (ISHLT); Galie N, Hoeper MM, Humbert M, Torbicki A, Vachiery J-L, Barbera JA, Berghetti M, Corris P, Gaine S, Gibbs JS et al. Guidelines for the diagnosis and treatment of pulmonary hypertension. *Eur Respir J* 2009; 34 (6): 1219 – 1263
18. 2015 ESC/ERS Guidelines for the diagnosis and treatment of pulmonary hypertension. The joint task force for diagnosis and treatment of pulmonary hypertension of the European Society of Cardiology (ESC) and the European Resouratory Society (ERS): Endorsed by: Association for European Paediatric and Congential Cardiology (AEPC), International Society for Heart and Lung Transplantation (ISHLT). Galie N, Humbert M, Vachiery J-L, Gibbs S, Lang I, Torbicki A, Simonneau G, Peacock A, Vonk Noordegraaf A, Beghetti M, et al. *Eur Heart J* 2016; 37: 67-119.
19. Jenkins, D. Pulmonary endarterectomy: the potentially curative treatment for patients with chronic thromboembolic pulmonary hypertension. *European Respiratory Review* 2015; 24: 263-271.
20. Khera AV, Chaffin M, Aragam KG, Haas ME, Roselli C, Choi SH, Natarajan P, Lander ES, Lubitz SA, et al. Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. *Nature genetics* 2018; 50: 1219.
21. McKenna SP, Doughty N, Meads DM, Doward LC, Pepke-Zaba-J. The Cambridge Pulmonary Hypertension Outcome Review (CAMPHOR): A measure of health-

related quality of life and quality of life for patients with pulmonary hypertension. *Qual Life Res* 2006; 15: 103–15.

22. R Core Team. R: A language and environment for statistical computing. R Found Stat Comput Vienna, Austria 2019: https://www.r-project.org/
23. Ishwaran H, Kogalur UB, Blackstone EH and Lauer MS. Random survival forests. *Annals of Applied Statistics* 2008: 841- 860.
24. Shanin Y, Alabed S, Quadery SR, Lewis RA, Johns C, Alkhanfar D, Sukhanenko M, Alandejani F, Garg P, Elliot CA, et al. CMR measures of left atrial volume index and right ventricular function have prognostic value in chronic thromboembolic pulmonary hypertension. *Frontiers in Medicine* 2022: https://doi.org/10.3389/fmed.2022.840196
25. Ludden T, Alberts TAM, Breel JS, de Klerk ES, Javaid SK, Boekholdt SM, et al. Exploring the impact of left ventricular diastolic dysfunction on postoperative cardiac surgery outcomes, with a focus on sex disparities: a comprehensive literature review. *Front. Anesthesiol* 2023; 2:1280189. doi: 10.3389/fanes.2023.1280189
26. Gerges C, Pistritto A-M, Gerges M, Friewalk R, Hartig V, Hofbauer TM, et al. Left ventricular filling pressure in chronic thromboembolic pulmonary hypertension J Am Coll Cardiol 2023; 81 (7): 653 – 664.
27. James G, Witten D, Hastie T, Tibshirani R. An introduction to statistical learning. Springer 2013; 11.

Tables and Figures

Table 1: Patient demographics and characteristics pre- and post-pulmonary endarterectomy

	Pre-PEA		Post-PEA		
	n	**Value**	**n**	**Value**	**p-value**
Total n	1334		1241		
Age at PEA, yrs	1333	61 ± 21			
Male sex, n (%)	1333	715 (54)			
BMI, kg/m^2	953	29 ± 8			
FEV$_1$/FVC, %	540	72 ± 13			
Smoker*, n (%)	953	495 (52)			
Comorbidities, n (%)					
Atrial arrhythmia	1286	120 (9)			
Systemic hypertension	1286	342 (27)			
Diabetes Mellitis	1286	142 (11)			
Chronic renal disease	1122	49 (4)			
Ischaemic heart disease†	1309	159 (12)			
History of malignancy	1283	123 (10)			
Thrombophilia	1334	90 (7)			
Thyroid dysfunction	1286	116 (9)			
Pulmonary vasodilator, n (%)	1119	314 (28)			
Haemodynamics					
Mean PAP, mmHg	1266	45 ± 15	1068	25 ± 13	<0.001
PVR, dynes cm s^{-5}	1208	674 ± 484	1041	244 ± 210	<0.001
PAWP, mmHg	1055	11 ± 5	1042	10 ± 5	0.0087
CI, l/min/m^2	1138	2.1 ± 0.8	1025	2.3 ± 0.7	<0.001
Functional status					

NYHA, 1/2/3/4 %	1220	0/14/75/11	998	29/44/26/1	<0.001
6MWD‡, metres	802	309 ± 206	993	365 ± 163	<0.001
CAMPHOR					
Symptoms	1245	12 ± 11	990	4 ± 9	<0.001
Activity	1245	11 ± 10	990	6 ± 10	<0.001
Quality of Life	1245	10 ± 12	990	4 ± 11	<0.001
Intra-operative					
CPB time, mins	986	321 ± 67			
DHCA time, mins	843	37 ± 15			
Other surgery, n (%)					
CABG	1332	99 (7)			
AVR	1332	15 (1)			
MVR	1332	12 (1)			
ASD/PFO closure	1331	5 (0)			
Complications, n (%)					
CPAP	1193	305 (26)			
Haemofiltration	1127	65 (6)			
ECMO	1265	75 (6)			
Pneumonia	1187	175 (15)			
Return to theatre	1191	85 (7)			
Reperfusion injury	872	66 (8)			
Intubation, days	1231	3087 (251)			
ICU stay, days	955	4 ± 3			
Total inpatient stay, days	985	13 ± 9			
Inpatient death, n (%)	1094	43 (4)			

Values are expressed as median ± IQR. Percentages may not add to 100 due to rounding. Variables taken at time of diagnostic right heart catheterisation.

PEA, pulmonary endarterectomy; BMI, body mass index; PAP, pulmonary artery pressure; PAWP, pulmonary arterial wedge pressure; PVR, pulmonary vascular resistance; CI, cardiac index; NYHA,

New York Heart Association functional class; 6MWD, 6-minute walk distance; CPB, cardiopulmonary bypass; DHCA, deep hypothermic circulatory arrest.

* At any time point prior to PEA

† History of myocardial infarction, coronary artery stenting, coronary artery bypass grafting or coronary artery lesion/s requiring intervention following routine pre-PEA angiography in those > 40 years

‡ Patients referred from Sheffield Pulmonary Vascular Diseases Unit were excluded as incremental shuttle walk test and not 6MWD performed pre-PEA

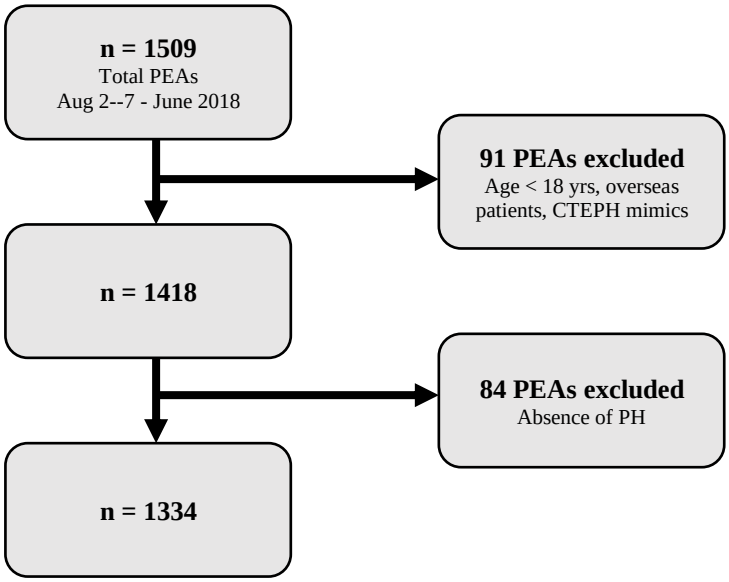
Table 2: Random forest AUROC (95% Confidence Intervals) for predictor sets against outcomes: 90-day mortality (90DM), 5-year mortality (5YM) change in CAMPHOR score (PROchange)

	90DM	**5YM**	**PROchange**
EUROSCORE	0.65 (0.59 - 0.71)	0.67 (0.63 - 0.70)	0.50 (0.46 - 0.54)
NONINV	0.82 (0.78 - 0.86)	0.75 (0.73 - 0.78)	0.47 (0.44 - 0.50)
PREOP	0.81 (0.76 - 0.85)	0.74 (0.72 - 0.77)	0.46 (0.43 - 0.49)
DISCHARGE	0.87 (0.83 - 0.91)	0.78 (0.76 - 0.80)	0.48 (0.46 - 0.51)
ALL	NA	0.85 (0.84 - 0.87)	0.61 (0.59 - 0.63)

Table 3: Top 5 variables of importance on random forest modelling from NONINV and PREOP predictor sets for each outcome measure

	90DM	5YM	PROchange
NONIV	Age	Age	CAMPHOR Symptoms
	6MWD	Left atrial dilatation	CAMPHOR QoL
	CAMPHOR QoL	Current or ex-smoker	CAMPHOR Activity
	CAMPHOR Activity	6MWD	Age
	CAMPHOR Symptoms	PMHx ischaemic heart disease	Body Mass Index
PREOP	Age	Age	CAMPHOR Symptoms
	6MWD	Left atrial dilatation	CAMPHOR QoL
	Cardiac output	Current or ex-smoker	CAMPHOR Activity
	Cardiac index	6MWD	PVR
	PVR	PMHx ischaemic heart disease	Cardiac output

Figure 1: Consort flowchart for inclusion in PEA dataset



Overseas patients excluded due to paucity of longitudinal post-operative UK follow-up.

PEA, Pulmonary Endarterectomy; CTEPH, chronic thromboembolic pulmonary hypertension; PH, pulmonary hypertension.
CTEPH mimics include pulmonary sarcoma and vasculitis

Figure 2: ROC curves for: **A** – EUROSCORE predictor set against outcome 90DM; **B** – novel predictor sets against outcome 90DM; **C** – novel predictor sets against outcome 5YM; **D** – novel predictor sets against outcome PROchange.

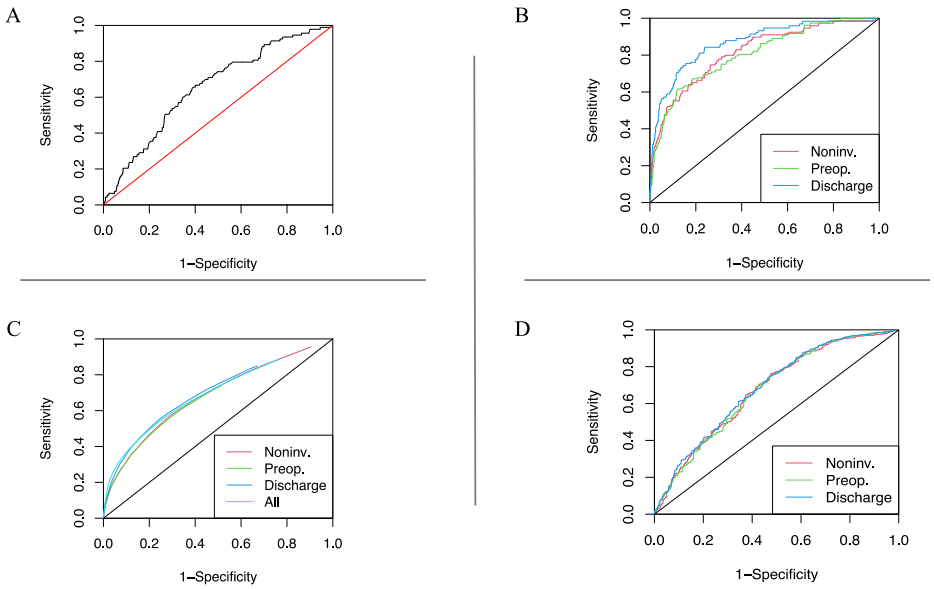


Figure: 3A Example of Kaplan-Meier survival curve output from derived online PEA risk prediction tool. **3B** Example of output for prediction of CAMPHOR score change following PEA from online risk prediction tool.