

PROTOCOL AND STATISTICAL ANALYSIS PLAN

Individualised Oxygen Targets after Cardiac Arrest: A secondary analysis of the LOGICAL Trial

Date: 7 July 2026

Parent Trial Registration: Australian and New Zealand Clinical Trials Registry (ACTRN12621000518864)

Protocol Version: 1.1

Authors: Nicholas J. Johnson, Andrew Udy, Bradley A. Drayton, Alexandra B. Spicer, Matthew M. Churpek, Carol L. Hodgson, Michael Bailey, Paul J Young

Corresponding Author: Nicholas J. Johnson, nickj45@uw.edu

1. BACKGROUND AND RATIONALE

1.1 Oxygen targets after Cardiac Arrest

Cardiac arrest affects millions of people worldwide each year and is associated with high mortality and a substantial burden of permanent neurological disability among survivors.^{1,2} Following return of spontaneous circulation, brain injury results from both the ischaemia sustained during the arrest and the reperfusion injury that occurs when circulation is restored.^{3,4} Oxygen plays a central role in both processes: insufficient oxygen delivery during and after arrest causes brain hypoxia and neuronal death, but excess oxygen in the reperfusion period may overwhelm endogenous antioxidant defences, generating free radicals and accelerating neuronal injury.^{5,6} This has led to interest in oxygen strategies that limit exposure to hyperoxaemia while avoiding hypoxaemia. Despite this biological rationale and numerous observational studies demonstrating associations between hyperoxemia and unfavorable outcomes, multiple randomised trials have failed to demonstrate that conservative oxygen therapy improves outcomes compared with liberal oxygen therapy after cardiac arrest.^{5,7–}

14

1.2 The LOGICAL Trial

The Low OxyGen Intervention for Cardiac Arrest injury Limitation (LOGICAL) trial was a multicentre, parallel-group, open-label, assessor-blinded randomised clinical trial conducted in 53 ICUs in Australia, New Zealand, and Ireland.¹² Eligible participants were mechanically ventilated adults in the ICU after cardiac arrest with suspected hypoxic–ischaemic encephalopathy (HIE).¹⁵ Patients were randomised to conservative or liberal oxygen therapy. In the conservative-oxygen group, the fraction of inspired oxygen ($F_{I}O_2$) was decreased to 0.21 where possible and an upper oxygen saturation via pulse oximetry (S_pO_2) alarm limit of 95% was applied whenever supplemental oxygen was administered. In the liberal-oxygen group, no upper S_pO_2 limit was applied and a minimum $F_{I}O_2$ of 0.30 was required during invasive mechanical ventilation.

The primary outcome was survival with a favourable functional outcome (Glasgow Outcomes Scale-Extended, GOS-E ≥ 5) at 180 days. Among 1,821 patients in the intention-to-treat population, a favourable outcome was observed in 313 of 819 patients (38.2%) in the conservative-oxygen group and 353 of 890 patients (39.7%) in the liberal-oxygen group (adjusted relative risk 0.97; 95% CI 0.87–1.09; $P=0.65$). Day 180 mortality and other secondary outcomes were similar by treatment group.

1.3 Rationale for Individualised Treatment Effect Analysis

The absence of an average treatment effect does not preclude clinically meaningful variation in the effects of treatment across patient subgroups. Several lines of evidence suggest that such heterogeneity of treatment effect may exist for conservative vs. liberal oxygen in post-cardiac arrest patients.

The pathophysiology of post-cardiac arrest brain injury is heterogeneous. Patients differ substantially in arrest duration, aetiology, initial rhythm, degree of haemodynamic compromise, and comorbidity burden, all of which plausibly influence the balance between the risks of hyperoxia-induced reperfusion injury and hypoxia-induced secondary brain injury.

Prior data support the hypothesis that heterogeneous responses to oxygen therapy might exist. An individualised treatment effect (ITE) analysis of two prior randomised trials of conservative versus liberal oxygen therapy in heterogeneous ICU patient populations derived an ITE model in the PILOT trial and applied it to ICU-ROX patients, predicted individualised treatment effects ranging from a 27.2% absolute reduction to a 34.4% absolute increase in 28-day mortality across individual ICU-ROX patients, despite both trials reporting no average treatment effect.¹⁶

The LOGICAL dataset is well-suited to ITE analysis. With 1,821 patients, a broad range of prospectively collected baseline characteristics, a clinically meaningful primary outcome assessed at 180 days, and follow-up

rates exceeding 93%, LOGICAL is the largest and most comprehensive trial of post-cardiac arrest oxygen targets conducted to date.

2. OBJECTIVES AND RESEARCH QUESTIONS

To test the hypothesis that the effect of peripheral oxygenation-saturation (SpO₂) targets on mortality and favourable functional outcome differ based on individual patient characteristics in a cohort of ventilated ICU adults who were unresponsive after cardiac arrest.

3. STUDY POPULATION AND DATA SOURCE

3.1 Population

All analyses will use the LOGICAL intention-to-treat population (n=1,821), defined as all randomised patients except those who did not consent or withdrew consent for use of any data. The primary outcome was available for 1,709 patients (93.8%). Multiple imputation will be applied for missing primary outcome data as described in the parent trial statistical analysis plan.

3.2 Pre-registration

This protocol and statistical analysis plan was finalised and posted in the public domain prior to any access to treatment-unblinded data by the team who conducted the analyses.

3.3 Relationship to Parent Trial

This analysis uses data from the LOGICAL trial and extends beyond the analyses reported in the primary manuscript. All analyses are additional to those in the parent trial and do not alter or supersede any parent trial findings.

4. BASELINE VARIABLES

The following baseline variables were collected prospectively in the LOGICAL trial and are available for analysis.

Demographic: age (continuous); sex (binary); ethnicity (categorical); skin tone (6-point ordinal scale)

Comorbidities (all binary): hypertension; diabetes; ischaemic heart disease; asthma or COPD; previous myocardial infarction; previous cardiac arrhythmia; previous percutaneous coronary intervention; congestive heart failure; previous CABG; previous TIA or stroke

Cardiac arrest characteristics: arrest location (categorical); out-of-hospital vs. in-hospital arrest (binary, derived); witnessed arrest (binary); bystander CPR (binary, available for out-of-hospital arrests only); first monitored rhythm (categorical); shockable vs. non-shockable rhythm (binary, derived); response time (continuous, minutes); time to defibrillation (continuous, minutes); time to sustained ROSC (continuous, minutes); cause of arrest (categorical); medical vs. other cause (binary, derived); STEMI present (binary)

Physiology at randomisation: PaO₂ (continuous, mmHg); FiO₂ (continuous); time from ROSC to randomisation (continuous, hours); time from ICU admission to randomisation (continuous, hours)

Co-interventions and illness severity: drugs given during resuscitation (adrenaline, amiodarone, vasopressin: binary each); targeted temperature management (binary); sepsis at baseline (binary); non-HIE acute brain injury at baseline (binary).

5. OUTCOMES

The primary outcome for this analysis is survival with a favourable functional outcome at 180 days based on the Glasgow Outcome Scale – Extended (GOSE-E). Levels on the GOS-E range from 1 (death) to 8 (“upper

good recovery” [no injury-related problems]). Categories were dichotomized into “death or survival with an unfavorable functional outcome” (which included “death,” “vegetative state,” “lower severe disability,” and “upper severe disability” [GOS-E levels of 1 to 4]) and “survival with a favorable functional outcome” (which included “lower moderate disability,” “upper moderate disability,” “lower good recovery,” and “upper good recovery” [GOS-E levels of 5 to 8]). Survival with a favourable functional outcome (GOS-E $\geq$ 5) was the primary outcome of the parent LOGICAL trial.

Secondary outcomes are all-cause mortality at 180 days and GOS-E modeled continuously.

6. ANALYTICAL APPROACH

Two complementary analytical approaches will be used:

1. Risk modelling
2. Effect modelling using machine learning

Skin tone will be examined as a predictor variable within both models and as a standalone subgroup analysis using a formal interaction test. All analyses are applied to the primary outcome. The skin tone subgroup analysis will also be applied to mortality at 180 days.

7. APPROACH 1: RISK MODELLING

7.1 Rationale

Patients' baseline risk of the primary outcome is a key determinant of absolute treatment effects. Even if the relative treatment effect is constant, patients at higher baseline risk experience larger absolute differences between treatments. Stratifying treatment effects by predicted risk can identify whether absolute benefits concentrate in particular risk groups and can reveal effect modification if relative effects also vary across risk strata.

7.2 Risk Model Development

A multivariable prognostic model will be developed to estimate each patient's predicted probability of an unfavourable outcome at 180 days using pre-randomisation baseline characteristics. The model will be developed using the full intention-to-treat population with both treatment groups combined.

Candidate predictor variables are: age; sex; out-of-hospital vs. in-hospital arrest; shockable vs. non-shockable rhythm; time to ROSC; medical vs. other cause; STEMI; witnessed arrest; bystander CPR (out-of-hospital arrests only); FiO₂ at randomisation; PaO₂ at randomisation.

Variable selection will use penalised logistic regression (LASSO) with 10-fold cross-validation to select the optimal penalty parameter, implemented using the glmnet package in R. Model calibration and discrimination will be reported (calibration intercept, slope, and unreliability index U p-value).

7.3 Evaluation of Treatment Effect Across Risk Strata

Patients will be divided into tertiles of predicted baseline risk. Within each tertile, the absolute risk difference and relative risk for the primary outcome will be estimated from log-binomial regression models adjusted for site (random effect), non-HIE brain injury at baseline, and sepsis at baseline.

To test whether the relative treatment effect varies across levels of predicted risk, a log-binomial model will be fitted with treatment group, the linear predictor from the risk model as a continuous variable, and their interaction. The P-value for this interaction will be reported.

8. APPROACH 2: EFFECT MODELLING

8.1 Overview

The effect modelling approach estimates the individualised treatment effect (ITE) for each patient: the predicted absolute difference in probability of an unfavourable outcome between assignment to conservative versus liberal oxygen therapy, based on that patient's baseline characteristics. Patients for whom the model predicts a lower probability of an unfavourable outcome with conservative oxygen have a negative ITE; patients predicted to do better with liberal oxygen have a positive ITE.

8.2 Test–Training Split

Patients will be randomly allocated to a training cohort (2/3, n=1,214) and a test cohort (1/3, n=607). Allocation will be stratified by treatment group, arrest location (out-of-hospital vs. in-hospital), rhythm (shockable vs. non-shockable), and arrest cause (medical vs. other) to ensure balance on key variables.

All model development, algorithm selection, and hyperparameter tuning will occur exclusively in the training cohort. Model performance will be assessed in the test cohort.

We acknowledge that internal validation is less stringent than external validation across independent trials. The extent to which findings generalise to other post-cardiac arrest populations will require future work combining LOGICAL data with data from other trials (or externally validating LOGICAL-derived models in other trial datasets).

9. PREDICTOR VARIABLES FOR EFFECT MODEL

All baseline variables will be used as candidate predictors. Variables with more than 40% missing data will be excluded from the primary model.

Missing predictor variables will be imputed using bagged trees (preProcess function, caret package in R) derived from the derivation cohort. The same imputation model will be applied to the test cohort.

10. MODEL DEVELOPMENT AND SELECTION

10.1 Algorithms

The following machine learning algorithms will be evaluated in the derivation cohort using 5-fold cross-validation, assessed using out-of-sample discrimination metrics (e.g. adjusted Qini statistic, Qini statistic, pROCini):

1. Rboost — XGBoost implementation of the R-learner
2. Causal Forest
3. Uplift Random Forest
4. Tian transformation with elastic net
5. X-learner with Bayesian Additive Regression Trees
6. Rlasso — lasso implementation of the R-learner
7. T-cforest — conditional random forest implementation of T-learner
8. S-cforest— conditional random forest implementation of S-learner
9. X-cforest— conditional random forest implementation of X-learner

Within each 5-fold cross-validation within the derivation cohort, hyperparameters will be tuned using an inner loop of 5-fold cross-validation. For each algorithm, five model iterations will be constructed using different random seeds to improve stability, and the mean ITE across iterations will be used for each patient.

10.2 Algorithm Selection

The single best-performing algorithm based on the out-of-sample discrimination statistics from 5-fold cross-validation in the derivation cohort will be selected. This selection will occur prior to any assessment in the test cohort.

10.3 Hyperparameter Tuning and Final Model

The final model will be trained using the same process as above (hyperparameter tuning using 5-fold cross-validation and averaging predictions across random seeds for stability) but now on the full training cohort and then applied to the test cohort.

11. MODEL VALIDATION AND PERFORMANCE ASSESSMENT

11.1 Predicted ITEs

The final model will generate a predicted ITE for each patient in the validation cohort. A negative ITE indicates predicted benefit from conservative oxygen therapy; a positive ITE indicates predicted benefit from liberal oxygen therapy. The distribution of predicted ITEs will be described and presented as a histogram.

11.2 Tertiles of Predicted ITE

Patients in the test cohort will be ranked by predicted ITE and divided into thirds using cut-points determined in the test cohort:

- Lower third: predicted benefit from conservative oxygen
- Middle third: predicted similar outcomes with either treatment
- Upper third: predicted benefit from liberal oxygen

11.3 Characterisation of Tertiles

Baseline characteristics of patients in each tertile will be summarised and compared using parametric and non-parametric tests, as appropriate, for continuous variables and chi-squared tests for categorical variables.

11.4 Variable Importance

Variable importance will be quantified as the mean absolute change in predicted ITE caused by replacing each variable with its median value from the derivation cohort. The normalised importance of each predictor will be displayed in a variable importance plot.

11.5 Partial Dependence Plots

For the four most important continuous predictor variables, partial dependence plots will show the effect of varying each variable on the average marginal predicted ITE, with bootstrapped 95% confidence intervals.

11.6 Spaghetti Plot

For 100 patients selected from the test cohort (one per percentile of predicted ITE), a spaghetti plot will show how the predicted ITE changes with the sequential addition of the 10 most important baseline variables, beginning at the average treatment effect for the test cohort and ending at each patient's final predicted ITE.

12. TESTING FOR INDIVIDUALIZED TREATMENT EFFECT

12.1 Primary Test

A logistic regression model will be fitted in the test cohort with unfavourable outcome at 180 days as the dependent variable, including treatment group, ITE tertile (categorical), and the treatment × ITE tertile interaction term. A likelihood ratio test will evaluate the significance of the interaction term (two-sided $\alpha = 0.05$).

12.2 Observed Treatment Effects Within Tertiles

For each ITE tertile in the test cohort, the following will be reported: number of patients by treatment group; primary outcome rate within each treatment group; absolute risk difference with 95% CI; and relative risk with 95% CI.

12.3 Estimated Benefit of Model-Directed Treatment

The absolute reduction in unfavourable outcomes that would have been achieved if each patient had received the treatment predicted to be best for them will be estimated as the difference between the observed outcome rate under random assignment and the counterfactual outcome rate under model-directed assignment.

12.4 Model Discrimination

Discrimination will be quantified using the adjusted Qini value, Qini value and the C-for-benefit statistic with bootstrapped 95% CI. The Qini curve will be presented graphically.

12.5 Model Calibration

Calibration will be assessed by comparing the observed absolute risk difference within each ITE tertile with the mean predicted ITE within the corresponding tertile, presented as a calibration plot.

13. SKIN TONE SUBGROUP ANALYSIS

A log-binomial regression model will be fitted including treatment group, skin tone (dichotomised as lighter [1–3] vs. darker [4–6]), the treatment × skin tone interaction term, presence of non-HIE acute brain injuries and conditions at baseline, sepsis at baseline, and site as a random effect. The treatment effect within each skin tone group will be reported as relative risk and absolute risk difference with 95% confidence intervals. The P-value for the interaction term will be reported. This analysis will be applied to both the primary outcome and mortality at 180 days. Skin tone will also be analysed as a categorical variable (6-point scale) with treatment as an interaction term.

The proportion of patients with missing skin tone data will be reported. If more than 10% of patients have missing skin tone, a sensitivity analysis restricted to patients with observed skin tone will be conducted.

14. SENSITIVITY ANALYSES

1. Exclusion of predicted baseline risk. The effect model will be re-fitted without the predicted baseline risk variable to assess its contribution to model performance.
2. Alternative training-test split. The full analysis will be repeated using a different random seed (to be specified prior to data access) to assess the stability of findings to the particular random split chosen.
3. Alternative split ratio. The analysis will be repeated using a 50/50 training-test split.
4. Complete-case analysis. The primary analysis uses multiple imputed outcome data. A sensitivity analysis restricted to patients with an observed primary outcome will be conducted.
5. Alternative skin tone threshold. The skin tone subgroup analysis will be repeated using a threshold of ≥ 4 vs. ≤ 3 in addition to the primary dichotomisation.

15. MISSING DATA

The primary outcome will be imputed using the multiple imputation approach described in the parent trial statistical analysis plan: chained equations with predictive mean matching using 5 nearest neighbours, 20 imputed datasets, separately by treatment group. All primary analyses will be conducted on multiply imputed datasets with results combined using Rubin's rules.

Missing predictor variables will be imputed using bagged trees derived from the derivation cohort and applied to both cohorts.

16. STATISTICAL SOFTWARE

All analyses will be performed in R version 4.4.1 or later using the following packages: rlearner, grf, uplift, xgboost, glmnet, caret, ggplot2, tidycmprsk, and ggsurvfit. Analysis code will be made publicly available upon publication.

A two-sided alpha of 0.05 applies to the primary interaction test only. All other P-values and confidence intervals are presented for exploratory purposes and should not be used as the basis for hypothesis testing. No correction for multiple comparisons is planned.

17. ETHICS AND GOVERNANCE

This analysis uses de-identified data from the LOGICAL trial under the data sharing provisions of the parent trial. No additional ethics approval is required. A data access agreement will be signed by all analysts. This protocol will be posted publicly prior to any treatment-unblinded analysis.

18. ROLES AND RESPONSIBILITIES

Name	Role
Nicholas Johnson	First author; protocol and manuscript writing
Andrew Udy	Clinical content review
Brad Drayton	Primary analyst
Alexandra Spicer	Machine learning methods
Matthew Churpek	Machine learning methods supervision
Carol Hodgson	Co-chief investigator LOGICAL
Michael Bailey	Data preparation; analysis oversight
Paul Young	Senior author; Co-chief investigator LOGICAL

19. REFERENCES

1. 2024 Heart Disease and Stroke Statistics: A Report of US and Global Data From the American Heart Association [Internet]. [cited 2024 Jan 24];Available from: <https://www.ahajournals.org/doi/epdf/10.1161/CIR.0000000000001209>
2. Coute RA, Nathanson BH, Panchal AR, et al. Disability-Adjusted Life Years Following Adult Out-of-Hospital Cardiac Arrest in the United States: A Burden of Disease Estimate. Circulation: Cardiovascular Quality and Outcomes [Internet] 2019;12(3). Available from: <https://www.ahajournals.org/doi/10.1161/CIRCOUTCOMES.118.004677>
3. Sekhon MS, Ainslie PN, Griesdale DE. Clinical pathophysiology of hypoxic ischemic brain injury after cardiac arrest: a “two-hit” model. Crit Care 2017;21:90.

4. Sandroni C, Cronberg T, Sekhon M. Brain injury after cardiac arrest: pathophysiology, treatment, and prognosis. *Intensive Care Med* 2021;47(12):1393–414.
5. Johnson NJ, Carlborn DJ, Gaieski DF. Ventilator Management and Respiratory Care After Cardiac Arrest: Oxygenation, Ventilation, Infection, and Injury. *Chest* 2018;153(6):1466–77.
6. Bray J, Skrifvars MarkusB, Bernard S. Oxygen targets after cardiac arrest: A narrative review. *Resuscitation* 2023;189:109899.
7. Bernard SA, Bray JE, Smith K, et al. Effect of Lower vs Higher Oxygen Saturation Targets on Survival to Hospital Discharge Among Patients Resuscitated After Out-of-Hospital Cardiac Arrest: The EXACT Randomized Clinical Trial. *JAMA* [Internet] 2022 [cited 2022 Nov 6];Available from: <https://doi.org/10.1001/jama.2022.17701>
8. Holmberg MJ, Nicholson T, Nolan JP, et al. Oxygenation and ventilation targets after cardiac arrest: A systematic review and meta-analysis. *Resuscitation* 2020;152:107–15.
9. Robba C, Badenes R, Battaglini D, et al. Oxygen targets and 6-month outcome after out of hospital cardiac arrest: a pre-planned sub-analysis of the targeted hypothermia versus targeted normothermia after Out-of-Hospital Cardiac Arrest (TTM2) trial. *Crit Care* 2022;26(1):323.
10. Schmidt H, Kjaergaard J, Hassager C, et al. Oxygen Targets in Comatose Survivors of Cardiac Arrest. *N Engl J Med* 2022;387(16):1467–76.
11. Young PJ, Hodgson CL, Rasmussen BS. Oxygen targets. *Intensive Care Medicine* [Internet] 2022 [cited 2022 May 6];Available from: <https://link.springer.com/epdf/10.1007/s00134-022-06714-0>
12. The LOGICAL Investigators and the Australian and New Zealand Intensive Care Society Clinical Trials Group. Conservative Oxygen for Unresponsive Patients after Cardiac Arrest. *N Engl J Med* 2026;NEJMoa2513814.
13. Johnson NJ, Dodampahala K, Rosselot B, et al. The Association between Arterial Oxygen Tension and Neurological Outcome after Cardiac Arrest. *Therapeutic Hypothermia and Temperature Management* 2017;7(1).
14. Wang HE, Prince DK, Drennan IR, et al. Post-resuscitation arterial oxygen and carbon dioxide and outcomes after out-of-hospital cardiac arrest. *Resuscitation* 2017;120.
15. Young PJ, Hodgson CL, Mackle D, et al. Protocol summary and statistical analysis plan for the low oxygen intervention for cardiac arrest injury limitation (LOGICAL) trial. *Crit Care Resusc* 2023;25(3):140–6.
16. Buell KG, Spicer AB, Casey JD, et al. Individualized Treatment Effects of Oxygen Targets in Mechanically Ventilated Critically Ill Adults. *JAMA* 2024;331(14):1195–204.
17. Sjoding MW, Dickson RP, Iwashyna TJ, Gay SE, Valley TS. Racial Bias in Pulse Oximetry Measurement. *New England Journal of Medicine* 2020;383(25):2477–8.
18. Valbuena VSM, Seelye S, Sjoding MW, et al. Racial bias and reproducibility in pulse oximetry among medical and surgical inpatients in general care in the Veterans Health Administration 2013-19: multicenter, retrospective cohort study. *BMJ* 2022;378:e069775.

20. APPENDIX: PLANNED TABLES AND FIGURES

Table 1: Baseline Characteristics of Training and Test Cohorts

Characteristic	Training cohort (n≈1,214)	Test cohort (n≈607)	P value
Age — yr, mean (SD)			
Male sex — no. (%)			
Out-of-hospital cardiac arrest — no. (%)			
Shockable first rhythm — no. (%)			
Medical cause — no. (%)			
Time to ROSC — min, median (IQR)			
STEMI — no. (%)			
Witnessed arrest — no. (%)			
Bystander CPR — no. (%)			
Targeted temperature management — no. (%)			
Sepsis at baseline — no. (%)			
Non-HIE brain injury at baseline — no. (%)			
PaO2 at randomisation — mmHg, median (IQR)			
FIO2 at randomisation, mean (SD)			
Time from ROSC to randomisation — hr, median (IQR)			
Skin tone (1–6), median (IQR)			
Skin tone ≥4 — no. (%)			
Randomised to conservative oxygen — no. (%)			
Unfavourable outcome at 180 days — no./N (%)			

Table 2: Risk Model Performance

Metric	Value
C-statistic (95% CI)	
Hosmer-Lemeshow P value	
Variables retained in final model	

Table 3: Treatment Effects Within Tertiles of Predicted Baseline Risk

Tertile	Conservative oxygen n/N (%)	Liberal oxygen n/N (%)	Absolute risk difference (95% CI)	Relative risk (95% CI)
Low risk (tertile 1)				
Medium risk (tertile 2)				
High risk (tertile 3)				
P for interacti on				

Table 4: Baseline Characteristics by Predicted ITE Tertile (Test Cohort)

Characteristic	Lower third (n≈202)	Middle third (n≈202)	Upper third (n≈203)	P value
Age — yr, median (IQR)				
Male sex — no. (%)				
Out-of-hospital arrest — no. (%)				
Shockable rhythm — no. (%)				
Medical cause — no. (%)				
Time to ROSC — min, median (IQR)				
STEMI — no. (%)				
Sepsis at baseline — no. (%)				
Non-HIE brain injury — no. (%)				
Targeted temperature management — no. (%)				
PaO2 at randomisation — mmHg, median (IQR)				
FIO2 at randomisation, median (IQR)				

Characteristic	Lower third (n≈202)	Middle third (n≈202)	Upper third (n≈203)	P value
Skin tone (1–6), median (IQR)				
Predicted baseline risk, median (IQR)				
Randomised to conservative oxygen — no. (%)				

Table 5: Primary Outcome by ITE Tertile and Treatment Group (Test Cohort)

	Lower third	Middle third	Upper third
Conservative oxygen — n/N (%) unfavourable			
Liberal oxygen — n/N (%) unfavourable			
Absolute risk difference (95% CI)			
Relative risk (95% CI)			
P value for interaction (likelihood ratio test)			

Table 6: Model Performance (Test Cohort)

Metric	Value
Adjusted Qini value	
C-for-benefit (bootstrapped 95% CI)	
Predicted ITE range	
Predicted ITE, median (IQR)	
Estimated absolute reduction in unfavourable outcomes with model-directed treatment (95% CI)	

Table 7: Skin Tone Subgroup Analysis

Skin tone group	Conservative oxygen n/N (%) unfavourable	Liberal oxygen n/N (%) unfavourable	Absolute risk difference (95% CI)	Relative risk (95% CI)	P for interaction
Lighter (1–3)					
Darker (4–6)					

Figure 1

Distribution of predicted individualised treatment effects in the test cohort. Histogram with vertical dashed lines marking tertile cut-points. Negative values indicate predicted benefit from conservative oxygen; positive values indicate predicted benefit from liberal oxygen.

Figure 2

Spaghetti plot. For 100 patients (one per percentile of predicted ITE), sequential replacement of baseline characteristics with derivation-cohort medians, showing the contribution of the 10 most important variables to each patient's predicted ITE.

Figure 3

Variable importance plot. Normalised importance (0–100) of each predictor variable in the final model.

Figure 4

Observed treatment effects by predicted ITE tertile. Absolute risk difference in unfavourable outcome at 180 days between conservative and liberal oxygen within each tertile, with 95% CI. P value for interaction shown.

Figure 5 (Supplemental)

Qini curve in the test cohort, with random-assignment baseline. Adjusted Qini value shown.

Figure 6 (Supplemental)

Calibration plot. Mean predicted ITE within each tertile compared with observed absolute risk difference within each tertile.

Figure 7 (Supplemental)

Partial dependence plots for the four most important continuous predictor variables, with bootstrapped 95% confidence intervals.

Figure 8 (Supplemental)

Panel A: Calibration plot for the baseline risk model. Panel B: Absolute risk difference (conservative minus liberal) and 95% CI within each tertile of predicted baseline risk.