

Statistical Analysis Plan

Haemoglobin concentration as an effect modifier of conservative oxygen therapy in mechanically ventilated adults

Secondary analysis of the ICU-ROX randomised trial with linkage to the ANZICS Adult Patient Database

SAP version	1.0
Date	28 September 2026
Status	Draft updated after inspection of available linked dataset; for investigator/statistician review
Primary analysis	Treatment × lowest first-24-hour haemoglobin interaction for 180-day all-cause mortality
Prepared for	ICU-ROX investigators

SAP status. This document is intended to be finalised before the haemoglobin-by-treatment analyses described below are undertaken. It is a post hoc secondary analysis of an existing randomised trial; all outcomes and exposure data have already been collected.

Revision compared to v0.2. Updated after inspection of the available merged ICU-ROX/APD dataset. Data-informed changes concern observed linkage/Hb availability, centre-level structural Hb missingness, haemoglobin units and validation, transfusion-data interpretation, explicit 180-day outcome derivation, and model-specific handling of ventilator-free days and centre. The primary treatment × continuous HMGNLO interaction for 180-day mortality is unchanged.

Core analysis decision

The primary effect modifier will be the lowest haemoglobin concentration recorded during the first 24 hours of ICU admission (ANZICS APD HMGNLO), analysed as a continuous variable. The primary outcome will be all-cause mortality at 180 days. The primary interaction model will use a linear haemoglobin term to provide a parsimonious, directly comparable randomised-trial replication of the PILOT analysis; a restricted cubic spline interaction will be a prespecified secondary analysis to assess non-linearity.

1. Background and rationale

ICU-ROX. ICU-ROX randomised 1000 critically ill adults expected to require invasive mechanical ventilation beyond the day after enrolment to conservative or usual oxygen therapy. The conservative strategy sought to minimise unnecessary oxygen exposure, including reducing FiO_2 to 0.21 when compatible with the specified lower oxygenation limit and using an upper SpO_2 alarm limit of 97%. The trial found no statistically significant difference in ventilator-free days to day 28 or in mortality at 180 days between treatment groups [1,2].

Biological rationale. Systemic oxygen delivery is determined by cardiac output and arterial oxygen content. Haemoglobin concentration is the principal determinant of the haemoglobin-bound component of arterial oxygen content. It is therefore biologically plausible that tolerance of lower arterial oxygenation could vary with haemoglobin concentration, even when the average effect of an oxygen strategy in an unselected trial population is neutral.

Prior observational evidence. Ramanan and Fisher analysed 219,723 mechanically ventilated adults in the ANZICS Adult Patient Database and reported that haemoglobin modified the association between oxygenation and hospital mortality. Lower haemoglobin was associated with greater mortality at lower PaO_2 and $\text{PaO}_2 / \text{FiO}_2$ values; the authors considered these observational findings hypothesis-generating and explicitly called for evaluation of oxygen and transfusion targets in higher-risk groups [3]. The study used first-24-hour APD physiology and is particularly relevant because the present analysis uses the same registry definition of early haemoglobin.

Prior randomised evidence. The PILOT trial compared lower, intermediate and higher SpO_2 targets during invasive mechanical ventilation [4]. A 2026 secondary analysis restricted to the lower (88–92%) and higher (96–100%) target groups evaluated baseline haemoglobin as a continuous treatment-effect modifier. Baseline haemoglobin did not modify the effect of oxygen target on ventilator-free days to day 28 (interaction $P=0.97$), and a prespecified sensitivity analysis using haemoglobin <100 versus ≥ 100 g/L (equivalent to <10 versus ≥ 10 g/dL in the original report) was also negative (interaction $P=0.83$) [5]. That analysis used a single baseline haemoglobin value and ventilator-free days as its primary outcome.

Rationale for the present analysis. The proposed ICU-ROX analysis provides a randomised evaluation of the haemoglobin–oxygen hypothesis in a distinct trial design and population, using early nadir haemoglobin rather than a single baseline value and 180-day mortality as the primary outcome. ICU-ROX compared conservative oxygen therapy with usual care rather than two widely separated SpO_2 targets. This analysis should therefore be interpreted as a replication and extension of the question, not as a first test of haemoglobin-related heterogeneity of oxygen treatment effect. Contemporary analyses of oxygen trials have also highlighted the possibility of clinically important heterogeneity of treatment effect across baseline physiological characteristics [6,7].

2. Objectives and hypotheses

2.1 Primary objective

To determine whether the lowest haemoglobin concentration recorded during the first 24 hours of ICU admission modifies the effect of assignment to conservative versus usual oxygen therapy on all-cause mortality at 180 days.

2.2 Primary hypothesis

The effect of conservative versus usual oxygen therapy on 180-day mortality varies according to early nadir haemoglobin concentration. The formal interaction test will be two-sided. A biologically plausible directional hypothesis is that lower haemoglobin may reduce tolerance of lower oxygenation; however, the analysis will not impose a directional alternative.

2.3 Secondary objectives

- Estimate the treatment effect on 180-day mortality across the observed distribution of early nadir haemoglobin on both relative and absolute effect scales.
- Assess whether the treatment–haemoglobin relationship is non-linear.
- Evaluate treatment-effect modification by haemoglobin for 90-day all-cause mortality, survival time to day 180, and ventilator-free days to day 28.
- Assess robustness to alternative haemoglobin definitions and handling of missing haemoglobin data.
- Describe whether the achieved separation in oxygen exposure between randomised groups appears materially different across the haemoglobin distribution only if the required longitudinal ICU-ROX exposure data are available in an analysis-ready source.

3. Study design and data sources

3.1 Parent trial

ICU-ROX was a multicentre, parallel-group randomised trial conducted in Australia and New Zealand. Allocation was 1:1 using variable blocks and was stratified by centre [1,2]. The present study is a secondary treatment-effect heterogeneity analysis using the original randomised allocation.

3.2 ICU-ROX case report form data

ICU-ROX collected demographics, comorbidities, admission characteristics, pre-randomisation physiology, APACHE II score, selected diagnoses and organ support, repeated oxygenation/ventilation data, and clinical outcomes. Pre-randomisation respiratory variables include SpO_2 , FiO_2 , PEEP, highest PaO_2 in the 12 hours before randomisation and PaO_2 immediately before randomisation. The CRF also records 180-day vital status and date of death when applicable [8].

3.3 ANZICS Adult Patient Database linkage

Of the 965 participants in the ICU-ROX intention-to-treat/source analysis population, 914 (94.7%) have a successfully linked ANZICS Adult Patient Database (APD) record in the available merged dataset. HMGNLO is populated for 753 of the 914 linked participants (82.4%) before application of validity checks and outcome-availability criteria. The APD contains demographic, physiological, biochemical, diagnostic, treatment and outcome variables for ICU episodes [9].

3.4 Haemoglobin variable

Primary effect modifier: HMGNLO, the lowest haemoglobin concentration recorded during the first 24 hours of ICU admission. The APD dictionary specifies that HMGNHI and HMGNLO are the highest and lowest haemoglobin values during the first 24 hours; if only one haemoglobin is available it is entered as both high and low, and a value from up to one hour before ICU admission may be used only if no first-24-hour result is available [9].

Interpretation: HMGNLO will be interpreted as early nadir haemoglobin, not necessarily a strictly pre-randomisation or pre-transfusion measurement. The available merged dataset contains transfusion summary variables, but these do not reliably establish whether transfusion occurred before or after the first-day haemoglobin nadir. Transfusion will therefore not be used as a primary adjustment or exclusion variable. Although some haemoglobin measurements may occur after randomisation, the oxygen strategy is not expected to materially alter haemoglobin concentration over the first ICU day. The treatment-by-haemoglobin analysis therefore treats HMGNLO as a baseline/early physiological effect modifier. The uncertain temporal relationship between transfusion and HMGNLO will be explicitly acknowledged in interpretation.

4. Analysis populations

4.1 Source ICU-ROX population

The source population will be the ICU-ROX analysis population for whom use of trial data is permitted under the parent study. Participants will remain classified according to randomised treatment assignment, irrespective of adherence to allocated oxygen strategy.

4.2 Primary haemoglobin analysis population

The primary analysis population will comprise source-population participants who: (1) have a successfully linked APD record; (2) have a valid non-missing HMGNLO value after prespecified data checks; and (3) have known 180-day vital status. No participant will be excluded on the basis of haemoglobin concentration alone, diagnosis, transfusion status, or post-randomisation treatment adherence. Because HMGNLO is structurally unavailable for participants from at least one centre (Wellington), the primary

analysis population represents the linked ICU-ROX population from centres contributing usable haemoglobin data rather than the entire ICU-ROX source population.

4.3 Participant accounting

A flow diagram will report the 1000 participants randomised, the 965 participants in the ICU-ROX intention-to-treat/source analysis population, the 914 linked to APD, the number with HMGNLO recorded (753 before validity checks), the number with valid HMGNLO after data checks, the number with known 180-day vital status, and the final primary analysis population by randomised group. Centre-level structural absence of HMGNLO will be shown separately from participant-level missingness.

5. Outcomes

5.1 Primary outcome

All-cause mortality by day 180 after randomisation, analysed as a binary outcome (dead by day 180 vs alive at day 180). Deaths occurring after day 180 will be coded as alive at day 180 for this binary outcome. The derivation will use the ICU-ROX 180-day vital-status information and date of death where required to distinguish deaths within and beyond the 180-day window.

5.2 Secondary outcomes

- All-cause mortality by day 90 after randomisation.
- Time from randomisation to death through day 180. Follow-up time will be defined as the minimum of time from randomisation to death and 180 days; participants alive at day 180 and participants dying after day 180 will be censored at day 180.
- Ventilator-free days to day 28, using the parent ICU-ROX definition in which patients who die by day 28 receive zero ventilator-free days [1,2].
- Exploratory oxygen-exposure summaries will be considered only if the required longitudinal ICU-ROX oxygen-exposure data are available in an analysis-ready source. The currently available linked APD extract does not by itself define the planned summaries (for example, hours with $\text{SpO}_2 > 96\%$ or hours with $\text{FiO}_2 > 0.21$), so these analyses are not required for completion of the primary or secondary outcome analyses.

6. Variable definition and data preparation

6.1 Haemoglobin units and validity checks

- HMGNLO and HMGNI are supplied in the linked extract in g/dL and will be converted to g/L by multiplying by 10 before analysis. All analysis variables, tables, figures and reported effect estimates will use g/L.
- The APD permissible range of 1–25 g/dL corresponds to 10–250 g/L. Values outside the permissible range will be treated as data errors and set to missing unless a source correction is available.
- Values outside 50–180 g/L are flagged by the APD as outside the usual range but are not intrinsically invalid. Extreme values will be checked for internal consistency using HMGNI, haematocrit and other available laboratory information. An observed HMGNLO value equivalent to approximately 13 g/L is internally inconsistent with the corresponding haematocrit and will be treated as erroneous unless source verification supports it. Other extreme but internally plausible values will be retained.
- Internal consistency will be checked where HMGNI and/or haematocrit are available. HMGNLO should not exceed HMGNI, and gross discordance between haemoglobin and haematocrit will trigger review before the value is retained or set to missing.

6.2 Coding of randomised treatment

Treatment will be coded as 1 for conservative oxygen therapy and 0 for usual oxygen therapy. All treatment-effect estimates will compare conservative with usual oxygen therapy.

6.3 Centre

Randomisation centre will be retained as a design variable because randomisation was stratified by centre. Primary logistic regression models will include centre as a random intercept. If a mixed-effects model fails to converge, a fixed-effect centre term will be used; if sparse centres prevent stable estimation, a model without centre will be reported as the primary fallback with the centre-adjusted model reported where estimable. Model-specific handling of centre for Cox and proportional-odds analyses is described below.

7. Descriptive analyses

Descriptive analyses will precede outcome modelling and will be performed without reference to statistical significance of baseline differences.

- Report APD linkage and HMGNLO availability overall and by treatment group and centre. Structural centre-level absence of HMGNLO will be distinguished from sporadic participant-level missingness.
- Summarise HMGNLO by treatment group using mean (SD), median (IQR), range, and selected percentiles; show overlaid density or histogram plots.
- Compare participants with versus without available HMGNLO using ICU-ROX variables available independently of the APD. Standardised mean differences will be used rather than hypothesis tests. Because an entire centre has structurally unavailable HMGNLO, comparisons will be presented overall and, where informative, after restricting to centres in which HMGNLO is observed for at least some participants.
- Summarise baseline clinical characteristics by treatment group in the haemoglobin analysis population. Where variables are available, the baseline table will mirror the order, definitions and summary format used in Table 1 of the primary ICU-ROX publication [2], with HMGNLO added as the study-specific effect-modifier variable.
- Estimate the randomised-group difference in HMGNLO with a 95% CI as a diagnostic assessment of whether treatment allocation appears materially associated with the early nadir haemoglobin measurement. This will not be used as a formal balance test.

8. Primary analysis

8.1 Primary estimand

The primary estimand is the change in the relative effect of conservative versus usual oxygen therapy on the odds of death by day 180 per 10 g/L increase in early nadir haemoglobin, within linked ICU-ROX participants from centres contributing a valid observed HMGNLO value and with known 180-day vital status. In addition, model-based absolute risks and absolute risk differences will be estimated across the observed haemoglobin distribution. Generalisation beyond centres with observed haemoglobin will be made cautiously because part of the HMGNLO missingness is structural at centre level.

8.2 Primary regression model

A mixed-effects logistic regression model will be fitted with 180-day mortality as the dependent variable, centre as a random intercept, and the following fixed effects: randomised treatment, HMGNLO (continuous, measured in g/L and centred at 100 g/L), and the treatment × HMGNLO interaction. The primary model will use a linear haemoglobin term. The model may be written as:

$$\text{logit}[P(\text{death by day 180})] = \beta_0 + u(\text{center}) + \beta_1 T + \beta_2 (Hb - 100) + \beta_3 T(Hb - 100) \quad (Hb \text{ in g/L})$$

The primary null hypothesis is $H_0 : \beta_3 = 0$. A two-sided Wald test at $\alpha=0.05$ will be used. The interaction coefficient will be reported as an interaction odds ratio per 10 g/L higher haemoglobin with a 95% confidence interval and exact P value.

8.3 Why a linear interaction is primary

A linear haemoglobin interaction is prespecified as the primary analysis because it is parsimonious for the available sample size and event count and provides the closest randomised-trial replication of the 2026 PILOT haemoglobin analysis [5]. The analysis does not assume that the underlying mortality risk is linear in haemoglobin on the probability scale; treatment effects will also be examined flexibly as described below.

8.4 Clinical presentation of the primary model

From the fitted model, treatment effects will be presented at clinically interpretable haemoglobin values that lie within the central observed data range (provisionally 70, 80, 90, 100, 120 and 140 g/L; values outside the 5th–95th percentile range will not be emphasised). For each value, the following will be reported with 95% CIs:

- predicted 180-day mortality under conservative oxygen therapy;
- predicted 180-day mortality under usual oxygen therapy;
- absolute risk difference (conservative minus usual); and
- odds ratio for conservative versus usual oxygen therapy.

Predicted risks and risk differences will be calculated from the fitted model by marginal standardisation. Plots will show predicted mortality in each treatment group and the estimated treatment effect across haemoglobin, with pointwise 95% CIs. Estimates in sparsely populated tails will be visually identified and interpreted cautiously.

9. Secondary and sensitivity analyses

9.1 Flexible haemoglobin–treatment interaction

To assess non-linearity, HMGNLO will be modelled using a restricted cubic spline with three knots placed at the 10th, 50th and 90th percentiles of HMGNLO in the pooled primary analysis population. Treatment will be interacted with all spline basis terms. A joint likelihood-ratio or Wald test of the treatment × spline terms will assess evidence of treatment-effect heterogeneity. The resulting absolute and relative treatment-effect curves will be plotted. This analysis is secondary and will not replace the prespecified linear interaction test as the primary inferential analysis.

9.2 Replication of the PILOT haemoglobin threshold analysis

As a sensitivity analysis directly comparable with Flemington et al. [5], haemoglobin will be dichotomised at <100 versus ≥100 g/L and a treatment × haemoglobin-category interaction will be fitted for 180-day mortality and ventilator-free days. Results will be treated as supportive rather than primary because dichotomisation discards information and the threshold was not developed from ICU-ROX.

9.3 Alternative haemoglobin definitions

- Repeat the primary interaction model using HMGNHI (highest haemoglobin during the first ICU day).
- Repeat using the midpoint of HMGNLO and HMGNHI when both are available, as a crude measure of central first-day haemoglobin exposure.
- Repeat after excluding values <50 or >180 g/L to assess sensitivity to extreme but potentially valid APD measurements.

9.4 Other outcomes

90-day mortality: the primary logistic interaction model will be repeated with death by day 90 as the outcome.

Survival to day 180: a Cox proportional hazards model will include treatment, continuous HMGNLO, and their interaction. Follow-up will be administratively censored at 180 days, including for participants whose

death occurs after day 180. Centre will be handled using shared frailty or robust centre-clustered standard errors. Proportional hazards assumptions will be assessed using Schoenfeld residuals. This analysis will be supportive because an interaction in hazard ratios is not identical to interaction in 180-day risk.

Ventilator-free days to day 28: to facilitate comparison with the PILOT haemoglobin analysis, a proportional-odds model will be fitted with the observed ventilator-free-day values as the ordered dependent variable and treatment, continuous HMGNO, and treatment $\times$ HMGNO as predictors [5]. The available ventilator-free-day variable contains fractional-day values; these will be retained and ranked directly rather than rounded. Centre will be included as a fixed effect for this supportive analysis. A dichotomous $<100/\geq 100$ g/L sensitivity analysis will also be performed. The proportional-odds estimand will be reported with 95% CI; results will be interpreted as supportive.

9.5 Covariate-adjusted sensitivity analysis

Because treatment was randomised, the primary interaction model will not require broad covariate adjustment. A prespecified covariate-adjusted sensitivity model will include age, sex and APACHE II score, thereby explicitly adjusting for baseline illness severity and aligning, where possible, with the covariate-adjustment framework planned for the parent ICU-ROX trial [1,2]. APACHE II score will be entered as a continuous variable. Centre will continue to be included as a random intercept. To avoid redundant adjustment for individual components of APACHE II, baseline creatinine and oxygenation variables will not also be included in this core adjusted model. If a clinically important baseline imbalance is identified in the primary haemoglobin analysis population (absolute standardised mean difference >0.10), an additional expanded sensitivity model may include that variable; such additions will be clearly labelled and will not replace the prespecified age-, sex- and APACHE-II-adjusted model.

9.6 Oxygen-exposure consistency analysis

The primary and prespecified secondary clinical analyses do not depend on derivation of oxygen-exposure summaries. If an analysis-ready longitudinal ICU-ROX oxygen dataset containing the required exposure information is available, treatment separation (for example, hours with $\text{SpO}_2 >96\%$ and hours with $\text{FiO}_2 0.21$) may be described across haemoglobin using appropriate descriptive or regression methods. If such data are not available, this mechanistic analysis will be omitted and this will not be considered a deviation from the primary SAP.

10. Missing data

10.1 Primary approach

The primary analysis will be a complete-record analysis among participants with a linked APD record, valid observed HMGNO and known 180-day vital status. The number and proportion excluded for each reason will be reported by randomised group and, for HMGNO, by centre. Structural centre-level absence of HMGNO will be distinguished from participant-level missingness.

10.2 Assessment of missingness

Patterns of linkage failure and missing HMGNO will be described by centre, year and treatment group. In the available linked dataset, HMGNO is absent for all participants from Wellington, demonstrating that an important component of missingness is structural at centre level rather than ordinary participant-level missing-at-random missingness. Participants with and without HMGNO will be compared using variables captured on the ICU-ROX CRF, both overall and, where informative, after restricting to centres with at least some observed HMGNO. Particular attention will be paid to treatment allocation and severity of illness.

10.3 Multiple-imputation sensitivity analysis

Because HMGNO is structurally unavailable for an entire centre, HMGNO will not be multiply imputed for participants from centres with no observed HMGNO: the within-centre haemoglobin distribution cannot be learned from those data, and imputing it would rely heavily on between-centre extrapolation. The primary analysis will therefore remain complete-record. As a supportive sensitivity analysis, multiple imputation

may be undertaken among linked participants from centres in which HMGNLO is partially observed, if participant-level HMGNLO missingness within those centres is sufficient to make such an analysis informative. If undertaken, substantive-model-compatible fully conditional specification will be preferred to preserve the treatment $\times$ haemoglobin interaction. The imputation model will include randomised treatment, centre, 180-day mortality, age, sex, admission category, pre-randomisation physiology and creatinine, comorbidities, APACHE diagnosis, and HMGNHI where available. At least 50 imputed datasets will be generated, with estimates combined using Rubin's rules. APD-unmatched participants and participants from centres with no observed HMGNLO will not be included in the Hb imputation analysis.

10.4 Missing outcomes

Missingness of 180-day vital status was low in the primary ICU-ROX trial (14 of 965 participants; 1.5%). The primary analysis will therefore use participants with observed 180-day vital status, without imputation of the primary outcome. The number and proportion of participants with missing 180-day vital status will be reported overall and by randomised treatment group. Baseline characteristics of participants with and without known 180-day vital status will be described where informative. Given the small amount of outcome missingness, no multiple-imputation analysis of 180-day mortality is planned. If the proportion of missing 180-day vital status is materially greater in the APD-linked haemoglobin analysis population than in the primary trial population, additional sensitivity analyses addressing missing outcomes will be considered and specified before examination of treatment-by-haemoglobin interaction results.

11. Multiplicity, statistical significance and interpretation

There will be one primary inferential test: the treatment $\times$ continuous HMGNLO interaction for 180-day mortality in the primary model. A two-sided $P < 0.05$ will be considered evidence against the null hypothesis for this test. No multiplicity adjustment will be applied to secondary or sensitivity analyses; these will be labelled exploratory/supportive and interpreted primarily through effect estimates and 95% confidence intervals rather than nominal P values. Absence of a statistically significant interaction will not be interpreted as proof of homogeneity; the width of the confidence interval and range of compatible treatment effects across haemoglobin will be emphasised.

12. Sample size and precision

This is a fixed-sample secondary analysis and no new sample-size calculation determines enrolment. ICU-ROX randomised 1000 participants; 965 were included in the parent intention-to-treat/source analysis population, 914 (94.7%) are linked to APD in the available merged dataset, and HMGNLO is recorded for 753 linked participants (82.4%) before validity and outcome-availability checks. Thus, the primary interaction analysis will contain materially fewer participants and events than the parent trial and may have limited precision for detecting modest treatment-effect heterogeneity. The parent trial observed 180-day mortality of approximately 35% [2]. No post hoc power calculation will be used to judge the findings. Precision will be conveyed using confidence intervals around the interaction and treatment-effect curves.

13. Statistical software and reproducibility

Analyses will be performed in R (version 4.6.1 or later). Analysis code will be version controlled. All data derivations, exclusions, model specifications and deviations from this SAP will be documented before manuscript submission.

14. Planned tables and figures

Mock table shells are provided in Appendix B. Values shown as XX are placeholders and will be replaced with observed estimates.

- Figure 1. Participant flow from ICU-ROX randomisation through APD linkage, haemoglobin availability and primary analysis inclusion.
- Figure 2. Distribution of HMGNLO overall and by randomised treatment group.

- Figure 3. Model-based 180-day mortality across haemoglobin for conservative and usual oxygen therapy (primary linear interaction; flexible spline overlaid or presented separately).
- Figure 4. Estimated treatment effect across haemoglobin: absolute risk difference and odds ratio with 95% CIs.
- Table 1. Baseline characteristics of the primary analysis population by randomised treatment group, aligned where possible with Table 1 of the primary ICU-ROX publication and including HMGNLO.
- Table 2. APD linkage, HMGNLO availability and derivation of the primary analysis population by treatment group, with a centre-level panel showing structural and participant-level HMGNLO missingness.
- Table 3. Primary and secondary treatment × haemoglobin interaction estimates.
- Supplementary Table S1. Characteristics of participants with versus without available HMGNLO, overall and, where informative, restricted to centres with at least some observed HMGNLO.
- Supplementary Table S2. Sensitivity analyses: spline interaction, Hb <100/≥100 g/L, HMGNHI, midpoint Hb, extreme-value exclusion, covariate adjustment and, if informative, multiple imputation restricted to centres with partially observed HMGNLO.

15. Planned interpretation

The primary analysis will be framed as an assessment of whether early nadir haemoglobin modifies the causal effect of randomised oxygen strategy on 180-day mortality among linked ICU-ROX participants with valid observed HMGNLO. Because HMGNLO may include values obtained after randomisation and the available transfusion summaries do not establish transfusion timing relative to the Hb nadir, the modifier will be described as early nadir haemoglobin rather than untreated baseline haemoglobin. HMGNLO is missing for a substantial minority of linked participants (161/914, 17.6% before validity checks), including structural absence for an entire centre; consequently, generalisability of the HTE analysis to the full ICU-ROX population will be considered explicitly. The central causal protection comes from randomised treatment assignment; however, subgroup/effect-modifier analyses remain vulnerable to limited precision, scale dependence and chance. Results will be interpreted together with the observational APD findings of Ramanan and Fisher [3] and the negative randomized PILOT secondary analysis of Flemington et al. [5].

16. References

1. Mackle DM, Bailey MJ, Beasley RW, et al.; Australian and New Zealand Intensive Care Society Clinical Trials Group. Protocol summary and statistical analysis plan for the intensive care unit randomised trial comparing two approaches to oxygen therapy (ICU-ROX). *Crit Care Resusc.* 2018;20(1):22-32. doi:10.1016/S1441-2772(23)00728-7.
2. ICU-ROX Investigators and the Australian and New Zealand Intensive Care Society Clinical Trials Group. Conservative oxygen therapy during mechanical ventilation in the ICU. *N Engl J Med.* 2020;382(11):989-998. doi:10.1056/NEJMoa1903297.
3. Ramanan M, Fisher N. The association between arterial oxygen tension, hemoglobin concentration, and mortality in mechanically ventilated critically ill patients. *Indian J Crit Care Med.* 2018;22(7):477-484. doi:10.4103/ijccm.IJCCM_66_18.
4. Semler MW, Casey JD, Lloyd BD, et al.; PILOT Investigators and the Pragmatic Critical Care Research Group. Oxygen-saturation targets for critically ill adults receiving mechanical ventilation. *N Engl J Med.* 2022;387(19):1759-1769. doi:10.1056/NEJMoa2208415.
5. Flemington AB, Wang L, Casey JD, Rice TW, Semler MW, Qian ET. Baseline hemoglobin level does not modify the effect of oxygen saturation target on outcomes for critically ill adults receiving mechanical ventilation. *Am J Respir Crit Care Med.* 2026;212(3):533-536. doi:10.1093/ajrccm/aamaf121.
6. Granholm A, Schjørring OL, Møller MH, et al. Lower versus higher oxygenation targets in critically ill patients with severe hypoxaemia: secondary Bayesian analysis to explore heterogeneous treatment effects in the Handling Oxygenation Targets in the Intensive Care Unit (HOT-ICU) trial. *Br J Anaesth.* 2022;128(1):55-64. doi:10.1016/j.bja.2021.09.010.

7. Martin DS, Grocott MPW. Heterogeneity of treatment effect: the case for individualising oxygen therapy in critically ill patients. *Crit Care*. 2025;29:50. doi:10.1186/s13054-025-05254-5.
8. ICU-ROX Investigators. ICU-ROX Case Report Form, version 3.1, 25 February 2015. Study document.
9. Australian and New Zealand Intensive Care Society Centre for Outcome and Resource Evaluation. ANZICS CORE Adult Patient Database Data Dictionary, version 6.1. April 2022.

Appendix A. Key prespecified analysis decisions

Issue	Prespecified decision	Rationale
Primary modifier	Lowest Hb in first 24 h (HMGNO), continuous	Represents early nadir oxygen-carrying capacity; directly available from APD.
Primary outcome	180-day all-cause mortality	Patient-important long-term outcome; distinguishes this analysis from the PILOT Hb analysis.
Primary functional form	Linear treatment × Hb interaction	Parsimonious and directly comparable with the prior randomized PILOT analysis.
Non-linearity	3-knot restricted cubic spline interaction, secondary	Allows clinically important departures from linearity without making the primary model too data-hungry.
Hb threshold	<100 vs ≥100 g/L, sensitivity only	Included for direct comparison with Flemington et al.; not used as primary because of information loss.
Transfusion	No primary adjustment/exclusion; exploratory description if useful	Transfusion summary variables are available, but timing relative to first-day Hb nadir cannot be reliably established.
Missing Hb	Complete-record primary; no imputation for centres with structurally absent Hb; optional MI within centres with partially observed Hb	Observed data show centre-level structural missingness, so broad MAR imputation across all centres would require unsupported between-centre extrapolation.
Multiplicity	One primary interaction test at two-sided $\alpha=0.05$	All other analyses supportive/exploratory.
Covariate adjustment	Age, sex and APACHE II score; centre random intercept	Prespecified sensitivity analysis. APACHE II provides explicit adjustment for baseline illness severity and aligns with the parent ICU-ROX adjustment framework.

Appendix B. Mock tables

These table shells define the intended presentation of the principal results. All entries shown as XX are placeholders. Continuous variables will be summarised as mean (SD) or median [IQR] as appropriate; categorical variables as n (%). Baseline tables will not include hypothesis-test P values.

Mock Table 1. Baseline characteristics of the primary haemoglobin analysis population by randomised treatment group

Characteristic	Conservative oxygen (N=XX)	Usual oxygen (N=XX)
Age — yr	XX±XX	XX±XX
Male sex — no. (%)	XX (XX.X)	XX (XX.X)
Source of admission to ICU — no. (%)		
Emergency department	XX (XX.X)	XX (XX.X)
Hospital ward	XX (XX.X)	XX (XX.X)
Transfer from another ICU	XX (XX.X)	XX (XX.X)
Transfer from a non-ICU ward of another hospital	XX (XX.X)	XX (XX.X)
Operating room		
After elective surgery	XX (XX.X)	XX (XX.X)
After emergency surgery	XX (XX.X)	XX (XX.X)
Median hr from initiation of invasive ventilation to randomisation (IQR)	XX (XX–XX)	XX (XX–XX)
APACHE II score	XX.X±XX.X	XX.X±XX.X
Diagnosis subgroup		
Admitted after surgery — no. (%)	XX (XX.X)	XX (XX.X)
Acute brain disease — no. (%)	XX (XX.X)	XX (XX.X)
PaO ₂ :FiO ₂ ratio <300 — no./total no. (%)	XX/XX (XX.X)	XX/XX (XX.X)
Suspected hypoxic–ischaemic encephalopathy — no. (%)	XX (XX.X)	XX (XX.X)
Physiological features		
SpO ₂ — %	XX.X±XX.X	XX.X±XX.X
Median PaO ₂ (IQR) — mm Hg	XX (XX–XX)	XX (XX–XX)
PaO ₂ :FiO ₂ ratio	XX±XX	XX±XX
Median PEEP (IQR) — cm H ₂ O	XX (XX–XX)	XX (XX–XX)
Study-specific linked variable		
Lowest haemoglobin in first 24 h (HMGNLO) — g/L	XX (XX–XX)	XX (XX–XX)

Baseline variables and ordering mirror Table 1 of the primary ICU-ROX publication where available. Summary conventions are likewise retained (mean±SD, median [IQR], or n [%] as shown). HMGNLO is added as the study-specific linked variable. Hb = haemoglobin; FiO₂ = fraction of inspired oxygen; SpO₂ = peripheral oxygen saturation; PaO₂ = arterial oxygen tension; PEEP = positive end-expiratory pressure.

Mock Table 2. APD linkage, haemoglobin availability, and derivation of the primary analysis population

Participant accounting	Conservative oxygen	Usual oxygen	Overall
Randomised, n	500	500	1000
ICU-ROX intention-to-treat/source analysis population, n	484	481	965
Successfully linked to APD, n (%)	XX (XX)	XX (XX)	914 (94.7)
HMGNLO available among linked participants, n (%)	XX (XX)	XX (XX)	753 (82.4)
Known 180-day vital status, n (%)	XX (XX)	XX (XX)	XX (XX)
Included in primary haemoglobin analysis, n (%)	XX (XX)	XX (XX)	XX (XX)

Percentages for linkage and HMGNLO availability use the relevant preceding population as denominator. In the available extract, 914/965 source-population participants are linked to APD and HMGNLO is populated for 753/914 linked participants before validity checks. The final table will report randomised-group-specific counts and will include a centre-level panel distinguishing structural centre-level absence (including Wellington) from sporadic participant-level missingness.

Mock Table 3. Primary and secondary treatment-by-haemoglobin interaction analyses

Outcome	Model / effect scale	Interaction estimate per 10 g/L higher Hb	95% CI	P for interaction
180-day mortality (primary)	Logistic regression; ratio of treatment ORs	XX	XX to XX	0.XX
90-day mortality	Logistic regression; ratio of treatment ORs	XX	XX to XX	0.XX
Time to death to day 180	Cox regression; ratio of treatment HRs	XX	XX to XX	0.XX
Ventilator-free days to day 28	Prespecified regression model; interaction coefficient	XX	XX to XX	0.XX

The primary inferential test is the treatment × continuous HMGNLO interaction for 180-day mortality. Hb will be centred at 100 g/L; interaction estimates are expressed per 10 g/L higher Hb.

Mock Table 4. Model-estimated effect of conservative versus usual oxygen therapy on 180-day mortality at representative haemoglobin concentrations

Lowest Hb, g/L	Mortality: conservative oxygen, %	Mortality: usual oxygen, %	Risk difference, % points (95% CI)	Odds ratio (95% CI)
70	XX.X	XX.X	XX.X (XX.X to XX.X)	XX.XX (XX.XX to XX.XX)
80	XX.X	XX.X	XX.X (XX.X to XX.X)	XX.XX (XX.XX to XX.XX)
90	XX.X	XX.X	XX.X (XX.X to XX.X)	XX.XX (XX.XX to XX.XX)
100	XX.X	XX.X	XX.X (XX.X to XX.X)	XX.XX (XX.XX to XX.XX)
120	XX.X	XX.X	XX.X (XX.X to XX.X)	XX.XX (XX.XX to XX.XX)
140	XX.X	XX.X	XX.X (XX.X to XX.X)	XX.XX (XX.XX to XX.XX)

Representative Hb values will be restricted to concentrations with adequate empirical support. If 70 or 140 g/L falls outside a stable data region, the displayed values will be revised before results are generated, without changing the underlying continuous model.

Mock Table 5. Prespecified sensitivity analyses for the primary 180-day mortality interaction

Sensitivity analysis	N	Deaths	Interaction estimate	95% CI	P for interaction
Primary linear HMGNLO model	XX	XX	XX	XX to XX	0.XX
3-knot restricted cubic spline interaction	XX	XX	Global interaction $\chi^2=XX$	—	0.XX
Hb <100 vs ≥ 100 g/L	XX	XX	Ratio of treatment ORs=XX	XX to XX	0.XX
Highest first-24-hour Hb (HMGNHI)	XX	XX	XX	XX to XX	0.XX
Midpoint of HMGNLO and HMGNHI	XX	XX	XX	XX to XX	0.XX
Exclude Hb <50 or >180 g/L	XX	XX	XX	XX to XX	0.XX
Adjusted for age, sex and APACHE II score	XX	XX	XX	XX to XX	0.XX
Multiple imputation within centres with partially observed HMGNLO (if informative)	XX	XX	XX	XX to XX	0.XX

For the spline model, the global interaction test will jointly test all treatment-by-spline terms. Secondary and sensitivity analyses will be interpreted as supportive/exploratory without multiplicity adjustment.