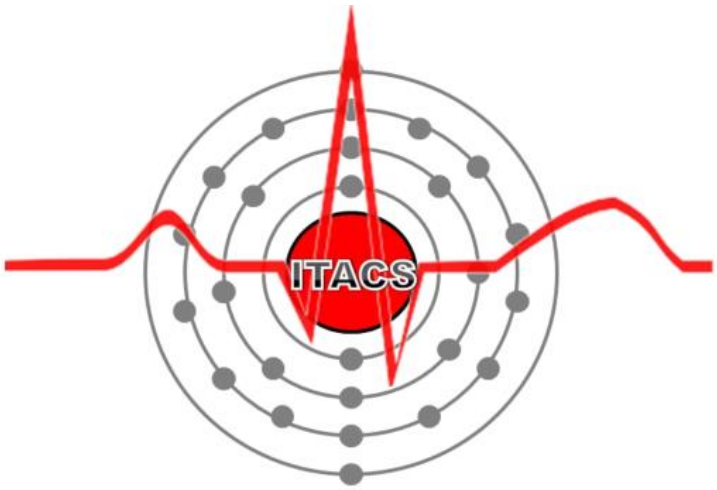

The ITACS Trial

IV iron for Treatment of Anaemia before Cardiac Surgery

www.itacs.org.au

STATISTICAL ANALYSIS PLAN

SAP Author: Prof Stephane Heritier, School of Public Health and Preventive Medicine, Monash University

SAP Version: 1.0

SAP Date: December 19, 2024

Table of Contents

Abbreviations and Definitions

ABBREVIATION	DEFINITION
ACEi	Angiotensin-Converting Enzyme inhibitors
ASA	American Society of Anaesthesiologists physical status
AE	Adverse event
CABG	Coronary Artery Bypass Graft
CI	Confidence interval
CONSORT	Consolidated Standards of Reporting Trials
CRF	Case Report Form
CRP	C-reactive protein
DAH	Days alive and at home
EF	Ejection fraction
EQ-5D-5L	5-level EQ-5D
eGFR	estimated Glomerular Filtration Rate
FBE	Full Blood Examination
FFP	Fresh Frozen Plasma
GGT	Gamma-glutamyl Transferase amount
Hb	Haemoglobin
ICU	Intensive Care Unit
LOS	Length of stay
MCV	Mean Corpuscular Volume
RBC	Red Blood Cell count
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
SUSAR	Suspected Unexpected Serious Adverse Reaction
TAVI	Transcatheter Aortic Valve Implantation
TSAT	Transferrin saturation
TSH	Thyroid Stimulating Hormone
WHODAS 2.0	World Health Organisation Disability Assessment Schedule 2.0

1 Introduction

1.1 Background

Approximately 20% to 30% of patients awaiting cardiac surgery are anaemic. Anaemia is a serious condition increasing the likelihood of requiring a red cell transfusion and is typically associated with increased complications, intensive care, and longer hospital stay after surgery. Iron deficiency is the most common cause of anaemia and it is therefore plausible that administration of preoperative intravenous (IV) iron improve patient outcomes after cardiac surgery. The trial was specifically designed to test the hypothesis that intravenous iron for treatment of anaemia before cardiac surgery (ITACS) will reduce complications and facilitate recovery in anaemic patients awaiting elective surgery. A patient-centred outcome was chosen as the primary endpoint to emphasize the importance given to the patient's experience during their journey to recovery

1.2 Purpose of the analyses

The purpose of this Statistical Analysis Plan (SAP) is to outline the pre-planned analyses to be completed to support the main publication of the ITACS study. Versions of the SAP will be tracked, with a clear distinction between changes before and after start of the statistical analysis.

2 Study design

ITACS is a multicentre, international, double-blind randomised clinical trial comparing the efficacy, safety and cost effectiveness of preoperative intravenous (IV) iron,1000 mg, single dose infusion given 1-26 weeks before the planned date of surgery) with placebo in patients with anaemia before elective cardiac surgery. The trial will incorporate the intervention into normal preoperative surgical pathways for hospital patients. The Trial is registered at ClinicalTrials.gov, Identifier NCT02632760, and the rationale/design article was published in the American Heart Journal in 2021 (doi.org/10.1016/j.ahj.2021.05.008).

2.1 Objectives

Primary objective: To determine whether a single dose of IV iron (ferric carboxymaltose [Ferrinject®] or ferric derisomaltose [Monofer®] 1000 mg, given to anaemic patients before elective cardiac surgery improves the primary outcome (i.e.

days alive at home up to 90 days (DAH90) after surgery, a validated patient-centred outcome [1]).

Secondary objectives: To assess whether IV iron improves patient recovery because of correction of preoperative iron deficiency anaemia, reduced the need for blood transfusion and risk of complications, manifesting as earlier hospital discharge, lower mortality and reduced readmissions after surgery. We also aim to demonstrate that IV iron is cost-effective in this setting.

Study Hypotheses

Primary: A single dose of IV iron in the weeks before surgery corrects anaemia in patients undergoing elective cardiac surgery, and thereby improves patient recovery and their return to home when compared with placebo.

Secondary (Health Economics): Preoperative administration of IV iron will be 'dominant' when compared with current practice – that is, it will both save money and improve health outcomes.

2.2 Patient population

As indicated in the protocol and reported in our rationale/design publication [2] patients must satisfy all the following inclusion criteria and none of exclusion criteria to be eligible to be randomised to the study

Inclusion criteria

- Adult (≥18 years) with anaemia (male or female Hb <130 g/L)
- Expected to undergo elective on-pump or off-pump cardiac surgery
- Able to receive trial drug 1- 26 weeks prior to planned surgery.
-

Exclusion criteria

- Transcatheter aortic valve implantation (TAVI) and other catheter laboratory interventional procedures
- Pregnancy
- Known hypersensitivity to study drug (ferric carboxymaltose, iron isomaltoside, or equivalent) or its excipients
- Previously documented TSAT >50%
- Previously documented vitamin B12 or folate deficiency
- Known or suspected hemoglobinopathy/thalassemia

- Bone marrow disease
- Hemochromatosis
- Renal dialysis
- Erythropoietin or IV iron in the previous 4 weeks

Note: Oral iron (tablets, capsules) therapy is not an exclusion criterion.

2.3 Randomisation

After consent has been obtained, patients are randomly assigned (1:1) to groups via a web-based service using computer-generated code to either receive IV iron or placebo. Group assignment is stratified by site, baseline haemoglobin ($<$ or $\geq$100 g/L), and planned surgery (isolated coronary artery or single valve surgery, or combined/other), using a mix of randomly permuted blocks of different sizes.

2.4 Baseline characteristics

Demographics are collected at baseline and include age, sex, weight (kg), weight (cm), ASA, and ethnicity (White/Caucasian, Indian/Pakistan, Asian, Black Other). Various laboratory parameters are also measured i.e. haemoglobin (g/L), creatinine (μmol/L), platelets, CRP (mg/L), Ferritin (ug/L), Transferrin (g/L), Serum Phosphate, (mmol/L), TSAT. A standard measure of health status, EQ-5D [3] is undertaken at baseline and is composed of 5 different domains with 5 questions per domain: mobility, self-care, usual activities, pain/discomfort; anxiety/depression. The EQ5D-VAS score is also collected. The WHODAS 2.0 assessment covering 6 domains of functioning [4] and the MacNew score [5] will also be calculated at baseline (prior to surgery). Finally, information on adverse events (AEs) is collected including AE severity (mild, moderate, severe), AE by system (neurological, cardiovascular, skin, other), SAE/SUSARs at the time of study drug administration.

2.5 Preoperative characteristics

Pre-existing conditions or past events prior to surgery will be recording including smoking history, heart failure, chronic lung disease, diabetes, previous MI.

Other standard pre-operative characteristics include 1) prior and concurrent medication i.e. ACE/ARB, β-blockers, Ca Channel blocker, statin, oral iron, IV parenteral iron, red cell transfusion; 2) laboratory parameters e.g. haemoglobin (g/L), haematocrit (L/L), creatinine (μmol/L), CRP (mg/L), ferritin (μg/L), transferrin (g/L), mean corpuscular volume [MCV] (fl), reticulocyte count(x10^9/L, TSAT (%), platelets, TSH (mIU/L), free thyroxine T4 (pmol/L, alanine transaminase (ALT) (U/L), vitamin B12 (pmol/L), serum folate (nmol/L), GGT (U/L), serum phosphate (mmol/L).

Finally, the risk of in-hospital mortality after cardiac surgery as evaluated by the EuroScore II [6] will also be measured preoperatively along with the following quality of life indicators: the functional status (measured by EQ-5D, 5 domains), EQ5D VAS score, the disability status (measured by WHODAS 2.0) and the MacNew score of Heart Disease Health-Related Quality of Life.

2.6 Intraoperative characteristics

The type of surgery will be recorded and divided according to the following categories: CABG only, open valve surgery alone, TAVI alone, other surgery alone, CABG and Valve surgery*, CABG and Other surgery*, Valve and other surgery*, CABG, Valve and Other surgery* (TAVI being excluded for all categories indicated with *). Information on broader categories i.e. CABG, valve repair or replacement, TAVI, other will also be collected along with surgical LV grade (good (EF > 50%), moderate (EF 31-50%), poor (EF 21-30%), very poor (EF < 20%). Other standard characteristics related to surgery are also reported, revision surgery, number of distal anastomosis, total bypass time (min), minimum temperature during bypass (c°), IABP/ECMO/VAD post bypass, lowest haemoglobin concentration during CPB (g/L), inotrope infusion post bypass, cell salvage used during bypass, RBC and units given, platelets and number of units given, fresh frozen plasma (FFP) and number of units given. The total surgical time will be calculated from the first incision until the last suture is inserted.

2.7 Endpoints

2.7.1 Primary endpoint

The primary endpoint, DAH90, is a patient-centred outcome calculated using day and time of surgery as the starting point (Day 0) and calculated thereafter as all days spent at home by the patient until day 90. A few key elements have to be considered to derive this endpoint: 1) a patient who dies within 90 days of surgery is scored as zero irrespective of whether they had spent some of that time at home; 2) in case of hospital readmission after discharge and home return, time spent in hospital due to readmission(s) will be deducted in the DAH90 calculation (up to the limit of 90 days); 3) home is considered to be the patient's normal place of residence or that of a family member, but not a rehabilitation or nursing home facility. This implies that any stay in rehabilitation or nursing home facility within 90 days of surgery will not be classed as "home" in the DAH90 calculation. Examples of DAH90 consistent with this definition are given in the protocol [2

2.7.2 Secondary endpoints

1) change in haemoglobin concentration from day of enrolment to day of surgery (preoperative response),
2) Correction of iron deficiency state (Δferritin, ΔTSat) from treatment to day of surgery
3) Change in haemoglobin from day of surgery to 6 weeks post surgery
4) Red cell transfusion (any and units) during hospital stay, and post-discharge to 90 days (any).
5) Postoperative complications (i.e. MI/stroke/other CVE/infection /thromboembolism) within 30 days of surgery
6) 15-item quality of recovery scale on postoperative day 3
7) a) Length of stay in hospital (LOS). Hospital stay is calculated from the start (date, time) of surgery until actual hospital discharge or death (in hospital).
   b) ICU stay (hours): ICU stay includes the initial ICU admission and additional time following any ICU readmission up to 30 days after surgery, or death.
8) Days alive and at home up to 30 days after surgery (DAH30)
9) Disability-free survival, from day of surgery to 6 months after surgery, where disability-free is defined as a WHODAS score <16% at each of 30 days, 90 days and 6 months post surgery.
10) Mortality at 30 days, 90 days and 6 months after surgery
11) Hospital readmission up to 90 days after surgery
12) a) patient-reported quality of life using the MacNew Heart Disease Health-Related Quality of Life Questionnaire (MacNew) up to the day of surgery,
    b) EQ-5D postoperatively at 30 days, 90 days and 6 months after surgery
    c) World Health Organization Disability Assessment Schedule 2.0 (WHODAS) postoperatively at 30 days, 90 days and 6 months after surgery.

2.7.3 Safety endpoints (to day 30 post surgery)

1) Anaphylactoid reactions
2) Blood transfusion reactions
3) Cardiovascular events a) MI , b) Stroke)
4) Thromboembolism: as a composite and separated into pulmonary embolism, deep vein thrombosis, Other.
5) Infection – a) surgical site, b) deep sternal wound and c) other

6) Adverse events (AE), including the type of AE (neurological, respiratory, gastrointestinal, vascular, renal, skin, musculoskeletal, lymphatic oedema, other [including hypophosphataemia]) and severity (mild, moderate, severe).

2.7.4 Heath economics

As explained in the protocol the main question of interest is whether the inclusion of preoperative single dose IV iron in patients with anaemia is a more cost-effective care pathway for elective cardiac surgery than current practice.
The main outcome measure for the economic evaluation will be quality-adjusted life years (QALYs) calculated based on utility weights (estimated by EQ-5D-5L) and life years survived. Other endpoints include: costs related to health care including indexed hospitalization, post-surgical care (medical services, pharmaceutical use, hospital and ambulance service, rehabilitation centre, nursing home, allied health service, etc.). These methods and results will be detailed in a second publication.

2.8 Power and sample size calculation

ITACS is a pragmatic clinical trial aligned to normal hospital pathways. Some preliminary information was available [7] to help calibrate the distribution of DAH90 in the control arm. The endpoint has a left-skewed distribution with a small spike at 0, essentially due to early deaths. Therefore, the sample was determined by simulations (B=10,000 replications) to detect a specific shift in the median DAH90 with power $>$ 90% after accounting for loss to followup and non-compliance. Specifically, DAH90 was obtained by generating first hospital length of stay (LOS) using a lognormal distribution, truncating it at 86 days (allowing for a minimum of 4 days in hospital), and computing DAH90 as 86 minus LOS for 95% of the data on average. The remaining data (5% on average) was generated as 0 to represent a small spike as observed in the original data [7]. A sample size of n=1000 patients was enough to detect a change of 1.45 in the median days with 94% power using a Wilcoxon rank-sum test. This calculation is adjusted for 5% loss to FU and 10% cross-over. Power is lower for median regression and depends on the bootstrap method employed to calculate the 95% CI; for instance 84% is obtained using the Markov chain marginal bootstrap [8] available in the R package.

3 Statistical analysis

3.1 Analysis principles

- Analyses will be conducted on a modified intention-to-treat (mITT) basis i.e. patients will be analysed according to the treatment they have been originally assigned to provided data is available (treatment policy estimand).

- All tests are two-sided and the nominal level of type I error will be 5%.
- The primary analysis of the effect of treatment on the primary outcome and key secondary endpoints will be adjusted for the stratification factors of baseline haemoglobin and planned surgery, with region adjusted for as a proxy for site to avoid model convergence issues.
- Subgroup analyses will be carried out irrespective of whether there is a significant treatment effect on the primary outcome.
- No imputation of missing values will be carried out for a percentage of missing values less than 5%. When the proportion of missing is substantial (>5%), a sensitivity analysis for the primary endpoint based on inverse probability weighting will be carried out.
- Analyses will be conducted primarily using Stata or R software.

3.2 Interim analyses

A maximum of 3 analyses included the final analysis are planned with the two interim analyses conducted when 33% and 66% of the data have become available. The O'Brien and Fleming Type I error spending function was used to determine the boundaries at each look. The final last level of significance with 3 looks will consequently be 0.0463.

3.3 Subject Disposition

The flow of study participants will be presented in a Consolidated Standards of Reporting Trials (CONSORT) diagram.

3.4 Demographic and baseline variables

A description of the following baseline characteristics will be presented by treatment group. Discrete variables will be summarised by frequencies and percentages. Percentages will be calculated according to the number of patients for whom data is available. If missing values are ≥5%, the denominator will be added in a footnote in the corresponding summary table. Continuous variables will be summarised by use of standard measures of central tendency and dispersion, either mean and standard deviation (SD) or median and interquartile range (IQR) as appropriate.

3.5 Preoperative and intraoperative variables

The preoperative and intra-operative data will be summarised and presented by treatment arm using frequencies and percentages for categorical variables, mean and standard deviation for symmetrical continuous variables, or median and quartiles (25th and 75th percentile) for non-symmetrical continuous variables. If variables have more than 5% missing values, we will report the non-missing sample either in the table or in the footnote of the table. No *p*-values will be reported to compare the two arms with the exception of iron-related variables, i.e. Hb, ferritin, transferrin, and mean corpuscular volume

3.6 Primary outcome

For analysis of DAH90 we will use a Wilcoxon rank sum test, and quantile regression for adjusted treatment effect of the median DAH90 and confidence interval, with the latter considered as the decisive analysis in case of discrepancy. Confidence intervals for the IV iron effect on the median DAH90 or other relevant quantiles will be obtained via bootstrapping with a minimum of B=4000 replicates. The nominal level 95% will be slightly corrected, i.e. 95.4% CI will be used to account for the two interim analyses. Stratification factors will be included as covariates in the quantile regression models.

3.7 Secondary outcomes

Continuous secondary outcomes (#1, 2, 3, 6, 12) will be analysed using multivariable linear regression possibly after transformation. The analysis will be adjusted for the baseline measurement of the outcome where available. Red cell transfusion units (#4), DAH to 30 days (#8) and length of stay outcomes (#7a, 7b) will be summarised with medians and interquartile ranges and with estimated differences between medians and their 95% CIs computed via quantile regression, and for the latter (#7a,b) with death prior to discharge set as the largest duration. Binary outcomes (#4, 5, 9, 10, 11, and safety endpoints) will be expressed as risk ratios and standard 95% CI using log-binomial regression or exact logistic regression to approximate these values when the number of events in either arm is fewer than 5. In case of convergence issues with log-binomial regression, log-poisson regression with robust standard errors will be used. Where possible, adjustment for region, baseline haemoglobin and planned surgery similar to the adjusted analysis for the primary outcome will be carried out in all the models above to be consistent with the primary outcome analysis.

Supplementary analyses of the binary outcomes to 30 days will be performed to take into account the competing risk of death. This will use cumulative incidence functions

with estimated incidence at 30 days, with the incidence compared across treatment groups using risk ratios with 95% CIs derived using the delta method.

3.8 Handling of missing data

No attempt at weighting or imputing missing outcomes will be attempted if the proportion of missing data is below 5% and missing will be considered missing at random. A complete case analysis will be used as the primary analysis if the potential impact of the missing data is negligible (< 5%). For a proportion of missing data greater than 5%, inverse probability weighting (IPW) where complete cases are weighted by the inverse of their probability of being a complete case [9] will be carried out, with weights determined empirically from variables predictive of missingness. This analysis will be conducted for the primary outcome but will not be systematically undertaken for other endpoints.Patients missing one item in the WHODAS (or EQ-5D) scales at any time point will have the value of the missing item imputed as the mean of the other items. Patients missing two or more items will have their WHODAS (or EQ-5D) set to missing. Patients who died before WHODAS (or EQ-5D) can be measured will have the corresponding endpoint set to missing.

3.9 Safety endpoints

The type of AEs and AE severity will be tabulated by treatment arm. The number and percentage of patients with at least one AE for each type will be presented. Severity will also be presented accordingly.

3.10 Subgroup analyses

Pre-specified sub-group analyses will be conducted to assess the effect of IV iron on DAH90 by patient subgroups as indicated below:

1) age category (approximate quartiles)
2) sex (male/female)
3) surgery type (CABG/valve/TAVI/other)
4) iron deficiency (yes/no) where iron deficiency is defined as ferritin (<30 mg/dL, or ferritin <100 mg/dL and C-reactive protein >5 mg/L and/or transferrin saturation <20%)
5) haemoglobin strata (≤100, 101-120, ≥120 g/L)
6) study drug formulation (carboxymaltose / isomaltoside)
7) Time from randomisation to operation (<4 weeks or ≥4 weeks).

Stratified effect estimates may be presented but the difference across subgroup will be tested via an interaction test in the corresponding median regression model for the primary endpoint. Forest plots will be constructed to illustrate results of the subgroup analyses.

4 References

[1] Myles PS, Shulman M, Heritier S, et al. Validation of days at home as an outcome measure after surgery: a prospective cohort study in Australia. BMJ Open 2017 doi:10.1136/bmjopen-2017-015828.

[2] Myles PS, Richards T, Klein A, Smith J, Wood E, Heritier S et al. Rationale and design of the intravenous iron for treatment of anemia before cardiac surgery (ITACS) trial. Am Heart J 2021, 239:64-72.

[3] [EQ-5D-5L - EQ-5D [Internet]. Euroqol.org. 2021 [cited 18 March 2021]. Available from: https://euroqol.org/eq-5d-instruments/eq-5d-5l-about/

[4] WHO Disability Assessment Schedule (WHODAS 2.0) [Internet]. World Health Organization. 2021 [cited 18 March 2021]. Available from: https://www.who.int/classifications/international-classification-of-functioning-disability-and-health/who-disability-assessment-schedule

[5] Valenti L, Lim L, Heller RF, Knapp J. An improved questionnaire for assessing quality of life after acute myocardial infarction. Qual Life Res 1996;5:151-61.

[6] Nashef SA, Roques F, Sharples LD, et al. EuroSCORE II. Eur J Cardiothorac Surg 2012;41:734-44 discussion 744-5.

[7] Myles PS, Smith JA, Forbes A, Silbert B, Jayarajah M, Painter T, et al. Tranexamic acid in patients undergoing coronary-artery surgery. N Engl J Med 2017;376(2):136-148.

[8] He X and Hu F. Markov chain marginal bootstrap. J. Amer. Statist. Assoc. 2002; 97(459), 783-795.

[9] Seaman SR, White IR. Review of inverse probability weighting for dealing with missing data. Stat Methods Med Res. 2013 Jun;22(3):278-95. doi: 10.1177/0962280210395740. Epub 2011 Jan 10. PMID: 21220355.