Determining potential pathological mechanisms associated with fluid resuscitation with albumin in patients with traumatic brain injury: statistical analysis plan

Stephane Heritier[1], John A Myburgh MBBCh PhD FANZCA FCICM[2] and D James Cooper MBBS MD FRACP FCICM[3] on behalf of the SAFE-TBI Study Management Committee.

1. Head of Statistical Research, George Institute for Global Health
2. Director, Division of Critical Care and Trauma, George Institute for Global Health
3. Director ANZIC Research Centre, School of Public Health, Monash University

Corresponding authors:
Prof John A Myburgh
Email: jmyburgh@george.org.au
Ph: +61 2 96570 348
Fax: +61 2 96570 301

Abstract

Background: A post-hoc analysis of Saline versus Albumin Fluid Evaluation (SAFE) study identified that 2-year mortality was higher in patients with traumatic brain injury (TBI) resuscitated with albumin (SAFE-TBI study). Additional data collection and analyses of the TBI patients were conducted to explore potential pathological mechanisms, primarily the development of raised intracranial pressure (ICP).

Objective: To describe the pre-specified statistical analysis (SAP) plan of this subsidiary analysis finalised prior to unblinding treatment allocation and to which the investigators adhered during data analysis.

Methods: The primary analyses on the effect of treatment allocation on ICP will be unadjusted. Secondary analyses will be adjusted for same covariates used in SAFE-TBI. Challenges in the interpretation of repeated measurements over time influenced by informative dropout was recognised *a priori* and will be addressed by conducting a hierarchy of analyses categorised by descriptive statistics, average profile plots, repeated measures analyses based on mixed models and pattern mixture models primarily in the population of patients having a monitor inserted. Progression of computerised tomographic scores will be analysed by logistic or ordinal regression. Cumulative therapeutic interventions to determine a therapeutic intensity score will be determined from sensitivity analyses.

Results: A comprehensive SAP was developed for additional analyses of the SAFE-TBI study. This plan provides a pre-determined framework for the complex modelling

techniques presented before the effect of treatment on selected primary and secondary outcomes were revealed.

Conclusion: We have developed a pre-determined SAP for an additional post hoc analysis of the SAFE-TBI II study. This plan will be followed to avoid potential bias arising from prior knowledge of the results.

Key words: intracranial pressure, SAFE-TBI, post hoc analysis, hierarchy of analyses, pattern mixture models

Introduction

A post-hoc analysis of 460 patients with traumatic brain injury (TBI) randomised into the Saline vs Albumin Fluid Evaluation (SAFE) study[1] demonstrated that 2-year mortality rates were significantly increased in patients who received 4% albumin for intravenous resuscitation in the Intensive Care Unit (ICU) compared to 0.9% saline (relative risk of death [RR] 1.63, 95% confidence intervals [CI] 1.20 to 2.38, p=0.003).(SAFE-TBI)[2] Despite these compelling data to guide the choice of resuscitation fluid in patients with TBI, the biological mechanisms for the observed differences in mortality are unclear.

The aim of this subsequent post-hoc analysis of SAFE-TBI was to determine whether potential biological mechanisms by which albumin was associated with increased mortality could be determined. Specifically, this includes the development of raised intracranial pressure (ICP) due to intrinsic pathophysiological processes and/or associated therapies that may be used to control raised ICP that may be toxic independently or within a tiered therapeutic intensity approach.

The statistical challenges of conducting this additional analysis were recognised at from the outset during the design of the study. Mixed-effects models, which incorporate repeated measurements of a particular continuous endpoint (here ICP) over time in the same patients is a well established method for studying the relationship between treatment and the endpoint over time. However, given the strong association between TBI and mortality for critically ill patients, the occurrence of death could result in what is called "*informative censoring*". In addition, ICP may stop being recorded for various reasons possibly linked to unobserved outcome values (e.g. the patient is in such poor state that the clinical team stopped monitoring ICP). This situation is known as "*informative dropout*". In both cases we have informative missing data for the longitudinal process, a violation of the underlying assumptions of the mixed linear models that potentially lead to biased results. Rubin developed a typology of missing data problems, based on a model for the probability of an observation being missing[3]. Data are described as missing completely at random (MCAR) if the probability that a particular observation is missing does not depend on the value of any observable variable(s). Data are missing at random (MAR) if, given the observed data, the probability that observations are missing is independent of the actual values of the missing data. The reasons for dropout are usually varied. For example, in 11 clinical trials of similar design, considered by Mallinckrodt *et al.* (2003)[4] with the same drug and involving patients with the same disease state, the rate of and the reasons for dropout varied considerably. There are no universally applicable methods for handling missing data that are missing non-at-random (NMAR) although sensitivity analyses and simulations can be used to evaluate their performance. The main problem is that they need somehow to accommodate dropout in the modelling process. Resulting inferential procedures will often depend on implicit and untestable assumptions regarding the distribution of the unobserved outcome measurements (here ICP or other concomitant therapies) given the observed measurements and one can wonder to what extent the lack of information due to incompleteness in the data can be recovered. In the class of NMAR methods that accounts for informative dropout, *pattern-mixture* models or PMM[5] are appealing as they condition on the dropout pattern and try to extrapolate beyond the last measurement time.

Other challenges exist in the analysis of the SAFE-TBI data, They include the inclusion of post-randomisation variables, testing several endpoints over time (multiplicity), accounting for intermittent MAR data and subgroup analyses. To address these, we present an *a priori* statistical analysis plan that was written by an independent statistician not previously involved in the SAFE or SAFE-TBI studies in combination with the principal investigators with no access to the follow-up data by treatment arm.

Decisions related to the research question(s) were made in a blinded fashion until the statistical analysis plan was finalised.

Study overview

This study was an additional analysis of the SAFE-TBI database that contains 460 patients randomised into the SAFE study (n=6997) with an admission diagnosis of trauma and an associated traumatic brain injury (defined as a post-resuscitation Glasgow Coma Score [GCS] <14, plus an abnormal cranial computerised tomography [CT] scan consistent with TBI). Of these patients, 231 (50.2%) received albumin and 229 (49.8%) received saline for fluid resuscitation in the ICU.
In addition, post-randomisation data from baseline to 14 days post-randomisation were obtained retrospectively from patient records.

Study outcomes

The *primary outcome* for this study was the effect of fluid resuscitation on ICP as determined by end-hourly measurements for the duration of ICP monitoring until 14 days post randomisation.

Secondary outcomes were classified into intrinsic and extrinsic variables:

1. Intrinsic variables relate primarily to intracranial patho-physiological processes that were considered to impact directly on raised intracranial pressure:
    a. Changes in intracranial volume represented by the duration and daily volume of cerebrospinal fluid drainage via an external ventricular drain used concomitantly for ICP monitoring
    b. Coagulopathic processes represented by the highest daily activated partial thromboplastin time (APTT) and international normalised ratio (INR) and lowest daily platelet count
    c. Alterations in brain swelling represented by a 3-point score (stability, progression, regression[6]) in CT scores for diffuse axonal injury from baseline to the first CT post randomisation.

2. Extrinsic variables relate to administered therapies directed at preventing or treating intracranial hypertension, defined as an ICP>20mmHg:
    a. Haemodynamic augmentation with catecholamines (noradrenaline, adrenaline or dopamine) and/or vasopressin presented by total daily dosages.
    b. Suppression of cerebral metabolism with barbiturates (thiopentone with and without pentobarbitone) presented by total daily dosages
    c. Sedation and analgesia with intravenous opiates (morphine, fentanyl), sedatives (propofol, benzodiazepines) presented by total daily dosages
    d. Hypothermia presented by the average daily temperature

e. Osmotherapy with mannitol or hypertonic saline and the time-weighted average daily serum sodium
f. Hyperventilation defined as PACO2 < 28 and the average daily arterial carbon dioxide tension ($PaCO_2$).

Two *tertiary outcomes* were defined:

1. The determination of a composite therapeutic intensity level (TIL) according to a hierarchy described by Marmarou et al [7] from selected secondary extrinsic binomial variables
2. Quantification of global secondary ischaemic/hypoxic brain insults in the generation of intracranial hypertension presented by daily mean arterial pressure and oxygen tension (PaO_2).

Study design

Analysis principles

Analyses would be conducted using primarily SAS or R software on an intention-to-treat basis and unadjusted, except where indicated.

All tests are two-sided with nominal level of α=5%.

Simple analyses would be conducted first followed by analyses of increasing complexity. A hierarchy of analysis beginning with the most plausible mechanisms and variables of interest will be determined.

Missing values will not be imputed unless specified otherwise and p values will not be adjusted for multiplicity.

A blind review of data will be conducted before finalising the statistical analysis plan to facilitate the selection and feasibility of the statistical methods.

Study populations

Three post-randomisation patient populations were considered:

1. Patients with ICP monitors: *n*=321 (dataset 1)
2. Patients with severe TBI defined as a post resuscitation GCS < 9; n=317 (dataset 2)
3. Patients presented in SAFE-TBI[2]; n=460 (dataset 3)

Given that the primary outcome changes in ICP, dataset 1 would be the principal population of interest.

Patient characteristics and comparisons at baseline

These are presented in Table 1.
Discrete variables will be summarised by frequencies and percentages.
Percentages will be calculated according to the number of patients with available data. Where values are missing, the denominator will be added with a footnote in the corresponding summary table. Continuous variables will be summarised using either mean ± SD or median and interquartile range (IQR). Durations will also be summarised by medians and IQR.

Daily measurements and date issues

The concept of 'short day 1' has been consistently used for daily measurements in the SAFE study. Day 1 can be any duration from one hour to 24 hours.

As a result, some inconsistencies are found between the survival time and the day the last measurement for a specific endpoint was taken. For example, ICP may be recorded one day after the patient died which can be corrected by systematically adding 1 to the survival time.

Primary outcome

The following hierarchy of analyses will apply to the primary outcome:

1. Descriptive analysis (Figure 2):
    a. These include box plot or bar plot per study day, overall average plots for daily means and average profile plots for daily means per pattern.
    b. Similar plots by treatment arm will be produced once the models described in pattern mixture model below have been completed and the statistical analysis plan finalised.
    c. Average profile plots will help determine if dropout patterns exists and should be modelled.

2. Repeated measures analysis based on mixed models valid under the MAR assumption (Table 2a). As stated before MAR means that, conditionally on the observed data, missingness is independent of the unobserved measurements (i.e. the missingness process can be ignored in the analysis):
    a. This is considered as the reference technique even though the MAR assumption might not be satisfied[8].
    b. The overall difference between treatment groups will be tested using the study day as a factor in the model with possible interaction with treatment.
    c. No difference will be tested for each day individually. However, two pre-defined days of interest will be tested (*i.e.* day 3 and 7), based on probable clinical changes in ICP.
    d. For random effects, a random intercept will be specified and an autoregressive AR(1) structure will be used for the error term. Such a correlation structure is supposed to capture the time-dependence in the data. In case of convergence problems, a model with independent errors will be fitted to avoid further computational issues.

3. Pattern mixture models or PMM[9]:
   This step will be completed if there is evidence from the data that drop-out or censoring (possibly due to death) is not missing at random for a particular endpoint. In a standard PMM, data are modelled according to some pattern identified through profile plots drawn per dropout category. Our primary option was to consider a week as an appropriate cut-off (i.e. dropout is classified according to whether the last day of ICP measurement is in week 1 (day 0-7) or in week 2 (day 8-14)). Preliminary plots done in a blinded fashion showed that either a linear or a quadratic trend was an appropriate pattern. The PMM is otherwise based on a mixed linear model with random intercept and slope to allow each patient to have their own line. Results for ICP will be reported as displayed in Table 2b. Other PMMs based on survival categories (day 0-7, day 8-28 censored at 28) or a combination of dead/alive at day 28 versus dropout

week is also allowed in a second step. These models are more delicate to interpret as the conditioning depends on survival that can itself be affected by treatment[10]. Irrespective of the findings, limitations of this type of modelling will be acknowledged in future publications.

The following analysis principles will apply to the primary outcome:

1. The primary analysis will be unadjusted and performed on dataset 1 (patients with ICP monitors). This means that only the study day, treatment and possibly the interaction study day by treatment will be included in the model.
2. No imputation for missing ICP data will be carried out for patients in dataset 1 dying on the first day (approximately 8%). To adjust for intermittently missing ICP data, a simple form of imputation may be carried out as a sensitivity analysis.
3. A secondary analysis will be adjusted for the same covariates as described in SAFE-TBI (age > 60 years, post resuscitation GCS ≤ 8, pre-randomisation systolic blood pressure < 90mmHg and CT evidence of traumatic subarachnoid hemorrhage).
4. Adjustment for potential significant imbalances at baseline will be conducted.
5. The analysis will be repeated on dataset 2 (patients with severe TBI) or at least in those patients of dataset 2 who underwent intracranial pressure monitoring.

Secondary outcomes

A similar strategy to that described for the primary outcome will be used for intrinsic and extrinsic variables with the following restrictions/modifications:

1. Repeated measures analysis based on mixed models will be conducted on all variables but will be based on nonlinear mixed models for binary/ordinal outcomes (Table 3a)
2. If a particular endpoint depends on study day through a functional form (e.g. a linear or quadratic trend), this form will be introduced in the PMM as a sensitivity analysis. In the absence of a clear relationship, the time effect will be treated as a factor to avoid uncertain parametric specifications.
3. PMMs will be performed on all continuous variables unless no clear pattern emerges from the plots. Similar technical conditions apply. Table 3b specifies how the results will be reported for each endpoint; a similar format to the one used for ICP (Table 2b) will be used
4. PMMs will generally not be performed on binary indicators as data are generally too flimsy to identify any kind of pattern. If large enough numbers are available this restriction will be waived.
5. Results of PMMs will normally be contrasted with those of repeated measures analysis based on a MAR assumption (as defined earlier).
   In addition, other PMMs may also be fitted in a similar fashion to what is suggested for ICP.
6. Changes in CT scan scores (Table 4):
   a. Prior to the analysis of changes in CT scan scores from baseline to the first scan post randomisation, 25 randomly selected scans will be scored by two independent assessors. Agreement in three related

groups within the Marshal score (diffuse axonal injury 2-4, evacuated or non-evacuated mass lesions and presence of traumatic subarachnoid haemorrhage) will be assessed with a kappa statistic and if insufficient ($\kappa<0.5$), an independent neuroradiologist will score all CT scans for analysis.

b. In case of poor agreement, CT scans will be rated independently by an independent radiologist after study completion. As some of the new ratings could be classified as DI, one of the eligibility criteria would 'de facto' be violated. Such patients will be reclassified as DII at baseline.
c. As there is no progressive relationship between scores for diffuse axonal injury and mass lesions, only changes for diffuse axonal injury will be used to determine temporal changes in brain swelling.
d. Three-level outcome (stability, regression or progression) will be modelled by ordinal logistic regression[11] or polytomous logistic regression if the proportional odds assumption is not met.
e. The analysis will be based on *patients having complete data* with *the exclusion of patients who cannot progress (n=260)* as recommended in reference [6]. The primary analysis for these two endpoints will be adjusted for the baseline measurement.
f. No imputation will be carried out for the CT scans missing either at baseline or post-randomisation.

The following analysis principles will apply to the secondary outcomes (apart from CT scans):

The primary analysis will be unadjusted for all endpoints.

1. Using these pattern mixture models, the effect of treatment will be tested in dataset 1 (patients with ICP monitoring). As all this modelling is viewed as a sensitivity analysis, tests to be carried out cannot be completely specified.
2. Whether the observed patterns are similar across treatment arms either for all patterns or for a specific pattern of interest will be considered.
3. A secondary analysis will be adjusted for the same factors as the primary endpoint. In a second step, a similar analysis will be carried out on patients of dataset 2 and dataset 3 having relevant data.

Tertiary outcomes

During the SAFE study, all aspects of patient treatment were left to the discretion of the attending clinician, including strategies and therapies for intracranial hypertension. As there is no standardised approached for a tiered therapeutic intensity level, individual components of therapies with the tiered therapeutic intensity level will be analysed separately.

Additional issues

The following analysis principles will apply to the secondary outcomes:

1. Missing data:
    a. The exact number of patients involved in the analysis of each extrinsic component will be given.

   b. There will be no imputation for missing data, although a sensitivity analysis may be considered.

2. Treatment of outliers, "zeroes" and unreliable total daily dosages
   a. Wide outliers due to transcription errors will be deleted prior to the analysis and no imputation made.
   b. Where >50% total daily dosages are recorded as zero (eg for thiopentone), modelling as a continuous outcome will not be possible and will be treated as a binary endpoint.

References:

[1] The SAFE study investigators (2004) *A Comparison of Albumin and Saline for Fluid Resuscitation in the Intensive Care Unit.* N Engl J Med, **350**:2247-2256

[2] The SAFE study investigators (2007) "*Saline or Albumin for Fluid Resuscitation in Patients with Traumatic Brain Injury*", New England Journal of Medicine, **357**:874-84.

[3] Little RJA and Rubin DB (2002). Statistical Analysis With Missing Data, 2nd Edition, J. Wiley & Sons Ltd: Chichester.

[4] Mallinckrodt CH, Clark W, Carroll RJ, Molenberghs G. (2003) *Assessing response profiles from incomplete longitudinal clinical trial data under regulatory considerations*. Journal of Biopharmaceutical Statistics **13**, 179–190.

[5] Little RJA. (1993). *Pattern-mixture models for multivariate incomplete data.* Journal of the American Statistical Association **88**, 125–134.

[6] Chieregato A, Fainardi E, Morselli-Labate AM, Antonelli V, Compagnone C, Targa L, Kraus J, Servadei F (2005) *Factors associated with neurological outcome and lesion progression in traumatic subarachnoid hemorrhage patients,* Neurosurgery, **56**(4):671-80

[7] Marmarou A, Anderson RL, Ward JD et al. (1991). "*NINDS Traumatic Coma Data Bank: intracranial pressure monitoring methodology*", Journal of Neurosurgery, 75 (1991), pp. S21–27

[8] Molenberhgs G, Thijs H, Jansen J, Beunckens C, Kenward MG, Mallinckrodt C and RJ Carroll (2004). "*Analyzing incomplete longitudinal clinical trial data"*, Biostatistics **5**:445-464

[9] Verbeke G. and G. Molenberghs (2000). "*Linear Mixed models for longitudinal data*", Springer: New York.

[10] Pauler DK, McCoy S and Moinpour C (2003). "*Pattern mixture models for longitudinal quality of life studies in advanced-stage disease*", Statistics in Medicine 2003;**22**:795-809.

[11] Ananth, CV and Kleinbaum, DG (1997). *Regression models for ordinal responses: A review of methods and applications.* International Journal of Epidemiology **26**:1323-1333

List of tables

Table 1: Baseline characteristics

Variable	Albumin	Saline	p
Age - years (mean ± SD)			
Age >55y (mean ± SD)			
Male sex			
APACHE II			
AIS score			
Mean arterial pressure			
Heart rate			
Central venous pressure			
Serum albumin			
Glasgow Coma Score			
GCS (median IQR)			
GCS 3-8			
GCS 9-12			
GCS 13			
GCS motor			
CT Scan Score			
DAI II			
DAI III			
DAI IV			
NEML			
EML			
Traumatic subarachnoid haemorrhage			
Intracranial pressure insertion			

Table 2: Primary outcome measure (intracranial pressure)

2a: Missing at random (MAR) model

Days	Albumin	Saline	p
3			
7			
14			
All days 1-14			

Data are presented as adjusted means (± standard error)

Table 2b: Pattern mixture models

Random effect	Pattern	Albumin	Saline	p
Intercept	week1			
	week2			
	Global			
Slope	week1			
	week2			
	Global			

Table 3a: Missing at random (MAR) model for secondary outcomes

Outcome	Days	Albumin	Saline	p
Intrinsic variables				
APTT	3			
	7			
	14			
	1 - 14			
INR	3			
	7			
	14			
	1 to 14			
CSF Drainage	3			
	7			
	14			
	1 to 14			
Extrinsic variables				
TDD morphine *	3			
	7			
	14			
	1 to 14			
TDD noradrenaline *	3			
	7			
	14			
	1 to 14			
TDD propofol *	3			
	7			
	14			
	1 to 14			
TDD midazolam *	3			
	7			
	14			
	1 to 14			
Temperature	3			
	7			
	14			
	1 to 14			
Sodium	3			
	7			
	14			
	1 to 14			
$PaCO_2$	3			
	7			
	14			
	1 to 14			

* transformed by log (x+1)

Table 3b: Pattern mixture models for secondary outcomes

Parameter	Random effect**	Pattern	Albumin	Saline	p
Intrinsic outcomes					
APTT	intercept	week1			
		week2			
		Global			
	Slope	week1			
		week2			
		Global			
INR	intercept	week1			
		week2			
		Global			
	Slope	week1			
		week2			
		Global			
CSF Drainage	intercept	week1			
		week2			
		Global			
	Slope	week1			
		week2			
		Global			
	Quadratic	week1			
		week2			
		Global			
Extrinsic outcomes					
TDD morphine*	intercept	week1			
		week2			
		Global			
	Slope	week1			
		week2			
		Global			
TDD noradrenaline*	intercept	week1			
		week2			
		Global			
	Slope	week1			
		week2			
		Global			
TDD propofol*	Intercept	week1			
		week2			
		Global			
	Slope	week1			
		week2			
		Global			
TDD midazolam*	Intercept	week1			

		week2			
		Global			
	Slope	week1			
		week2			
		Global			
Temperature	Intercept	week1			
		week2			
		Global			
	Slope	week1			
		week2			
		Global			
Sodium	Intercept	week1			
		week2			
		Global			
	Slope	week1			
		week2			
		Global			
	Quadratic	week1			
		week2			
		Global			
$PaCO_2$	Intercept	week1			
		week2			
		Global			
	Slope	week1			
		week2			
		Global			
	Quadratic	week1			
		week2			
		Global			

* transformed by log (x+1)

ICP = intracranial pressure, TDD = total daily dose, MAP = mean arterial pressure, PMM = pattern mixture model

* transformed by f(y)=log(y+1)

** time to be removed if numerical problems observed

Table 4: CT scans

		Albumin	Saline	OR (95% CI)	p-value
Clinical progression	yes (n, %)				
	no				
3-level change	progression				
	stability				
	Regression				

List of Figures