Renal Replacement Therapy Dependence at Day 28 in Critically Ill Patients with Acute Kidney Injury Receiving Continuous or Intermittent Renal Replacement Therapy

Statistical Analysis Plan

Ary Serpa Neto,[1,2] Thummaporn Naorungroj,[2,3] Rinaldo Bellomo[1,2]

1. Australian and New Zealand Intensive Care Research Centre, School of Public Health and Preventive Medicine, Monash University, Melbourne, Victoria, Australia.

2. Department of Intensive Care, Austin Hospital, Melbourne, Australia.

3. Department of Intensive Care, Siriraj Hospital, Mahidol University, Bangkok, Thailand.

Corresponding author:

Rinaldo Bellomo

Department of Intensive Care, Austin Hospital,

145 Studley Rd, Heidelberg, Victoria, Australia

Tel: +61-3-9496 5992

Fax: + 61-3-9496 3932

Email: rinaldo.bellomo@austin.org

INTRODUCTION

The present study outlines the protocol and statistical analysis plan for a pooled analysis of two randomized controlled trial aiming to compare the renal replacement therapy (RRT) dependence at day 28 in critically ill patients with acute kidney injury (AKI) receiving continuous renal replacement therapy (CRRT) versus patients treated with intermittent hemodialysis (IHD). We hypothesize that critically ill patients receiving CRRT and alive at day 28 have a lower incidence of RRT dependence at day 28 compared with critically ill patients treated with IHD.

METHODS

Study design

This is a pooled analysis of individual patient data from RENAL[1] and ATN[2] trial. The RENAL trial was a multicenter randomized clinical trial comparing the efficacy of two different intensities of CRRT (40 mL/kg/h of effluent flow vs. 25 mL/kg/h of effluent flow) in critically ill patients with AKI. The study was conducted in 35 ICU in Australia and New Zealand from December, 2005 to November, 2008, and did not find difference among the treatment arms.[1] The ATN trial was a multicenter randomized clinical trial comparing the efficacy of two different intensities of RRT (35 mL/kg/h of effluent flow vs. 20 mL/kg/h of effluent flow) in critically ill patients with acute kidney injury. The study was conducted in 27 ICUs in USA from November, 2003 to July, 2007, and did not find difference among the treatment arms.[2]

Patients

All patients included in the original trials will be considered for inclusion in the present pooled analysis. Briefly, patients were eligible to participate in the RENAL trial if they were critically ill adults with AKI, were deemed to require CRRT by the physician in charge, and fulfilled at least one predefined inclusion criteria, including severe organ edema, oliguria, hyperkalemia, uremia, and/or severe metabolic acidosis.[1] In ATN trial, critically ill adults with AKI, deemed to require RRT, and with at least one nonrenal organ failure or sepsis were included.[2] For this secondary analysis, only survivors at day 60 will be included. The only exclusion criterion in the present analysis will be the absence of information on the modality of RRT (CRRT or IHD) implemented after randomization.

Data collection and definitions

Data will be obtained entirely from the existing datasets associated with the completed participating studies.[1,2] Patients will be classified in the CRRT or IHD group according to the modality of RRT received immediately after randomization. When not available, Acute Physiology and Chronic Health Evaluation (APACHE) III will be derived from APACHE II using pre-defined formulas.[3]

Outcomes

The primary outcome of the study will be RRT dependence at day 28, defined as patients still receiving RRT (independent of the mode) at day 28 after randomization. Secondary outcomes will include RRT dependence at day 60, ICU and hospital length of stay, and ICU, hospital mortality.

Statistical analysis

Categorical variables will be presented as numbers and percentages and compared using Fisher exact tests. Continuous variables will be presented as median (quartile 25% - quartile 75%), and compared using Wilcoxon rank-sum test. Rate of missing data will be reported.

Since it was expected that the two groups assessed will be different in several baseline characteristics, especially in disease severity, all analyses described below will be adjusted for a baseline risk model. To build the baseline risk model, first clinically relevant variables at baseline will be compared between patients who were dependent or not at day 28. Those variables with a $p < 0.05$ in this comparison will be selected for inclusion in the final models. If a variable is considered really important from a clinical perspective in a pre-analysis assessment, this variable will be included even in absence of a statistically significant difference between the groups. At the end, the following variables will be considered potential covariates for the final models, according to the criteria described above: age, gender, weight, type of admission (medical, surgical or other), APACHE III, hours between randomization and therapy, use of mechanical ventilation, presence of oliguria, presence of acidemia, presence of hyperkalemia, presence of sepsis, last urea before randomization, last creatinine before randomization, premorbid estimated glomerular filtration rate and intensity of treatment (as allocated in the original trials). The use of vasopressor will not be included because almost all patients receiving vasopressors were treated with CRRT in the original trials, potentially creating a bias in the analysis.

The primary outcome, RRT dependence at day 28, will be assessed considering a generalized linear model with binomial distribution, and presented as odds ratio and 95% confidence interval (CI). The same strategy will be used for the RRT dependence at day 60, ICU and hospital mortality. RRT-free days at day 28 will be compared using cumulative logistic models, and presented as common odds ratio (OR) with 95% CI. In this model, a common OR > 1.00 indicates a greater chance of having more RRT-free days at day 28 (better outcome). ICU and hospital length of stay will be compared between the groups with a Fine-Gray competing-risk model, with death before the event of interest treated as a competing risk. Results will be shown as sub-distribution hazard ratios (SHR) with their 95% CI. In the competing risk models, subjects who died remain in the risk set and the SHR denotes the instantaneous rate of the event of interest in subjects who are still alive (i.e., who have not yet experienced either death or the event of interest) and also those who have died.

In sensitivity analysis, to assess the consistency of findings, a covariate-balancing propensity score (CBPS) will be used to control for observed confounding factors that may influence outcomes.[4] The CBPS will be estimate for each patient with logistic regression using the variables described above and adding the use of vasopressor to the model. Patients with missing data will be excluded from this analysis. Based on the CBPS weighted estimators for the data, a CBPS-matched cohort will be constructed using nearest neighbor matching without replacement, with each patient from the IHD group matched to N patients of the CRRT group (1:N match). This strategy will be use due to the big difference in the number of patients expected in each group, and the N will be determined according to the data. Also, to decrease bias with the 1:N matching strategy, a variable ratio matching will be used.[5,6] A caliper width of 0.01 of the standard deviation of the logit of the CBPS will be used for the development of matching.[7,8] Matching performance will be assessed comparing the standardized mean difference of the variables between the groups (aiming to a value < 10%) and graphically shown in Love plots.

As additional sensitivity analyses for the primary outcome, the following additional analyses will be performed: 1) weighted regression by the inverse probability of treatment weighting; and 2) weighted regression by stabilized inverse probability of treatment weighting.[9] The calculation of the weighted regressions and more information on the models are shown below.

To confirm the findings in different cohorts, the primary and secondary outcomes will be re-assessed in the following cohorts: 1) patients receiving CRRT or IHD exclusively in the first three days of follow-up (patients who died or do not have information of modality in the first three days will be excluded from this analysis); 2) patients receiving CRRT or IHD exclusively during the whole follow-up available for each patient (including those who died early); 3) patients with cardiovascular SOFA ≤ 2 at baseline; and 4) restricting to patients in RENAL who would be randomized in ATN (sepsis OR at least one non-renal SOFA component ≥ 2 at baseline).

For all the models described, variables with less than 3% of missing will be imputed by the median value of the overall cohort. Missing data in premorbid estimated glomerular filtration rate will be imputed based on creatinine levels according to age and gender.[10] All analyses will be conducted in R v.4.0.2 (R Foundation, Vienna, Austria), and significance level will be set at 0.05.

Weighted regressions

Inverse probability of treatment weighting

In this method, the treatment effect is estimated in a population whose distribution of risk factors is equal to that found in all study subjects. The calculation of the IPTW was done according to Eq. 1:[11]

$$IPTW_T = \frac{1}{\hat{e}(X)} \; ; \; IPTW_{UN} = \frac{1}{1 - \hat{e}(X)} \quad (Eq. 1)$$

Where $IPTW_T$ is the IPTW for high CRRT group patients, $IPTW_{UN}$ is the IPTW for IHD group patients and ê(X) is the covariate balancing propensity score.

Stabilized inverse probability of treatment weighted

It may happen that treated subjects have a covariate balance propensity score near 0 or that untreated subjects have a covariate balancing propensity score near 1, making the relative IPTW excessively high and unstable. Computationally, as in any weighted regression, unstabilized IPTW changes the sample size of the original sample, generating an underestimate of the variance of the estimate of the effect, producing inappropriately narrow confidence intervals and leading to the lack of control of the probability of a type I error.[12]

Stabilized inverse probability of treatment weight (SIPTW) can be obtained by multiplying the IPTW by the marginal probability of receiving the actual treatment received. Moreover, it preserves the sample size of the original data, produces appropriate estimation of the variance of the main effect, and adequately controls the type I error rate. The calculation of the IPTW was done according to Eq. 2:[12]

$$SIPTW_T = \frac{p}{\hat{e}(X)} \; ; \; SIPTW_{UN} = \frac{1-p}{1-\hat{e}(X)} \quad (Eq. 2)$$

Where $SIPTW_T$ is the SIPTW for CRRT group patients, $SIPTW_{UN}$ is the SIPTW for IHD group patients, *p* is the probability of treatment without considering covariates (defined as n_t / N) and ê(X) is the covariate balancing propensity score.

REFERENCES

1. RENAL Replacement Therapy Study Investigators. Intensity of continuous renal-replacement therapy in critically ill patients. N Engl J Med 2009; 361:1627–38.

2. VA/NIH Acute Renal Failure Trial Network, Palevsky PMZJ, O'Connor TZ et al. Intensity of renal support in critically ill patients with acute kidney injury. N Engl J Med 2008; 359:7–20.

3. Schneider AG, Lipcsey M, Bailey M, Pilcher DV, Bellomo R. Simple translational equations to compare illness severity scores in intensive care trials. J Crit Care 2013; 28:885.e1-885.e8.

4. Imai K, Ratkovic M. Covariate balancing propensity score. J R Statist Soc B 2014; 76:243–63.

5. Rassen JA, Shelat AA, Myers J, Glynn RJ, Rothman KJ, Schneeweiss S. Optimal approaches to one-to-many propensity score matching in cohort studies. Pharmacoepidemiol Drug Saf 2012; 21(Supp 3):S34-5.

6. Ming K, Rosenbaum PR. Substantial gains in bias reduction from matching with a variable number of controls. Biometrics 2000 Mar; 56:118-24.

7. Austin PC. Optimal Caliper Widths for Propensity-Score Matching When Estimating Differences in Means and Differences in Proportions in Observational Studies. Pharm Stat 2011; 10:150-61.

8. Austin PC. Some Methods of Propensity-Score Matching had Superior Performance to Others: Results of an Empirical Investigation and Monte Carlo simulations. Biom J 2009; 51:171-84.

9. Kurth T, Walker AM, Glynn RJ, et al. Results of Multivariable Logistic Regression, Propensity Matching, Propensity Adjustment, and Propensity-based

Weighting under Conditions of Nonuniform Effect. Am J Epidemiol 2006; 163:262-70.

10. Tiao JYH, et al. The effect of age on serum creatinine levels in an aging population: relevance to vascular surgery. Cardiovasc Surg 2002; 10:445-51.

11. Xu S, Ross C, Raebel MA, Shetterly S, Blanchette C, Smith D. Use of stabilized inverse propensity score as weight to directly estimate relative risk and its confidence intervals. Value Health 2010; 2:273–7.

12. Kurth T, Walker AM, Glynn RJ, et al. Results of Multivariable Logistic Regression, Propensity Matching, Propensity Adjustment, and Propensity-based Weighting under Conditions of Nonuniform Effect. Am J Epidemiol 2006; 163:262-70.