Treatment of severe COVID-19 patients with convalescent sera: a multicentric, randomized, controlled, and open-label study

(COOP-COVID-19-MCTIC)

STATISTICAL ANALYSIS PLAN FOR the CLINICAL TRIAL

Version: 2.0

March 2nd, 2020

Authors

Pr Esper Georges Kallas, Pr Alice Song, Pr Vivian Avelino-Silva, Hospital das Clínicas, Faculdade de Medicina, Sao Paulo, Brasi

Pr S Chevret, Saint Louis Hospital, Paris, France

CONTENTS

1. INTRODUCTION

The COVID-19 pandemic has been spreading continuously, and in Brazil, until April 10, 2020, there were more than 17,000 cases with more than 1,000 deaths, with daily growth. To date, there are no vaccines available or effective specific treatment. The administration of convalescent serum as an adjunct to supportive treatment has shown, in a few series of cases, promising results, without adverse effects. However, effectiveness cannot be proven by the absence of a control group. There are many clinical trials underway with this therapeutic modality registered worldwide.

The present project proposes to evaluate the efficacy and safety of convalescent plasma in the treatment of serious and potentially serious cases of COVID-19 in a multicenter study.

To prevent outcome reporting bias and data–driven analysis results, the International Conference on Harmonization of Good Clinical Practice (ICH–GCP) recommends that clinical trials should be analyzed according to a pre–specified detailed Statistical Analysis Plan (SAP). This document presents the updated and finalized SAP of the COOP-COVID-19-MCTIC trial.

2. DESIGN

The protocol has been registered in clinicaltrials.gov (NCT04415086).

It is a multicenter parallel three-arm randomized clinical superiority trial evaluating the benefit and tolerance of the administration of convalescent plasma to potentially severe and severely ill patients with COVID-19.

3. RANDOMIZATION

Block randomization with stratification by study site will be performed with varying block sizes of 3, 6, and 9.

4. ANALYSIS SETS/ POPULATIONS/SUBGROUPS

The eligible population has been described in the main protocol, as follows.

4.1 Inclusion criteria

1. Age equal to or greater than 18 years;
2. Laboratory-proven COVID-19 infection by RT-PCR in any clinical sample
3. Time since symptoms onset less than 10 days at the time of screening;

4. Presence of COVID-19 pneumonia, with a typical, indeterminate or atypical compatible image on a chest CT scan (see definition below)
5. Presence of one of the following criteria:
    a. Need for > 3L of O2 in the catheter / mask or> 25% in the Venturi mask to maintain O2 saturation> 92%
    b. Presence of respiratory distress syndrome with PaO2 / FiO2 <300mmHg
        i. If intubated, within 48 hours of orotracheal intubation;
        ii. Absence of a history of serious adverse reactions to transfusion, for example, anaphylaxis

4.2 Exclusion criteria

1. Participation in another clinical antiviral or immunobiological agent study for the treatment of COVID-19.
2. IgA deficiency
3. Severe Renal Failure (eGFR <30)
4. Presence of heart failure that does not allow infusion of 400 ml of plasma at clinical discretion
5. Pregnancy or breastfeeding
6. Receipt of immunoglobulin in the last 30 days

5. STUDY OBJECTIVES

The overall scientific objectives of the analyses, are to estimate the benefit of the experimental arms, compared to the control, then, in case demonstration of efficacy, to assess whether one of the two experimental arms, performs better than the other

- General objective:
  To assess the efficacy and safety of convalescent plasma transfusion in hospitalized patients with severe and potentially severe COVID-19.

- Specific objectives
  1. To compare patients randomized to convalescent plasma and those receiving only standard therapy for clinical status using the ordinal scale of 10 categories in D7, D14 and D28 after transfusion in the plasma groups, and after 24 hours of randomization in the control group;
  2. To compare patients randomized to convalescent plasma and those receiving only standard therapy regarding the duration of mechanical ventilation, length of hospital stay in survivors up to 28 days, and time from D0 to death.
  3. To compare the primary outcome between the two plasma volumes assessed.

4. To compare the detection of SARS-CoV-2 in respiratory samples on days 0, 1, 3, 7, 14 and 28 in the three groups.
5. To compare serum IgG, IgM and IgA titers specific for SARS-CoV-2 on days 0 (before transfusion), 1, 3, 5, 7, 14 and 28 in the three groups.
6. To compare the detection of neutralizing antibodies on days 0 (before transfusion), 1, 14 and 28 in the three groups.
7. To compare adverse events in plasma groups and control group.

6. ENDPOINTS AND COVARIATES

6.1 Main endpoint

The main end point will be time to clinical improvement, defined as the time from the randomization until the decline of 2 categories in the WHO progression scale, or hospital discharge (whichever comes first).

WHO Progression scale	Descriptor	Score
Uninfected	Uninfected; non-viral RNA detected	0
Ambulatory	Asymptomatic; viral RNA detected	1
Ambulatory	Symptomatic; Independent	2
Ambulatory	Symptomatic; Assistance needed	3
Hospitalized: mild disease	Hospitalized; No oxygen therapy	4
Hospitalized: mild disease	Hospitalized; oxygen by mask or nasal prongs	5
Hospitalized: severe disease	Hospitalized; oxygen by NIV or High flow	6
Hospitalized: severe disease	Intubation and Mechanical ventilation, pO2/FIO2>=150	7

	OR SpO2/FIO2>=200	
Hospitalized: severe disease	Mechanical ventilation, (pO2/FIO2<150 OR SpO2/FIO2<200) OR vasopressors (norepinephrine >0.3 microg/kg/min)	8
Hospitalized: severe disease	Mechanical ventilation, pO2/FIO2<150 AND vasopressors (norepinephrine >0.3 microg/kg/min), OR Dialysis OR ECMO	9
Death	Dead	10

6.2 Secondary end points

1. Incidence of acute adverse events, as defined by the International Society of Blood Transfusion/International Haemovigilance Network
2. Evaluation according to an ordinal scale of 10 categories in D7, D14 and D28, duration of mechanical ventilation, length of hospital stay in survivors up to 28 days and time from the beginning of treatment to death (in the control group, this time will be determined at 24 hours after randomization).
3. Detection of SARS-CoV-2 in nasopharyngeal swab (and tracheal secretion if intubated patient) on days 0, 1, 3, 7, 14 and 28 after transfusion in groups B and C, and 24 hours after randomization in the group A (control).
4. Specific IgS, IgM and IgA titers for SARS-CoV-2 on days 0 (before transfusion), 1, 3, 5, 7, 14 and 28 after transfusion in groups B and C, and 24 hours after transfusion randomization in group A (control).
5. Detection of neutralizing antibodies on days 0 (before transfusion), 1, 14 and 28 after transfusion in groups B and C, and 24 hours after randomization in group A (control).

7. DATA SOURCE

The analysis will use the Intention-to-Treat (ITT) population, that is, all patients will be analyzed once randomized, in the group allocated by the randomization, whatever the observance of the treatment and potential protocol deviations.

The database will be locked as soon as all data are entered and all discrepant or missing data are resolved, after all efforts are employed to complete the database, and we consider that the remaining issues cannot be fixed.
At this step, the data will be reviewed before database locking. After that, the study database will be locked and exported for the statistical analysis. At this stage, permission for access to the database will be removed for all investigators, and the database is locked and archived.

8. STATISTICAL ANALYSIS

Analysis will be stratified on the randomization patient base line severity group.

8.1 Baseline characteristics

Summary statistics will be computed, namely percentages (on available measures) or median (with interquartile range, IQR) unless specified. Number of missing values will be reported (as "NA"). They will be reported according to the randomization arm, with no statistical testing as recommended.

8.2 Primary outcomes

Interim analyses will be performed in order to assess the futility, or the potential harm of trial continuation (see section 10 below).
Cumulative incidence of success (primary outcome defined as gain of 2 points at least in the WHO scale or discharge alive from the hospital), will be provided, with death free of success considered as a competing risk event.
Analyses will use Bayesian methods. Indeed, when comparing multiple products or treatment strategies against one another via a comparative trial, Bayesian analyses appear particularly well suited (Connor 2013; Gsponer 2014; Ryan 2019). Bayesian approaches have two main advantages compared with classical designs (Berry 2004; Spiegelhalter 1994): first, they facilitate implementation of interim success and futility criteria that are tailored to the subsequent decision making based on explicit probabilistic statements (that may offer strong benefit for their ability to calculate the probability that each treatment is the best or worst) (Jacob 2016), and second, they allow inclusion of prior information on the treatment differences and on the control group, actualized over the trial.

8.3 Prespecified Subgroup analyses of primary outcomes

Subgroup analyses will include outcome measures according to number of days of symptoms, total specific SARS-CoV-2 antibody titers, neutralizing antibodies titers, age, and body mass index (using 30 as the predetermined threshold of obesity). When no threshold has been determined in the literature, median and quartiles will be used.

Following Millen (2014) and the approach of Morita (2014) interaction criteria were computed, based on the posterior probability given the observed data in the trial of a treatment benefit. Measures of interaction, were defined as the ratio between the efficacy measures in each subset: θ_0/θ_1, where k = {0, 1} defines the partition of the sample.
Bayesian criteria for evaluation of the interaction condition from the posterior probability of the criterion were to be used (table 1).

Table 1: Criteria for efficacy in subset$_k$ and for interaction for a binary endpoint, where 1 denotes a defavorable outcome such as death

Probability	Evidence for
$P_{0k=}P(\theta_k > 1.05\|data)$	harm in subset k
$P_{1k} = P(\theta_k < 1\|data)$	any benefit in subset k
$P_{2k} = P(\theta_k < 0.8\|data)$	moderate benefit or greater in subset k
$P_{3=} = P(\theta_0/\theta_1 > 1.05\|data)$ $P_{3=} = P(\theta_1/\theta_0 < 0.95\|data)$	any interaction
$P_{4=} = P(\theta_0/\theta_1 > 1.25\|data)$ $P_{4=} = P(\theta_1/\theta_0 < 0.8\|data)$	moderate interaction or greater

θ_k refers to the relative risk of the outcome in subset $k = \{0, 1\}$

8.4 Analysis of secondary outcomes

Analysis will use statistical methods adapted to each outcome measure.

1. The cumulative hazard of adverse events will be estimated by semiparametric model using Andersen and Aalen estimator, compared by likelihood ratio test.
2. The evolution of WHO-scale will modelized and compared by linear mixed models. The number of days alive without mechanical ventilation, the length of hospital stay, and time from treatment to death, will be compared between groups by a Kruskal-Wallis test.
3. The detection of SAR-CoV-2, the Ig S, M and A titers, and of neutralizing antibodies over time will be modelized and compared between groups by a linear mixed model.

9. HANDLING OF MISSING VALUES AND OTHER DATA CONVENTIONS

All efforts will be made not to have missing data on study outcomes, and main covariates. In case of missing data, the reason why the data is missing will be documented.

Complete–case analysis will be carried out for all the outcomes. According to the volume of missingness, simple or multiple (>5% of missing values) imputation techniques for covariates will be used as sensitivity analyses.

10. STATISTICAL METHODOLOGY

10.1 Interim analyses

Interim monitoring for success will be performed once 30 patients have been consecutively enrolled and will be repeated after every additional 30 patients are enrolled. Interim analyses will be stratified on the patient base line severity group, that is performed in each stratum, independently, based on the binary secondary outcome measure of each subset, namely the decrease of at least 2 points in the 10-point WHO scale at day 7.
At each interim analysis, we will report the response rate for each treatment group with 95% credible intervals as well as the pairwise differences D in response rates between experimental arms and control arm with 95% credible intervals.

From a Bayesian perspective, the decision making will be based on posterior probabilities. Indeed, Bayesian posterior probabilities, will be computed at any point in the trial and provide current evidence about all possible questions, such as

P1=P(D>0 | data)	any beneficial effect
P2=P(D>0.15 | data)	clinically relevant effect
P3=P(D<0 | data)	inefficacy or harm
P4=P(D<-0.15 | data)	more than trivial harm

For this study, action triggers were set at the following:
Stop the arm with evidence for efficacy if P1>0.95
Stop with evidence for moderate or greater efficacy of the arm if P2>0.8
Stop with evidence for inefficacy of the arm if P3>0.8
Stop with evidence for harm of the arm if P4>0.75

We will use noninformative Beta (1,1), that is, uniform prior for the proportion of the outcome in each treatment arm, in each stratum. We will actualize those priors in posterior probabilities, then derived the previous probabilities P1-P4, pooling the two experimental arms, or pairwise differences for each arm, separately. If both appear relevant, they will be compared to each other. In case any of the treatment groups performs worse (that is, on the basis of P3 or P4), then we will consider the possibility of drop the loser

If there is a failure to achieve clinical improvement or death before D28, the participant will be censored on D28.

10.2 MEASURES TO ADJUST FOR MULTIPLICITY, CONFOUNDERS, HETEROGENEITY, ETC.

Briefly describe and justify these measures if applicable. For example, in the setting of observational data analyses, one possible adjustment measure might be propensity scoring.

11. SENSITIVITY ANALYSES

Evidence for differential treatment effect (also called heterogeneity of treatment effect) will be assessed by including interaction between treatment and each of the pre-specified covariates one-at-a-time.

12. RATIONALE FOR ANY DEVIATION FROM PRE-SPECIFIED ANALYSIS

Any difference will be considered and submitted as amendments to the protocol.

13. PROGRAMMING PLANS

All codes will be generated using R (https://www.R-project.org/).

14. REFERENCES

- Berry DA. Bayesian statistics and the efficiency and ethics of clinical trials. Stat Sci. 2004;19:175–87.
- Connor JT, Elm JJ, Broglio KR, and for the ESETT and ADAPT-IT Investigators. Bayesian adaptive trials offer advantages in comparative effectiveness trials: an example in status epilepticus. Journal of Clinical Epidemiology. 2013:66:S130-S137.
- Gsponer T, Gerber F, Bornkamp B, Ohlssen D, Vandemeulebroecke M, Schmidli H. A practical guide to Bayesian group sequential designs. Pharmaceut. Stat. 2014;13:71-80.
- Jacob L, Uvarova M, Boulet S, Bega I, Chevret S. Evaluation of a multi-arm multi-stage Bayesian design for phase II drug selection trials – an example in hemato-oncology. BMC Medical Research Methodology. 2016;16:67.

- Millen BA, Dmitrienko A, Song G. Bayesian assessment of the influence and interaction conditions in multipopulation tailoring clinical trials. J Biopharm Stat. 2014;24(1):94-109.
- Morita S, Yamamoto H, Sugitani Y. Biomarker-based Bayesian randomized phase II clinical trial design to identify a sensitive patient subpopulation. Stat Med. 2014; 33:4008-4018
- Ryan EG, Bruce J, Metcalfe AJ, et al. Using Bayesian adaptive designs to improve phase III trials: a respiratory care example. BMC Medical Research Methodology. 2019;(1).
- Spiegelhalter DJ, Freedman LS, Parmar MKB. Bayesian approaches to randomized trials (with discussion). J R Stat Soc, Ser A. 1994;157:357–416.
- Working Party for Haemovigilance. International Society of Blood Transfusion/Internation Haemovigilance Network. Available at https://www.isbtweb.org/fileadmin/user_upload/Proposed_definitions_2011_surveillance_non_infectious_adverse_reactions_haemovigilance_incl_TRALI_correction_2013_TACO_correction_2018.pdf with the updated TACO definition at https://www.isbtweb.org/fileadmin/user_upload/TACO_2018_definition_March_2019.pdf. Accessed on April 20 2020.

15. APPENDICES

- Definition and use of visit windows in reporting
- Definition of Analysis Populations/Sets
- Further Definition of Endpoints
- Statistical Methodology Details