An Open-label Randomized Multicenter Study to Evaluate the Efficacy of Early Administration of Tocilizumab (TCZ) in Patients With COVID-19 Pneumonia.

Statistical Analysis Plan

Version: 1.0.

Date: 27.th may 2020.

Author: Silvio Cavuto, MsC.

Trial registration: EudraCT (2020-001386-37) and ClinicalTrials.gov (NCT04346355).

Methods/design

The study is registered in the EudraCT database (2020-001386-37) and in the ClinicalTrials.gov Protocol Registration and Results System (NCT04346355). The main results of this trial will be published following the Consolidated Standards of Reporting Trials (CONSORT) statement (Schultz et al, 2010). The sponsor is the Azienda USL-IRCCS di Reggio Emilia, Italy.

Study aims

Primary aim

Evaluate the efficacy of TCZ administered early in patients with COVID-19 pneumonia compared to patients who receive standard therapy in the two weeks after entering the study.

Secondary aims

1. Compare the effectiveness of TCZ in terms of admission to ICU with invasive mechanical ventilation in two groups:
    a. patients to whom TCZ is administered early as per protocol;
    b. patients to whom TCZ is administered at clinical worsening (for an arterial partial pressure of oxygen / fraction of inspired oxygen (PaO_2 /FiO_2) ratio <150).
2. Compare the effectiveness of TCZ in terms of all-cause mortality in two groups:
    a. patients to whom TCZ is administered early as per protocol.
    b. b. patients to whom TCZ is administered at clinical worsening (for a PaO_2/FiO_2 ratio <150) or in the first 24 hours after admission to ICU.
3. Evaluate TCZ toxicity
4. Evaluate the levels of IL-6 and serum C-Reactive Protein (CRP) and their correlation with the outcome
5. Evaluate the levels of ferritin, Lactate Dehydrogenase (LDH) and D-dimer and their correlation with the outcome
6. Evaluate the progress of the PaO_2/FiO_2 ratio and their correlation with the outcome
7. Evaluate the trend over time of the lymphocyte count and their correlation with the outcome

Design and clinical centers

This is a randomized, open label, parallel, phase 2 trial comparing early administration of TCZ in hospitalized patients with COVID-19 pneumonia to late administration of TCZ (when COVID-19 pneumonia aggravates).

Patients allocated to the experimental arm will receive TCZ therapy within 8 hours from entering the study + standard therapy while those allocated to the control arm will receive standard therapy. In the event of clinical aggravation or entry into the ICU, patients will receive TZC. Clinical

aggravation is defined by an arterial partial pressure of oxygen/fraction of inspired oxygen (PaO_2/FiO_2) ratio <150 at one of the scheduled arterial blood gas (ABG) measurements or at an emergency measurement confirmed by a second ABG within 4 hours.
COVID-19 patients will be recruited in 24 clinical centers located in five Italian Regions: Emilia Romagna, Liguria, Lombardia, Toscana, Piemonte, and Veneto.

Study population

The study population includes patients with recent-onset COVID-19 pneumonia who require hospital care, but not invasive or semi-invasive mechanical ventilation procedures at enrolment. The inclusion criteria are:

- age> 18 years;
- Informed consent for participation in the study;
- Real time Polymerase Chain Reaction (Real-time PCR) diagnosis of Sars-CoV2 infection;
- Hospitalization due to clinical/instrumental diagnosis (high resolution chest computed tomography scan or chest x-ray or pulmonary ultrasound);
- Presence of acute respiratory distress syndrome with a PaO_2/FiO_2 ratio between 200 and 300 mm/Hg;
- Presence of exaggerated inflammatory response defined by the presence of at least 1 of the following criteria: at least one body temperature measurement > 38 °C in the past two days; serum CRP greater than or equal to 10 mg/dl; CRP increase of at least twice the basal value.

The exclusion criteria are:

- Patient with acute respiratory distress syndrome with PaO_2/FiO_2 ratio < 200 mm/Hg or in non-invasive ventilation or in invasive ventilation or presence of shock or presence of concomitant organ failure that requires admission to the ICU;
- Severe heart or kidney failure;
- Pregnant or breastfeeding patient;
- Patient who, in the opinion of the clinician or by the patient's express will, will not go to intensive care regardless of the evolution of the lung picture;
- Known hypersensitivity to TCZ or its excipients;
- Patient being treated with immuno-depressors or anti-rejection drugs;
- Known active infections or other clinical conditions that contraindicate TCZ and cannot be treated or resolved according to the physician's judgment;
- Glutamate-pyruvate transaminase (GPT) or glutamine oxaloacetic transaminase (GOT) > 5 times the upper limit of the norm;

- Neutrophils < 500/mmc;
- Platelets < 50.000/mmc;
- Diverticulitis or intestinal perforation
- Suspicion of latent tuberculosis

Assessments

Endpoints of the study

The primary endpoint is defined by the appearance, within 2 weeks of randomization, of one of these 3 events:

a) entry into ICU with invasive mechanical ventilation;
b) death from all causes;
c) clinical aggravation documented by the finding of a PaO_2/FiO_2 ratio < 150mm/Hg to one of the programmed ABG measurements or to an ABG emergency measurement and confirmed by a second examination within 4 hours.

The assessment of secondary endpoints is done in the order in which the secondary objectives of the study are presented:

1) entry into ICU with invasive mechanical ventilation;
2) all-cause mortality;
3) toxicity measured according to internationally recognized standard.

Clinical assessments

Evaluations are scheduled in the 2 weeks of the study or until the patient leaves the study due to death, transition to intensive care or clinical aggravation.

Day 1 is the day of randomization regardless of the time it was made (00-24). Exams done for the assessment of eligibility prior to randomization can be used for day 1. The required exams are those required by the protocol. The doctor is free to prescribe additional tests at his discretion. In case of early termination of the study due to ICU transfer or clinical aggravation, the evaluations foreseen by the protocol are suspended.

Timing of the programmed arterial EGAs and laboratory evaluations are described in Table 1. During the planned days of assessments temperature, saturation and respiratory frequency will be collected for all patients.

Registration and randomisation procedures

Eligible patients will be registered in a centralized database developed by the Information Technologies Service of the Azienda USL-IRCCS of Reggio Emilia. Randomisation will take place on a competitive and balanced basis by participating clinical center. Random lists stratified by center will be prepared using permuted blocks of various sizes in random order. Randomisation will be carried out by telephone access to the Clinical Trials and Statistics Unit of the Azienda USL-IRCCS 7 days a week, 24 hours a day. This operation also determines the registration of enrolled patients.

Statistical considerations

Sample size determination

The sample size was calculated congruently with the primary objective and in accordance with the following assumptions. Null hypothesis (H_0): $\pi_{ctrl} = \pi_{exp} = 0.20$ (where π_{ctrl} and π_{exp} are the "failure" rates in the control and experimental group respectively);

1. Alternative hypothesis (H_1): $\pi_{ctrl} = 0.20$ & $\pi_{exp} = 0.10$ which corresponds to a 50% reduction;
2. Statistical test: chi-square (asymptotic);
3. two-sided test;
4. alpha error = 0.05;
5. statistical power = 0.80;
6. allocation ratio: 1: 1

Based on these assumptions, a sample of 199 patients per group is required, 398 in total.

The sample size was calculated using PROC POWER implemented in the SAS / STAT package of the SAS software, version 9.4 for Microsoft OS.

Statistical analysis

Efficacy analyses

The statistical analyses are described by aim as follow:

Primary aim:

The primary efficacy analysis will be conducted on the randomized population for the purposes set out in the primary objective, following the intention-to-treat principle.

A secondary analysis will be conducted “per protocol” for completeness and with descriptive purposes.

The failure rate of the two arms will be compared using the chi-square test (in an asymptotic form). The Rate Ratio (RR) will also be calculated, with the related bilateral 95% confidence interval, to

facilitate the comparison of the rates mentioned above from a descriptive point of view: the confidence limits will be calculated using the (unmodified) Wald method.

In case of computational or distributional issues possibly affecting the planned asymptotic chi-square test, it will be performed the exact calculation of the related p-value. To this aim, the calculation will be based on the Mehta & Patel (1983) algorithm as implemented in PROC FREQ of SAS/STAT package. In this case, the calculation of the confidence limits for the RR will be changed too for coherence, shifting to unconditional confidence limits: so, the tail methods by Chan & Zhang (1999) will be applied to the p-values calculated following the Santner & Snell (1980) approach.

Secondary aims:

1. the rate of patients admitted to ICU with invasive mechanical ventilation will be analysed as stated for the primary aim;
2. the all-cause mortality rate will be analysed as stated for the primary aim;
3. toxicity will be described by tables and listings;
4. Level of IL-6 will be compared between groups using the t-test for unpaired data; in case of heteroskedasticity, the test will be fixed using the Satterthwaite method (Satterthwaite, 1946). The 95% 2-sided confidence interval will be provided for the mean of each IL-6 level. The same analysis will be conducted for the C-Reactive protein.
5. Levels of ferritin, LDH and D-dimer will be analysed as stated for the level of IL-6;
6. the progress of the PaO_2/FiO_2 ratio will be evaluated using the analysis of variance for repeated measures (Verbeke & Molenberghs, 2000). The underlying model will be estimated by the REsidual (restricted) Maximum Likelihood (REML) method (Brown & Prescott, 1999; Little et al, 2006). An unstructured R-side covariance matrix will be assumed, shifting to a simpler AR(1) structure in case of estimation troubles.
7. the trend over time of the lymphocyte count will be evaluated using the analysis of variance for repeated measures for Poisson variables (Diggle et al, 1994). The underlying models will be estimated by the Generalized Estimating Equations method (Liang & Zeger, 1986; Zeger et al, 1988) An unstructured within-subject correlation structure will be assumed, shifting to a simpler AR(1) structure in case of estimations troubles. To effectively describe the possible trend, a regression model over time will be also estimated assuming the same correlation structure.

The conventional level of 0.05 for statistical significance will be adopted to evaluate the p-values. Statistical analysis will be carried out by using R and SAS System.

Safety analysis

Safety analyses will be performed only on the safety population defined above.

For each patient and for each type of toxicity described according to Common Terminology Criteria for Adverse Events, the worst grade found during treatment will be used for descriptive analysis. These data will be described using lists and tables.

References.

Brown, H. and Prescott, R. (1999), *Applied Mixed Models in Medicine*, New York: John Wiley & Sons.

Chan, I. S. F., and Zhang, Z. (1999). “Test-Based Exact Confidence Intervals for the Difference of Two Binomial Proportions.” *Biometrics* 55:1202–1209.

Diggle, P. J., Liang, K. Y., and Zeger, S. L. (1994), *Analysis of Longitudinal Data*, Oxford: Clarendon Press.

Liang, K. Y. and Zeger, S. L. (1986), “Longitudinal Data Analysis Using Generalized Linear Models,” *Biometrika*, 73, 13–22.

Littell, R. C., Milliken, G. A., Stroup, W. W., Wolfinger, R. D., and Schabenberger, O. (2006), *SAS for Mixed Models,* Second Edition, Cary, NC: SAS Institute Inc.

Mehta, C. R., and Patel, N. R. (1983). “A Network Algorithm for Performing Fisher’s Exact Test in r x c Contingency Tables.” *Journal of the American Statistical Association* 78:427–434.

Santner, T. J., and Snell, M. K. (1980). “Small-Sample Confidence Intervals for $p_1 – p_2$ and p_1 / p_2 in 2 x 2 Contingency Tables.” *Journal of the American Statistical Association* 75:386–394.

Satterthwaite, F. W. (1946), "An Approximate Distribution of Estimates of Variance Components," *Biometrics Bulletin*, 2, 110–114.

Schulz KF, Altman DG, Moher D (2010). “CONSORT 2010 Statement: Updated guidelines for reporting parallel group randomised trials”, *Journal of Clinical Epidemiology,* 2010;63:834–840.

Verbeke, G. and Molenberghs, G. (2000), *Linear Mixed Models for Longitudinal Data,* New York: Springer.

Zeger, S. L., Liang, K. Y., and Albert, P. S. (1988), “Models for Longitudinal Data: A Generalized Estimating Equation Approach,” *Biometrics*, 44, 1049–1060.