



A phase 2, open label, multicenter, single arm study of tocilizumab on the efficacy and tolerability of tocilizumab in the treatment of patients with COVID-19 pneumonia (TOCIVID-19 trial): Statistical analysis plan

Paolo Chiodini^{a,*}, Laura Arenare^b, Maria Carmela Piccirillo^b, Francesco Perrone^b, Ciro Gallo^a

^a Medical Statistics Unit, University of Campania 'Luigi Vanvitelli', Napoli, Italy

^b Istituto Nazionale Tumori, IRCCS, Fondazione G. Pascale, Napoli, Italy

ARTICLE INFO

Keywords:

Statistical analysis plan
COVID-19 pneumonia
Tocilizumab
Phase 2 trial

ABSTRACT

Background: Tocilizumab, an IL-6 receptor antagonist, was suggested as a possible treatment of severe or critical COVID-19 pneumonia in a small Chinese study. The TOCIVID-19 trial evaluates efficacy and tolerability of tocilizumab in the treatment of patients with severe or critical COVID-19 pneumonia.

Methods: TOCIVID-19 is an academic multicenter, single-arm, open-label, phase 2 study. All the patients are being offered a single shot of 8 mg/kg of Tocilizumab (up to a maximum of 800 mg), with an eventual second administration at the discretion of the Investigator. A companion prospective cohort, added to corroborate internal validity, includes either patients not eligible for phase 2 or subjects eligible for phase 2 but exceeding the planned sample size. 14- and 30-days lethality rates are the two co-primary endpoints in the intention-to-treat (ITT) population. Secondary objectives are to evaluate mortality and clinical improvement in the modified-ITT population of subjects who received the drug. Details of the methodological and statistical approaches are reported here reflecting the amendments impelled by the continuously increasing knowledge on COVID-19 progression and challenges in data collection.

Conclusion: This paper provides details of planned statistical analyses for TOCIVID19 trial to reduce the risk of reporting bias and increase validity of the study findings.

TOCIVID-19 trial is registered in the EudraCT database with number 2020-001110-38 and in clinicaltrials.gov with ID NCT04317092.

1. Introduction

1.1. Background and rationale

This multicenter study on the efficacy and tolerability of tocilizumab in the treatment of patients with COVID-19 pneumonia (TOCIVID-19) has been realized in the context of the COVID-19 epidemics in Italy. Such epidemic has led to a rapidly increasing number of cases of interstitial pneumonia with a high case-fatality ratio and a tremendous burden on Italian intensive and sub-intensive care units. The number of deaths associated with COVID-19 in Italy is high, and is mainly concentrated in a few Regions (Lombardia, Piemonte, Emilia-Romagna, Veneto) [1,2].

Tocilizumab, an IL-6 receptor antagonist, is an anti-inflammatory drug used in rheumatology for the treatment of some forms of arthritis and in oncology to fight the 'cytokine storm' subsequent to immunotherapy. It was suggested as a possible treatment of severe or critical COVID-19 pneumonia by Chinese researchers [3] in a cohort of 21 patients whose clinical variables improved dramatically. Then a randomized trial comparing tocilizumab versus control started in China evaluating about 190 patients [4]; as far as we know the trial is still ongoing. After the initiation of the present study two other randomized phase 3 trials have been launched by the University of Oxford (EudraCT 2020-001113-21) and by the drug manufacturer (NC-T04320615).

Notwithstanding the previous limited information, tocilizumab was steadily used off-label in Italy to treat severe or critical COVID-19 pneumonia, and anecdotal initial positive results reported by physi-

* Corresponding author. Medical Statistics Unit, University of Campania 'Luigi Vanvitelli'. Via Luciano Armanni, 5 - 80138, Napoli, Italy.
E-mail address: paolo.chiodini@unicampania.it (P. Chiodini).

cians gave rise to a massive media campaign for a generalized use of drug. Thus, the main concern was to provide more accurate and methodologically sound information on the efficacy of tocilizumab.

1.2. Objectives

The TOCIVID-19 trial is a single-arm open label phase 2 study designed to assess the efficacy of tocilizumab in reducing mortality of patients with severe or critical COVID-19 pneumonia. The aim of this Statistical Analysis Plan (SAP) is to report in detail the methodological and statistical approaches to the analysis of the data whilst still blind to any analyses of efficacy study outcomes. This SAP is produced at the end of April 2020, and any deviations will be documented in the final clinical study report. It is based on the TOCIVID-19 study protocol, version 3, dated April 28th, 2020, and is in line, where applicable, with the literature guidelines [5].

2. Methods

2.1. Trial design

TOCIVID-19 is an academic, no-profit, multicenter, single-arm, open-label, phase 2 study. All the patients enrolled receive a single shot of 8 mg/kg of Tocilizumab (up to a maximum of 800 mg). A second administration of Tocilizumab at the same dose is allowed after 12 h if respiratory function has not recovered, at the discretion of the Investigator. Lethality rate is the primary endpoint. The study is registered in the EudraCT database with number 2020-001110-38 and in clinicaltrials.gov with ID NCT04317092.

The single-arm design was prompted by social needs and feasibility reasons. On one hand, although the rationale behind the use of tocilizumab is clinically plausible, previous Chinese experience is weak. Nevertheless, an aggressive media campaign has arisen favoring an indiscriminate request of the drug on an off-label basis from physicians. Such powerful plea for the drug in Italy, despite the absence of strong evidence, would have made achieving consent to participation in a controlled randomized trial very difficult. This belief was confirmed by ensuing events, particularly the massive registration of centers as soon as the registration was opened.

Thus, this study design represents the effort to strike a balance between the clinicians' demand on one side and the scientific need to have the most accurate information by governing the data collection on a formalized basis.

Because of the unprecedented situation of COVID-19 pandemic, inadequate information was available when the trial was planned, and this SAP reflects the subsequent amendments impelled by the continuously increasing knowledge.

The planned number of patients of phase 2 population was completed in less than 24 h, from March 19 to March 20, 2020. The core phase 2 trial was complemented by a further cohort of patients treated with Tocilizumab at the same doses, with a prospective and a retrospective component. The prospective cohort was defined as treated patients not enrolled in the phase 2 cohort, but prospectively registered *before* receiving Tocilizumab. The prospective cohort includes either subjects eligible for phase 2, but possibly exceeding the planned phase 2 sample size, or patients not eligible for phase 2 because of either practical circumstances or time of intubation superior to 24 h before registration. The retrospective cohort is defined as patients who received the drug but were registered *after* receiving Tocilizumab.

Many issues were encountered with centers regarding data collection, mainly because registration in the trial was the only way to obtain the drug in a very critical moment of the spread of the disease. Thus, we expect an important proportion of missing data. As of April 15, only 151 phase 2 patients were available for analysis, according to

baseline and treatment information. Efforts to obtain more data will continue until the date of data lock, and the exact number will be reported in the final paper. This scenario led us to add a 'validation cohort', involving patients registered in the prospective cohort, with the same eligibility criteria of phase 2, from March 20 to March 24, 2020 when the enrolment of patients was temporarily halted because of drug shortage. Hopefully, the findings of this validation cohort would corroborate the internal validity of the study main results.

This SAP only refers to the phase 2 and validation cohorts. The remaining patients from the prospective and retrospective cohorts will be investigated in separate analyses.

2.2. Eligibility criteria

Inpatients with virological diagnosis of SARS-CoV-2 infection (real-time PCR) and clinical or instrumental diagnosis of pneumonia who had an oxygen saturation at rest $\leq 93\%$ or intubated less than 24 h before registration were eligible. Neither age nor gender limits were contemplated.

3. Outcomes

3.1. Primary outcome

When the trial was first planned only one primary outcome was defined, i.e. 1-month death rate. However, biweekly reports of the Italian National Institute of Health (ISS) [6] on deceased COVID-19 patients recorded median times of 10 days and 5 days from onset of symptoms and death, and hospitalization and death, respectively. Accordingly, it appeared that delaying death assessment at 1 month might be unnecessary and death estimates at 14 days might be very informative, also being less prone to loss of information.

Accordingly, the April 28th protocol amendment introduced lethality rates measured 14 and 30 days after registration as co-primary outcomes of the intention to treat analysis in the TOCIVID-19 protocol version 3.

3.2. Secondary outcomes

Secondary outcomes are:

- lethality rates at two weeks and one month in the subgroup of phase 2 patients who actually received the experimental drug;
- lethality rates at two weeks and one month according to delay of treatment;
- time to death, defined as the time from registration to death or being alive within 30 days;
- respiratory function in terms of:
 - time to invasive mechanical ventilation (if not previously initiated), defined as the time from registration to first occurrence of mechanical ventilation;
 - time to definitive extubation (if previously intubated), defined as the time from registration to definitive extubation;
 - time to independence from non-invasive mechanical ventilation, defined as the time from registration to the definitive stopping of mechanical ventilation;
 - time to independence from oxygen therapy, defined as the time from registration to the definitive stopping of oxygen supplementation;
- duration of hospitalization, defined as the time spent in hospital from registration to death or discharge;
- longitudinal changes in clinical and laboratory variables outlined in the protocol as reported in medical records (IL-6 levels, CRP levels, PaO₂/FiO₂ ratio, body temperature, lymphocyte count, SOFA score, radiological response);

- prognostic value of baseline variables (age, IL-6 levels, CRP levels, lymphocyte count) in COVID-19 patients;
- safety outcomes as codified by Common Terminology Criteria for Adverse Events (CTCAE) v. 5.0.

3.3. Sample size

As of March 15, 2020, when only preliminary data were available, 330 patients were planned assuming a 15% 1-month lethality rate as H_0 , an alternative hypothesis (H_1) equal to 7.5% (i.e. halving the risk of death) with a very high power 0.99 and two-tailed alpha level equal to 0.05.

As more Italian data were available, it was evident that initial null hypothesis was seriously underestimated and should be refined.

A main concern was the accessibility of data referring to patients hospitalized with severe or critical COVID 19 pneumonia. As of April 12, 2020, from data of Minister of Health we calculated an overall death ratio of 23.3%, under the assumption that all patients deceased or cured were hospitalized [7].

Confidential individual data supplied by ISS [8], concerning the Veneto region, showed cumulative death rates of 15.6% (day 14) and 28.2% (day 30) calculated by the Kaplan-Meier product-limit method (as for April 15).

However, these data should still be slightly underestimated because of the presence of people with very short follow up and could not be simply generalized to the whole Italian population, as it seemed clear, from official data by Minister of Health, that Veneto has been one of the Italian regions with lower death rates (Table 1). Under the assumption that case mix of our sample is similar to case mix of the whole Italian population, the overall expected death rate in our trial should be equal to 21.7% and 39.2% at 14 and 30 days, respectively (multiplying by 1.39 the two Veneto estimates above).

This encouraged us to adopt, in agreement with the Independent Data Monitoring Committee (IDMC), an estimate of the null hypotheses (H_0) to 20% and 35% at 14 and 30 days, respectively. Likewise,

Table 1

Deaths from hospitalized COVID-19 patients as for April 15, 2020 (elaboration from official data of Minister of Health).

	Hospitalized cases	Deaths	death rate	ratio of death rates
Lombardia	42349	11377	0.27	1.56
Emilia	11165	2788	0.25	1.45
Piemonte	8738	2015	0.23	1.34
Veneto	5456	940	0.17	1.00
Toscana	2442	556	0.23	1.32
Marche	3551	807	0.23	1.32
Liguria	3400	746	0.22	1.28
Lazio	2723	311	0.11	0.66
Campania	1413	278	0.20	1.14
Trento	1476	318	0.22	1.25
Puglia	1276	288	0.23	1.31
Friuli	1337	212	0.16	0.92
Sicilia	1044	181	0.17	1.01
Abruzzo	849	240	0.28	1.64
Bolzano	846	223	0.26	1.53
Umbria	899	54	0.06	0.35
Sardegna	424	83	0.20	1.14
Calabria	319	71	0.22	1.29
Valle	533	121	0.23	1.32
Basilicata	131	21	0.16	0.93
Molise	88	15	0.17	0.99
Overall	90459	21645	0.24	1.39

Hospitalized cases: calculated as all positive cases minus positive subjects confined at home. Deaths: reported deaths. Death rates: # deaths out of # hospitalized cases. Ratio of death rates: ratio of death rate of each region and that of Veneto (assumed as reference), i.e. the death rate of Lombardia is 1.56 times the death rate of Veneto.

significance level for each co-primary test was decreased to 0.025 to account for multiplicity.

The number of 330 patients planned for the phase 2 study would still be adequate to recognize an absolute risk reduction (ARR) $\geq 10\%$, assumed as clinically relevant. The trial power (calculated with binomial enumeration) by ARR is reported in Table 2, according to different assumptions, assuming a two-tailed exact test with alpha level equal to 0.025.

4. Statistical analysis

4.1. Trial populations

According to the study design the following populations are defined:

- **Intention to treat (ITT) phase 2 population**, defined as all patients enrolled in the phase 2 cohort. ITT population also includes patients who have received the drug some days after registration due to the shortage of the study drug immediately after the start of the study, and patients who could not have get it at all.
- **Modified ITT (mITT) phase 2 population**, defined as all patients of the ITT phase 2 population who received at least one dose of the study drug.
- **ITT validation population**, defined as all patients consecutively and prospectively registered from March 20 to March 24 who were potentially eligible for the phase 2 study but could not be enrolled because of the completion of the phase 2 cohort. Analyses in the validation population will help to corroborate the phase 2 findings.
- **mITT validation population**, defined as all patients of the ITT validation population who have received at least one dose of the study drug.
- **Safety population**, defined as all patients in the phase 2 and validation populations who received at least one dose of the study drug.

4.2. Efficacy analyses for the primary outcome

Primary analyses will be performed according to ITT strategy, in line with the Treatment policy strategy for all intercurrent events [9], where the target of estimation is considered regardless of the occurrence of intercurrent events, and the estimand mirrors the decision to treat a patient rather than the effect of the treatment itself. The estimand framework explicitly acknowledges the impact on the study question of the choices made to deal with events that “occur after treatment initiation and either preclude the observation of the [end-point] variable or affect its interpretation” (intercurrent events).

As for this study we anticipate the following intercurrent events:

- Delay in administering the treatment due to shortage or administrative reasons;

Table 2

Trial power by absolute risk reduction, according to different H_0 assumptions, assuming a two-tailed exact test with alpha level equal to 0.025 and with a sample size of 330.

Death rate	Death rate at 14 days			Death rate at 30 days		
	15%	20%	25%	30%	35%	40%
Absolute risk reduction						
10%	> 99%	> 99%	99%	98%	95%	93%
12.5%	> 99%	> 99%	> 99%	> 99%	> 99%	99%
15%	> 99%	> 99%	> 99%	> 99%	> 99%	> 99%

- Discharge before d_{14} ;
- Discharge before d_{30} only for the 1-month outcome;
- Discharge before start of treatment, thus preventing treatment administration;
- Death before start of treatment, thus preventing treatment administration;
- Transfer to other wards;
- Second administration of Tocilizumab;
- Change in concomitant treatments.

Accordingly, the primary estimand and key secondary estimands are defined as follows:

4.2.1. Primary estimand

- *Treatment condition of interest*: all patients registered in the study, regardless of actual administration of tocilizumab, time of administration, dose and number of administrations
- *Population*: prospectively registered patients with severe or critical COVID-19 pneumonia as defined in the eligibility criteria (ITT phase 2 population and ITT validation population)
- *Variable*: death/alive status within 14 and 30 days from registration (primary endpoints)
- *Population-level summary*: case-fatality ratio with 97.5% confidence interval.

4.2.2. Time-to-death secondary estimand

- *Treatment condition of interest*: all patients registered in the study, regardless of actual administration of tocilizumab, time of administration, dose and occurrence of a second administration;
- *Population*: prospectively registered patients with severe or critical COVID-19 pneumonia as defined in the eligibility criteria (ITT phase 2 population and ITT validation population);
- *Variable*: death/alive status at each day within 30 days from registration;
- *Population-level summary*: cumulative probability of death with 95% confidence interval.

“Treatment availability as per normal practice” secondary estimands.

These secondary estimands differ from the primary estimand only for the handling of the intercurrent event “Delay in administering the treatment due to shortage or administrative reasons”, which is treated with a “hypothetical” strategy:

- *Treatment condition of interest*: all patients registered in the study when Tocilizumab is available as per normal practice, regardless of time of administration, dose and occurrence of a second administration;
- *Population*: prospectively registered patients with severe or critical COVID-19 pneumonia as defined in the eligibility criteria;
- *Variable*: death/alive status within 14 and 30 days from registration;
- *Population-level summary*: case-fatality ratio with 95% confidence interval.

This target of estimation seems justified by the intent to estimate the effect of the treatment in a situation where its availability is not affected by the very special circumstances encountered at the peak of the epidemics in Italy.

4.3. Data collection

Baseline information includes measurements of demographic and clinical characteristics of patients collected before the first administra-

tion of treatment. On-treatment (longitudinal) data are collected thereafter until discharge or the end-date of 30 days from registration.

The registered data are checked for completeness and accuracy, and amended, if appropriate. For missing, incoherent, and implausible data, a query is raised, which is to be solved by the staff personnel. As soon as information is considered no more susceptible to be improved in the short time, database will be locked and made available for data analysis. The main analysis is planned as soon as possible after a month from the registration of the last patient.

Data collection is web-based (<http://www.usc-intnapoli.net>) or by paper CRF. Data management is centralized at the Clinical Trials Unit of the National Cancer Institute of Napoli. The analysis is carried out using Stata version 14.0 (Stata Corp. College Station, TX, USA) and R version 3.6.1 (R Foundation for Statistical Computing, Vienna, Austria).

4.4. Methods of analysis and presentation of results

4.4.1. Data presentation

The flow of patients of the phase 2 and validation populations defined above will be summarized using a Consolidated Standards of Reporting Trials diagram [10]. Baseline characteristics of patients' populations will be described with the usual summary measures, as appropriate, for the two main populations (phase 2 and validation). Geographical distribution of patients of the two populations will be reported to ease the external validity of findings.

Differences between groups at baseline will be assessed for categorical variables using χ^2 and Fisher's exact as appropriate. For continuous variables *t*-test and ANOVA will be used and the Shapiro Wilks test will be performed to evaluate the normality, otherwise Wilcoxon test or Kruskal-Wallis test for independent data will be applied.

Protocol deviations and actual exposure to treatment.

Reasons for protocol deviations will be summarized in a table. Absolute and relative frequencies of patients not meeting inclusion/exclusion criteria will be reported.

Frequency distributions of actual administration, time of administration, dose and number of administrations will be reported.

4.4.2. Primary endpoint analysis

Proportion of death at day 14 and day 30 will be calculated with exact 97.5% Clopper-Pearson confidence intervals (CI). Test of the pre-specified null hypotheses at days 14 and 30 will be done by means of a two-sided binomial test with alpha level equal to 0.025, in phase 2 ITT population.

Every effort is made to verify the outcome (i.e. status as death/alive at 14 and 30 days) through clinical records and distant monitoring (telephone call). Patient discharged to home or to low-intensity care setting will be considered alive at the end-date of 30 days (in the absence of additional data confirming otherwise). Information on withdrawals and reasons for withdrawal will be described in a summary table. No imputation models will be implemented.

Information on the primary outcomes will be described for baseline subgroups defined by demographics and clinical variables (treatment status, age, gender, geographical area, important comorbidities, mechanical ventilation, and concomitant treatments) and compared with exact chi square test.

We assume *a priori* that treatment status is only dependent on contingent circumstances (i.e. drug shortage) and not on physicians' or patients' decision (i.e. is not likely affected by prognostic factors). However, a physicians' selection bias could still have happened, even though we cannot anticipate the direction. We will check our assumption by comparing the baseline characteristics of groups defined by treatment status, and multivariable logistic regression models will be performed with 14- and 30-day mortality as dependent variables and

treatment information (treatment status and time to treatment) and significant baseline characteristics (age, gender, geographical area, important comorbidities, mechanical ventilation, and concomitant treatments) as covariates.

4.4.3. Time-to-event analyses

For time-to-event analyses, estimates of cumulative probabilities of the event will be provided with 95% CIs. Cumulative probabilities of event will be estimated by using 1 minus Kaplan-Meier estimator or by using cumulative incidence estimator in the presence of other competing events [11]. Median time to event and interquartile range will be calculated for subjects with event. In the absence of additional vital status data, patient discharged earlier than 30 days to home or to low-intensity care setting will be considered alive at the end-date of 30 days. Dependency on waiting time to treatment will be investigated by non-parametric hazard curves of death in predefined subgroups of waiting times according to 'clock back' scale [12]. Time-to-event analysis will be performed on ITT and mITT populations. Information on time-to-event analyses will be described also for baseline subgroups defined by demographic and clinical variables. A sensitivity analysis will be performed by considering patients discharged earlier than 30 days to home or low-intensity care settings as competing events and using cumulative incidence estimator.

Multivariable analyses will be performed by the proportional hazard Cox model where treatment is entered as a time-dependent covariate, to properly account for immortal-time bias [13]. Proportionality assumption will be tested by Schoenfeld residuals.

For time to death, all patients will be included in the analysis and patient discharged earlier than 30 days to home or to low-intensity care setting will be considered alive at the end-date of 30 days.

For time to invasive mechanical ventilation, patients with invasive mechanical ventilation within 24 h from registration date will be excluded from the analysis and death will be considered as competitive event. Patient discharged earlier than 30 days to home or to low-intensity care setting will be considered without ventilation at the end-date of 30 days.

For time to definitive extubation, patients without invasive mechanical ventilation within 24 h from registration date will be excluded from the analysis and death will be considered as competitive event. Patient discharge to home or to low-intensity care setting will be considered as event when data of definitive extubation is missing.

For time to independence from non-invasive mechanical ventilation, only patients with non-invasive mechanical ventilation within 24 h from registration date will be included in the analysis and death will be considered as competitive event. Patient discharge will be considered as event if data of independence from non-invasive mechanical ventilation is missing.

For time to independence from oxygen therapy, all patients will be included in the analysis (due to eligibility criteria) and death will be considered as competitive event. Patient discharge to home or to low-intensity care (in both cases with independence from oxygen therapy) will be considered as event when data of independence from oxygen therapy is missing.

4.4.4. Longitudinal analysis

Longitudinal evaluations of clinical and laboratory values will be summarized at specific time-points. Depending on the type of variable, summary estimates (e.g. proportion, mean or median) will be calculated and reported at specific time-points. Mixed-effect regression models will be used to analyze the associations including longitudinal evaluations as dependent variable and time and the baseline characteristics as covariates. Interaction between time and the other covariates included will be tested. Mixed effects models use all available data over follow-up and can properly account for correlation between repeated measures.

Duration of hospitalization will be summarized as median and interquartile range.

4.4.5. Additional analyses

In the multivariable regression models, a sensitivity analysis could be performed by using a propensity score approach if the number of deaths will not be sufficient to allow the high number of covariates in multivariable models. The propensity score, defined as the probability of receiving treatment, is estimated by a logistic regression model with the baseline characteristics as covariates. The C-statistic of the propensity score, that indicates the degree to which the propensity score model discriminates between treated patients and untreated patients, will be calculated. The estimated propensity score will be added as a covariate in the regression models together with treatment information (status and time). This approach increases flexibility when the outcome is rare, and treatment is common allowing to adequately adjust for all baseline variables.

4.5. Missing data

A high rate of missing data is expected for several administrative reasons:

- (i) Clinical researchers are overburdened because of the increasing number of patients (the trial started during the ascending phase of the pandemic);
- (ii) Data managers and other supportive personnel are lacking in many participating centers, also due to restrictions consequent to pandemic;
- (iii) Some centers seem poorly motivated and participated to the trial just because it was the only chance to use the drug;
- (iv) In several COVID units clinical charts have been locked down because of risk of paper-mediated virus transmission and are not accessible for external use for several days
- (v) Researchers had a limited training on the use of the platform due to the short time between protocol approval and trial start (this was partially counteracted with publication on the web site of tutorial documents);
- (vi) Few fields in web-CRFs were mandatory to simplify the job of researchers and to reduce slowdowns of data input;
- (vii) No on-site monitoring is possible owing to pandemic confinement (central monitoring was performed by telephone, e-mail, and e-queries through the web-based platform);
- (viii) Hardware temporarily crashed twice due to extreme and unexpected information loading volume.

Hence, in agreement with IDMC, it was decided to remove from any analysis all patients recruited from uncooperative centers providing less than 25% of information on baseline characteristics and treatment administration. We had no information on reasons why the uncooperative centers provided so few data, but we cannot exclude that it was related to baseline (or outcome) patient's characteristics. Thus, rather than using only patients for which information was available, we assumed that information was unreliable for all subjects in that center, and, with the agreement of the IDMC, we decided to remove the whole center from all analyses. It was agreed that a boundary of 25% was low enough to still permit the participation of a sufficient number of patients.

All other patients are included in primary ITT analyses and frequency distributions of missing data will be reported for all variables assessed in the trial.

4.6. Safety analysis

All safety analyses will be conducted on the safety phase 2 population and safety validation population. In terms of safety, treatment emergent AEs and SAEs will be summarized with all safety data available. No formal hypothesis testing is planned.

4.7. Ethics

TOCIVID-19 trial is a no-profit study promoted by the National Cancer Institute of Napoli. The TOCIVID-19 protocol was first approved by the National Ethical Committee at the Lazzaro Spallanzani Institute on March 18th, 2020 [14]. Two amendments followed on March 24th, 2020 and April 28th, 2020: the March 24th amendment mainly addressed the problems raised by the very fast enrolment and drug shortage, while the second one essentially modified statistical analysis, introducing the 14-days co-primary endpoint and defining the efficacy populations for ITT analyses.

Authorship statement

PC and CG drafted the statistical analysis plan. CG is the senior statistician and supervises all statistical aspects of the trial. PC and LA are responsible for data quality assessment. MCP is the study coordinating physician. FP is the senior chief investigator, conceived the project and developed the protocol. All authors provided feedback on drafts of this paper and read and approved the final manuscript.

Funding

The trial sponsor is the Istituto Nazionale dei Tumori di Napoli. No specific funding was available. The study drug was provided by the pharmaceutical company (Roche) free of charge.

Declaration of competing interest

PC, LA and CG have no competing interests. FP and MCP coordinate three academic clinical trials in oncology, promoted by the Istituto Nazionale Tumori di Napoli, that are supported by Roche (clinicaltrials.gov id: NCT01706120, NCT01802749, NCT02633189).

References

- [1] Boccia S, Ricciardi W, Ioannidis JPA. What other countries can learn from Italy during the COVID-19 Pandemic. *JAMA Intern Med* Published online April 7, 2020.
- [2] P. Pasquariello, S. Stranges, Excess mortality from COVID-19: a commentary on the Italian experience, *Int. J. Publ. Health* 65 (5) (2020) 529–531.
- [3] Xiaoling Xu, Mingfeng Han, Tiantian Li et al. Effective Treatment of Severe COVID-19 Patients with Tocilizumab. *ChinaXiv*: 202003.00026v1, now published in PNAS <https://doi.org/10.1073/pnas.2005615117>.
- [4] Xiaoling Xu (PI). A multicenter, randomized controlled trial for the efficacy and safety of tocilizumab in the treatment of new coronavirus pneumonia (COVID-19). Registration number on ChiCTR: ChiCTR2000029765 of February 13 2020.
- [5] C. Gamble, A. Krishan, D. Stocken, et al., Guidelines for the content of statistical analysis plans in clinical trials, *J. Am. Med. Assoc.* 318 (23) (2017) 2337–2343.
- [6] <http://www.epicentro.iss.it/coronavirus/>. (Accessed 10 April 2020).
- [7] http://www.salute.gov.it/imgs/C_17_notizie_4454_0_file.pdf. (Accessed 13 April 2020).
- [8] P. Giorgi Rossi, E. Ferroni, S. Spila Alegiani, O. Leoni, G. Pitter, D. Cereda, M. Marino, M. Pellizzari, J. Sultana, G. Trifirò, M. Massari, the ITA-COVID19 working group, Survival of hospitalized COVID-19 patients in Northern Italy: a population-based cohort study by the ITA-COVID19 Network, *medRxiv* (2020), <https://doi.org/10.1101/2020.05.15.20103119>.
- [9] International council on harmonization of technical requirements for registration of pharmaceuticals for human use. Addendum on estimands and sensitivity analysis in clinical trials to the guideline on statistical principles for clinical trials E9(R1) available at https://database.ich.org/sites/default/files/E9-R1_Step4_Guideline_2019_1203.pdf.
- [10] K.F. Schulz, D.G. Altman, D. Moher, CONSORT Group, CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials, *BMJ* 340 (2010) c332.
- [11] D.Y. Lin, Non-parametric inference for cumulative incidence functions in competing risks studies, *Stat. Med.* 16 (8) (1997) 901–910.
- [12] D.P. Bernasconi, P. Rebora, S. Iacobelli, et al., Survival probabilities with time-dependent treatment indicator: quantities and non-parametric estimators, *Stat. Med.* 35 (2016) 1032–1048.
- [13] I. Poguntke, M. Schumacher, J. Beyersmann, M. Wolkewitz, Simulation shows undesirable results for competing risks analysis with time-dependent covariates for clinical outcomes, *BMC Med. Res. Methodol.* 18 (2018) 79.
- [14] <https://www.aifa.gov.it/web/guest/-/gestione-degli-studi-clinici-in-italia-in-corso-di-emergenza-covid-19-coronavirus-disease-19->