Cohort Multiple randomized controlled trials open-label of immune modulatory drugs and other treatments in COVID-19 patients

General Statistical Analysis Plan for CORIMUNO-19 TOCIDEX Trial

Version 1.0

November 16, 2020

Redacted by Raphaël Porcher and Gabriel Baron
Validated by Philippe Ravaud

Coordinating Investigator:	**Pr Olivier Hermine,** Service Hématologie Adulte Necker Hospital, Paris, France Tel: +33 144495282 +33603707920 E-mail: olivier.hermine@aphp.fr
Co-coordinating Investigator	**Pr Xavier Mariette,** Service Rhumatologie Hôpital du Kremlin Bicêtre Tel: +33145213758 +33623268104 E-mail: xavier.mariette@aphp.fr **Immune COVID 19 group**
Scientific Director:	**Dr Pierre Louis Tharaux** INSERM U970, PARCC HEGP Tel +33689502948 E-mail: pierre-louis.tharaux@inserm.fr

Methodologist and statisticians:	**Pr Philippe Ravaud** Centre d'Epidémiologie Clinique. Hôpital Hôtel Dieu 1 Place du Parvis Notre Dame 75004 Paris E-mail: philippe.ravaud@aphp.fr **Pr Raphael Porcher** Centre d'Epidémiologie Clinique. Hôpital Hôtel Dieu 1 Place du Parvis Notre Dame 75004 Paris E-mail: raphael.porcher@aphp.fr **Pr Matthieu Resche-Rigon** Hôpital Saint-Louis 1 av Claude vellefaux 75010 Paris E-mail: matthieu.resche-rigon@u-paris.fr
Sponsor : AP-HP and by delegation:	**Délégation à la Recherche Clinique et à l'Innovation (DRCI)** Saint-Louis Hospital 1, avenue Claude Vellefaux, Paris, France DRCD-Siège : Cécile Kedzia - E-mail: cecile.kedzia@aphp.fr Damien Vanhoye - E-mail: damien.vanhoye@aphp.fr
Monitoring and data Management:	Trial monitoring URC Hôpitaux Universitaires de l'Est Parisien Hôpital Saint-Antoine 184 rue du Faubourg Saint Antoine 75571 Paris cedex 12 **Pr Tabassome Simon** E-mail: tabassome.simon@aphp.fr Data management URC Lariboisière Saint-Louis Site Saint - Louis Hôpital Saint-Louis, 1 av Claude Vellefaux, 75010 Paris **Pr Matthieu Resche-Rigon** E-mail: matthieu.resche-rigon@u-paris.fr

1 Summary

	CORIMUNO-19 – TOCIDEX (TOCI + Dexamethasone Versus Dexamethasone)
Rationale for using Tocilizumab + Dexamethasone and Dexamethasone in severe patients infected with COVID-19	**CORIMUNO-19 - TOCIDEX** The SRAS-CoV-S protein induces direct up-regulation of IL-6, IL-1 and TNFα, some of the most potent pro-inflammatory cytokines. Dexamethasone (DXM) is a steroid with a high anti-inflammatory activity. In the RECOVERY study, 2104 patients randomly allocated to receive dexamethasone were compared with 4321 patients concurrently allocated to usual care. Overall, 454 (21.6%) patients allocated dexamethasone and 1065 (24.6%) patients allocated usual care died within 28 days (age-adjusted rate ratio [RR] 0.83; 95% confidence interval [CI] 0.74 to 0.92; P<0.001). The proportional and absolute mortality rate reductions varied significantly depending on level of respiratory support at randomization (test for trend p<0.001): Dexamethasone reduced deaths by one-third in patients receiving invasive mechanical ventilation (29.0% vs. 40.7%, RR 0.65 [95% CI 0.51 to 0.82]; p<0.001), by one-fifth in patients receiving oxygen without invasive mechanical ventilation (21.5% vs. 25.0%, RR 0.80 [95% CI 0.70 to 0.92]; p=0.002), but did not reduce mortality in patients not receiving respiratory support at randomization (17.0% vs. 13.2%, RR 1.22 [95% CI 0.93 to 1.61]; p=0.14). Therefore, in patients hospitalized with COVID-19, DXM reduced 28-day mortality among those receiving invasive mechanical ventilation or oxygen support at randomization, but not among patients not receiving respiratory support. These data provide strong evidence that DXM could become the SOC, since it is the only drug tested in a randomized way that showed improvement of survival. Tocilizumab (TCZ) is an anti-human IL-6 receptor monoclonal antibody that inhibits signal transduction by binding sIL-6R and mIL-6R. The main approved indication is for rheumatoid arthritis, in association or not with methotrexate. The IV approved dose in RA is 8 mg/kg every month. TCZ is also approved in the treatment of juvenile inflammatory arthritis and in the treatment of refractory giant cell arteritis. Interestingly, this later indication concerns aged patients and, in this population, the safety profile was the same as in younger patients. In 2017, the U.S. Food and Drug Administration approved TCZ for the treatment of cytokine release syndrome (CRS) consisting in a systemic inflammatory response caused by the massive release of pro-inflammatory cytokines in response to iatrogenic (e.g. CAR-t therapies) or infective stimuli. In our previous study that randomized TCZ versus standard of Care (SOC), we have shown that TCZ was superior to SOC. 131 patients were randomized to receive SOC (n=67) or SOC + TCZ (n=64). The posterior probability of a reduction of non-invasive or mechanical ventilation or death at day 14 with TCZ was 95.0%, thus achieving predefined efficacy threshold (posterior median hazard ratio (HR) [90% credible interval], 0.58 [0.33-1.00]). Reduction of mechanical ventilation or death was of the same magnitude: HR [90% CrI], 0.58 [0.30-1.09]. With a median of follow-up of 28 days, seven deaths (11.1%) were observed in the TCZ group and 11 (16.4%) in the SOC group. No increase in serious adverse events was observed in the TCZ arm. In addition, none of the patients who received the combination of DXM + TCZ (n=10) experienced either death or ventilation support, suggesting that this combination might improve results obtained with TCZ. Interestingly, no increase of infectious events was observed.

	Therefore, based on these two studies, we could define a new SOC with DXM and test whether or not the combination of DXM and TCZ improve outcome of patients with severe COVID-19.
Diagnosis and inclusion and Exclusion criteria for the Tocilizumab trial	**Inclusion Criteria for the TOCI-DEX trial:** 1. Patients included in the CORIMUNO-19 cohort 2. Patients belonging to the following group: - *Group 1: Cases meeting all of the following criteria* • *Requiring ≥ 3L/min of oxygen* • *WHO progression scale = 5* • *No NIV or High flow* **Exclusion Criteria for the TOCIDEX trial:** ● Patients with exclusion criteria to the CORIMUNO-19 cohort. ● Known hypersensitivity to Tocilizumab or to any of their excipients. ● Known hypersensitivity to Dexamethasone or to any of their excipients. ● Pregnancy ● Current documented bacterial infection ● certain evolving viral diseases (especially active herpes, chickenpox or shingles), ● psychotic states still not controlled by treatment, ● live vaccines, ● Patient with any of following laboratory results out of the ranges detailed below at screening should be discussed depending of the medication: ○ Absolute neutrophil count (ANC) ≤ 1.0 x 10^9/L ○ Hemoglobin level: no limitation ○ Platelets (PLT) < 50 G /L SGOT or SGPT > 5N
Randomisation and Treatment procedures	Within this group all consecutive patients meeting the inclusion criteria will be randomized 1:1 either in the TCZ + DXM arm or in the new defined standard of care (Soc) control arm containing DXM in a set of 60 patients in total (30 in the TCZ + DXM arm, 30 in the control DXM arm). Interim analysis will be carried out, and if the interim analysis indicates to continue the subtrial, a new set of 60 patients will be included on the same basis (30 in the TCZ + DXM arm, 30 in the control DXM arm). Inclusions of new sets will stop when statistical analyses conclude on futility or efficacy or by DSMB decision.
Duration of follow-up	90 days
Criteria for efficacy	**Measures** A core set of clinical measures will be recorded daily the first 2 weeks and then every week. The core measures include measures of WHO progression scale, oxygenation, mechanical and supportive ventilation (Optiflow and NIV). For patients who are eligible for an intervention trial (in both the intervention and control arms), this daily measurement will include trial-specific measures related to the trial outcomes of interest. **Primary endpoint:** Survival without needs of mechanical ventilation at day 14. A new DNR order will be considered as an event at the date of the DNR.

Secondary Endpoints:

1. WHO Ordinal Scale at day 7 and day 14

The 10-point scale is defined as follow:

OMS 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 OR SpO2/FIO2>=200	7
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

2. Overall survival at 14, 28, 60 and 90 days
3. Survival without needs of ventilator utilization (including nn-invasive ventilation and Optiflow) at day 14.
4. Cumulative incidence of discharge alive at 14 and 28 days
5. Cumulative incidence of oxygen supply independency at 14 and 28 days

Exploratory outcomes;
Biological parameters improvement including CRP, neutrophil and lymphocytes counts

Criteria of safety

- Number of serious adverse events (SAEs)
- Number of Grade 3 and 4 AEs.

Investigational medication discontinuation (for any reason)

Statistical Method	Bayesian monitoring and analysis of the trial will be used. The primary outcome will be therefore analyzed using a Bayesian Cox model adjusted for age. The treatment effect will be summarized in terms of hazard ratio (HR) for the experimental vs. control arm. Every 60 patients randomized (30 in each arm) reach a minimal 7 days follow-up, an interim analysis is triggered, where several posterior probabilities will be calculated: 1) posterior probability of benefit $P_1 = P(HR < 1 \mid data)$; 2) posterior probability of at least a fair benefit $P_2 = P(HR < 0.8 \mid data)$, 3) posterior probability of inefficacy or harm $P_3 = P(HR > 1 \mid data)$. At the interim analysis, the trial can be stopped for futility if $P_2 < 0.10$ or $P_3 > 0.80$. At the first interim analysis, the trial can be stopped for efficacy if $P_1 > 0.99$, the. At the subsequent interim analyses, **efficacy boundaries are set to $P_1 > 0.95$ or $P_2 > 0.80$**. Decision boundaries are non-binding, and the DSMB can recommend continuing recruitment or stopping even if the aforementioned boundaries are not crossed. Overall, the trial is planned for a **total sample size of 660 participants** (330 per arm).

2 Major amendments to the protocol

An amendment has been submitted and approved on 2020-11-16, to modify the sample size (extending Bayesian interim analyses every other 60 patients randomized), and specifying survival without needs of invasive ventilation at day 14 as primary outcome.
In the original protocol, the maximum sample size was 180 patient, with a Bayesian interim analysis after 60 and 120 patients recruited, and the primary outcome was survival without needs of ventilator utilization (including non-invasive ventilation and Optiflow) at day 14.
Owing to the following reasons, the maximum sample size was increased up to 660, and the primary outcome changed to survival without needs of invasive ventilation at day 14:

1. Scientifically and medically, survival without mechanical ventilation better reflects the most severe worsening of patients' condition in the target population, and has been more often used in other trials compared to the original outcome.
2. Operationally, the experience of the CORIMUNO trials conducted until now has shown that the use of non-invasive ventilation and high-flow oxygen devices has been more difficult to record due to the large heterogeneity of the devices used.

Survival without needs of ventilator utilization (including non-invasive ventilation and Optiflow) was added to secondary outcomes.

The design was also modified as follows:
While we keep the core principles and stopping rules of the design, there are several issues to be considered, in the light of preceding CORIMUNO trials and other trials published or registered:

1. Targeting a very large effect as the current TOCIDEX trial (powered for HRs between 0.30 and 0.46) is likely too optimistic, and smaller effect sizes (e.g. HRs of 0.60–0.70) clearly represent important benefits for patients;
2. Small trials (total sample size < 200) could be justified in order to provide quick answers during the first wave of the pandemics, but are now not regarded as providing robust answers at the international level;
3. Switching the outcome to MV or death instead of any form of ventilator utilization or death implies a decrease in the expected event rate, thereby requiring a reassessment of the sample size.

Using the results of previous CORIMUNO trials (CORIMUNO-TOCI and CORIMUNO-ANA), we hypothesized that the proportion of patients with MV or death at day 14 would be 25%. A frequentist sample size calculation suggests that a total sample size of 634 patients is necessary to demonstrate a HR of 0.65 with power 80%, using a one-sided 5% type I error rate. Keeping the strategy of analysis every 60 patients of CORIMUNO trials, an indicative maximum sample size would be 660 (330 per arm).

The previous Bayesian design is then extended to analysis every 60 patients, triggered when data are available for at least 7 follow-up days for the first 60 patients.
At each analysis, the three previously planned posterior probabilities are computed:

- Posterior probability of benefit $P_1 = P(HR < 1 \mid data)$;
- Posterior probability of at least a fair benefit $P_2 = P(HR < 0.8 \mid data)$;
- Posterior probability of inefficacy or harm $P_3 = P(HR > 1 \mid data)$.

At each analysis, the following actions are triggered according to the thresholds given below, adapted from the Statistical Design and Analysis Plan for Sequential Parallel-Group RCF for COVID-19 (Harrell & Lindsell, 2020. http://hbiostat.org/proj/covid19/bayesplan.html):

- Stop with evidence for efficacy if $P_1 > 0.95$ ($P_1 > 0.99$ at the first analysis because we consider that there is a need for very convincing evidence if the sample size is very limited, and this was already planned in the previous version of the protocol);
- Stop for futility if $P_2 < 0.10$ or $P_3 > 0.80$;
- Stop with evidence for efficacy if $P_2 > 0.80$ (only actionable when at least 180 patients have been randomized, i.e. starting from the third analysis).

3 Analysis population

3.1 Flow diagram

At the final analysis of trial, a flow chart will be constructed according to the CONSORT 2010 reporting guidelines. It will describe:

- The number of eligible patients, randomized patients and the number of patients who have actually followed the study;
- The intervention arm allocated per randomization;
- Early cessation of the intervention and their causes and drop-outs;
- The number of patients excluded from the analysis.

3.2 Definition of the analysis population

The final analysis will be carried out according to the intention to treat (ITT) principle, i.e. each randomised participant will be analysed in the group assigned to him/her by randomisation, regardless of the actual treatment received or other protocol deviations. In particular patients randomised while not meeting eligibility criteria will be kept in the analysis. For interim analyses, only patients with at least 7 days theoretical follow-up and data recorded for the first follow-up visit (7 days) will be analysed.
No data will be analysed for patients who have withdrawn their consent during the study and have expressed opposition to the analysis of their data. If necessary, the data concerning these patients that have been collected will be destroyed. The existence of these patients will nevertheless be documented in the study flow chart.

3.3 Sample size

The maximum sample size has been fixed for the whole trial at 660 (330 per arm), with a Bayesian analysis every 60 patients.
This sample size was based on considering that a frequentist sample size calculation would suggest that a total sample size of 634 patients is necessary to demonstrate a HR of 0.65 with power 80%, using a one-sided 5% type I error rate, and keeping the strategy of analysis every 60 patients of CORIMUNO trials.

4 Analysis principles

4.1 General principles for analysis of outcomes

The final results will be reported according to the recommendations of CONSORT 2010.
All outcomes will be analysed in superiority analyses, and the analyses will be adjusted for age and centre (randomisation stratification variable), the latter as a random effect.
The primary efficacy analyses will rely on computing the posterior distribution of the adjusted hazard ratio between the experimental and control arms for the primary outcome, using a proportional hazards model. These posterior distributions will be graphically displayed, and summarized by their medians and two-sided 90% credibility intervals (the Bayesian counterparts of confidence intervals).

For secondary efficacy and safety outcomes, frequentist (i.e. non-Bayesian) analyses will be used. No correction for multiplicity and no hierarchical testing procedures are planned in analysing secondary outcomes. These analyses will therefore be considered as exploratory in nature.

4.2 Participants' characteristics at inclusion

The characteristics of patients collected at inclusion will be described globally and by randomization group, using means, standard deviations, medians, interquartile intervals, minimum and maximum for quantitative variables and by their numbers and percentages by modality for qualitative variables.
The number of missing data for each variable will also be reported. No statistical tests for comparison between groups will be carried out.

4.3 Handling of missing or incoherent data

Given their nature and the trial settings, it is not be expected that primary outcome data would be missing. In the case of a follow-up shorter than 14 days, they will be naturally handled using methods for censored data. No imputation will be used for secondary efficacy and safety outcomes.

4.4 Statistical software

The analyses will be carried out using the R software version 4.0.1 or later (The R Foundation for Statistical Computing, Vienna, Austria), SAS version 9.4 or later (SAS Institute Cary, NC) and JAGS version 4.3.0 or later.

5 Primary outcome analysis

5.1 Definition of the primary outcome

The primary outcome is survival without needs of invasive ventilation at day 14.

5.2 Outcome analysis

5.2.1 Modelling

This section describes the Bayesian analysis of the primary outcome used for trial monitoring and final analysis.
A Bayesian Cox model will be estimated using Markov chain Monte Carlo (MCMC) methods, adjusted for age and centre (modelled as a random effect). The primary analysis will use a flat prior, and different sceptical or enthusiastic priors will be used as sensitivity analyses (see specification of the priors in the §5.2.2 below). In addition, a frequentist Cox model (adjusted for age and with random centre effect) will also be used.

5.2.2 Settings for Monte Carlo Markov Chain Bayesian analyses

The main analysis will use a Gaussian prior distribution with mean 0 and variance 10^2 for the log hazard ratio. The prior for the log hazard ratio for age will also be a Gaussian prior, with mean 0 and variance 10^2. Four different chains with different starting values will be run, with a burn-in of 10,000 iterations, and 100,000 additional iterations and a thinning interval of 10, leading to keeping 10,000 values per chain, 40,000 in total. The convergence of the models will be assessed using the Gelman-Rubin statistic and by visual inspection of the trace of coefficients
As a sensitivity analysis, we will investigate different prior distributions, namely two sceptic priors centred on 0 with variance set so that a $P(HR < 0.2) = P(HR > 5) = 0.05$ (SD 0.975) or $P(HR < 0.2) = P(HR > 5) = 0.025$ (SD 0.82), and one enthusiastic informative prior centred on the targeted treatment effect (HR of 0.65) and informative with $\sigma = 0.975$.

5.2.3 Presentation of results

The posterior distribution of the hazard ratio will be displayed, and summarized by its median and two-sided 90% and 95% credibility intervals. Kaplan-Meier plots or cumulative incidence of the outcome will also be estimated in each arm, in a frequentist approach. Posterior probabilities of any benefit and at least a fair benefit will also be presented.

5.3 Stopping rules

At each interim analysis, the posterior distribution of the hazard ratio θ will be used to compute different posterior probabilities:

- Posterior probability of any benefit $P_1 = P(\theta < 1 \mid data)$;
- Posterior probability of at least a fair benefit $P_2 = P(\theta < 0.8 \mid data)$;
- Posterior probability of inefficacy or harm $P_3 = P(\theta > 1 \mid data)$.

At each analysis, the following actions are triggered according to the thresholds given below, adapted from the Statistical Design and Analysis Plan for Sequential Parallel-Group RCF for COVID-19 (Harrell & Lindsell, 2020. http://hbiostat.org/proj/covid19/bayesplan.html):

- Stop with evidence for efficacy if $P_1 > 0.95$ ($P_1 > 0.99$ at the first analysis because we consider that there is a need for very convincing evidence if the sample size is very limited, and this was already planned in the previous version of the protocol);
- Stop for futility if $P_2 < 0.10$ or $P_3 > 0.80$;
- Stop with evidence for efficacy if $P_2 > 0.80$ (only actionable when at least 180 patients have been randomized, i.e. starting from the third analysis).

Any decision to stop or continue recruitment will be advised by the DSMB based of the aforementioned posterior probabilities, as well as safety data.

5.4 Calculation of the outcome

The day of randomisation will be counted as day 1. The time to mechanical ventilation or death will be computed starting from day one up to day 14 included.
A close data monitoring will be carried out to limit missing information on the use of ventilation as much as possible.
A new Do-Not-Resuscitate (DNR) order, i.e. a DNR order posterior to the date of randomisation and that has been noted as having been effectively used to limit care will be considered as an event for the primary outcome, at the date of limitation of care.

5.5 Subgroup analyses

The initial protocol specified that, at the end of the study, subgroup analyses would be performed according to antiviral therapies at baseline. Given the lack of efficacy of antivirals in the trials, no such analysis will be carried out.

6 Secondary efficacy outcomes analysis

6.1 Definitions

- WHO progression scale at 7 and 14 days;
- Overall survival at 14, 28, 60 and 90 days;
- Time to discharge at 14 and 28 days;
- Time to oxygen supply independency at 14 and 28 days.

Biological parameters improvement including CRP, neutrophil and lymphocytes counts are not secondary outcomes, but exploratory outcomes.

6.2 Methods for analysis

6.2.1 Time-to-event outcomes

Time-to-event outcomes will be analysed using Cox or Fine-Gray regression models adjusted for the same variables as the primary outcome; results will be expressed as hazard ratios with 95% confidence interval. Competing risks analyses (Fine-Gray model) will be used for time to discharge, and time to oxygen supply independency, for which death will be considered as a competing event. When several timepoints are mentioned, separate models will be estimated at these timepoints. Point estimates of survival in each arm will be presented together with Kaplan-Meier survival curves.

6.2.2 WHO ordinal scale

For the WHO ordinal scale, a proportional odds models will be used to compare the distribution of ordinal scores at day 7 and at day 14, adjusted for age and centre. The distribution of scores will be described at 7, and 14 days.

6.2.3 Biological and physiological outcomes

For biological outcomes, only descriptive analyses will be performed.

7 Safety analysis

7.1 Definitions

Adverse events are spontaneously declared on the CRF. For each adverse event, the following information is collected:

- Classification of the adverse event (AE) as a serious adverse event (SAE);
- Seriousness criteria for SAEs;
- Intensity (severity): mild, moderate or severe;
- Start/end dates;
- Investigator judgement on relationship with the study treatment, concomitant treatment, pre-existing disease and COVID-19;
- Modification of study treatment;
- Symptomatic treatment;
- Outcome.

Moreover, major safety endpoints are monitored: blood cells and platelets counts and liver transaminases, are monitored frequently, every three days systematically:

- Neutrophil count;
- Platelet count;
- Liver enzymes: ALT and AST;
- Occurrence of skin rashes;
- Systolic and diastolic blood pressure;
- Ventilator asynchronization.

7.2 Analysis

Adverse events and their characteristics will be described using numbers and percentages per treatment arm. The proportion of participants with each of the reported events, as well as the proportions of participants with at least one SAE will be compared using Fisher's exact tests. The total number of AE/SAEs and SAEs will also be described for each arm, and compared using Poisson models (with a robust error variance if necessary).

8 Summary of changes since previous versions

This Statistical Analysis Plan was developed from the Statistical Analysis Plan version 2.1 of previous CORIMUNO-19 trials.
Subsequent changes to the SAP will be summarised here whenever relevant.

Cohort Multiple randomized controlled trials open-label of immune modulatory drugs and other treatments in COVID-19 patients

General Statistical Analysis Plan for CORIMUNO-19 TOCIDEX Trial

Version 1.2

July 1, 2021

Redacted by Raphaël Porcher and Gabriel Baron
Validated by Philippe Ravaud

Coordinating Investigator:	**Pr Olivier Hermine,** Service Hématologie Adulte Necker Hospital, Paris, France Tel: +33 144495282 +33603707920 E-mail: olivier.hermine@aphp.fr
Co-coordinating Investigator	**Pr Xavier Mariette,** Service Rhumatologie Hôpital du Kremlin Bicêtre Tel: +33145213758 +33623268104 E-mail: xavier.mariette@aphp.fr **Immune COVID 19 group**
Scientific Director:	**Dr Pierre Louis Tharaux** INSERM U970, PARCC HEGP Tel +33689502948 E-mail: pierre-louis.tharaux@inserm.fr

Methodologist and statisticians:	**Pr Philippe Ravaud** Centre d'Epidémiologie Clinique. Hôpital Hôtel Dieu 1 Place du Parvis Notre Dame 75004 Paris E-mail: philippe.ravaud@aphp.fr **Pr Raphael Porcher** Centre d'Epidémiologie Clinique. Hôpital Hôtel Dieu 1 Place du Parvis Notre Dame 75004 Paris E-mail: raphael.porcher@aphp.fr **Pr Matthieu Resche-Rigon** Hôpital Saint-Louis 1 av Claude vellefaux 75010 Paris E-mail: matthieu.resche-rigon@u-paris.fr
Sponsor : AP-HP and by delegation:	**Délégation à la Recherche Clinique et à l'Innovation (DRCI)** Saint-Louis Hospital 1, avenue Claude Vellefaux, Paris, France DRCD-Siège : Cécile Kedzia - E-mail: cecile.kedzia@aphp.fr Damien Vanhoye - E-mail: damien.vanhoye@aphp.fr
Monitoring and data Management:	Trial monitoring URC Hôpitaux Universitaires de l'Est Parisien Hôpital Saint-Antoine 184 rue du Faubourg Saint Antoine 75571 Paris cedex 12 **Pr Tabassome Simon** E-mail: tabassome.simon@aphp.fr Data management URC Lariboisière Saint-Louis Site Saint - Louis Hôpital Saint-Louis, 1 av Claude Vellefaux, 75010 Paris **Pr Matthieu Resche-Rigon** E-mail: matthieu.resche-rigon@u-paris.fr

1 Summary

	CORIMUNO-19 – TOCIDEX (TOCI + Dexamethasone Versus Dexamethasone)
Rationale for using Tocilizumab + Dexamethasone and Dexamethasone in severe patients infected with COVID-19	**CORIMUNO-19 - TOCIDEX** The SRAS-CoV-S protein induces direct up-regulation of IL-6, IL-1 and TNFα, some of the most potent pro-inflammatory cytokines. Dexamethasone (DXM) is a steroid with a high anti-inflammatory activity. In the RECOVERY study, 2104 patients randomly allocated to receive dexamethasone were compared with 4321 patients concurrently allocated to usual care. Overall, 454 (21.6%) patients allocated dexamethasone and 1065 (24.6%) patients allocated usual care died within 28 days (age-adjusted rate ratio [RR] 0.83; 95% confidence interval [CI] 0.74 to 0.92; P<0.001). The proportional and absolute mortality rate reductions varied significantly depending on level of respiratory support at randomization (test for trend p<0.001): Dexamethasone reduced deaths by one-third in patients receiving invasive mechanical ventilation (29.0% vs. 40.7%, RR 0.65 [95% CI 0.51 to 0.82]; p<0.001), by one-fifth in patients receiving oxygen without invasive mechanical ventilation (21.5% vs. 25.0%, RR 0.80 [95% CI 0.70 to 0.92]; p=0.002), but did not reduce mortality in patients not receiving respiratory support at randomization (17.0% vs. 13.2%, RR 1.22 [95% CI 0.93 to 1.61]; p=0.14). Therefore, in patients hospitalized with COVID-19, DXM reduced 28-day mortality among those receiving invasive mechanical ventilation or oxygen support at randomization, but not among patients not receiving respiratory support. These data provide strong evidence that DXM could become the SOC, since it is the only drug tested in a randomized way that showed improvement of survival. Tocilizumab (TCZ) is an anti-human IL-6 receptor monoclonal antibody that inhibits signal transduction by binding sIL-6R and mIL-6R. The main approved indication is for rheumatoid arthritis, in association or not with methotrexate. The IV approved dose in RA is 8 mg/kg every month. TCZ is also approved in the treatment of juvenile inflammatory arthritis and in the treatment of refractory giant cell arteritis. Interestingly, this later indication concerns aged patients and, in this population, the safety profile was the same as in younger patients. In 2017, the U.S. Food and Drug Administration approved TCZ for the treatment of cytokine release syndrome (CRS) consisting in a systemic inflammatory response caused by the massive release of pro-inflammatory cytokines in response to iatrogenic (e.g. CAR-t therapies) or infective stimuli. In our previous study that randomized TCZ versus standard of Care (SOC), we have shown that TCZ was superior to SOC. 131 patients were randomized to receive SOC (n=67) or SOC + TCZ (n=64). The posterior probability of a reduction of non-invasive or mechanical ventilation or death at day 14 with TCZ was 95.0%, thus achieving predefined efficacy threshold (posterior median hazard ratio (HR) [90% credible interval], 0.58 [0.33-1.00]). Reduction of mechanical ventilation or death was of the same magnitude: HR [90% CrI], 0.58 [0.30-1.09]. With a median of follow-up of 28 days, seven deaths (11.1%) were observed in the TCZ group and 11 (16.4%) in the SOC group. No increase in serious adverse events was observed in the TCZ arm. In addition, none of the patients who received the combination of DXM + TCZ (n=10) experienced either death or ventilation support, suggesting that this combination might improve results obtained with TCZ. Interestingly, no increase of infectious events was observed.

	Therefore, based on these two studies, we could define a new SOC with DXM and test whether or not the combination of DXM and TCZ improve outcome of patients with severe COVID-19.
Diagnosis and inclusion and Exclusion criteria for the Tocilizumab trial	**Inclusion Criteria for the TOCI-DEX trial:** 1. Patients included in the CORIMUNO-19 cohort 2. Patients belonging to the following group: - *Group 1: Cases meeting all of the following criteria* • *Requiring ≥ 3L/min of oxygen* • *WHO progression scale = 5* • *No NIV or High flow* **Exclusion Criteria for the TOCIDEX trial:** ● Patients with exclusion criteria to the CORIMUNO-19 cohort. ● Known hypersensitivity to Tocilizumab or to any of their excipients. ● Known hypersensitivity to Dexamethasone or to any of their excipients. ● Pregnancy ● Current documented bacterial infection ● certain evolving viral diseases (especially active herpes, chickenpox or shingles), ● psychotic states still not controlled by treatment, ● live vaccines, ● Patient with any of following laboratory results out of the ranges detailed below at screening should be discussed depending of the medication: ○ Absolute neutrophil count (ANC) ≤ 1.0 x 10^9/L ○ Hemoglobin level: no limitation ○ Platelets (PLT) < 50 G /L SGOT or SGPT > 5N
Randomisation and Treatment procedures	Within this group all consecutive patients meeting the inclusion criteria will be randomized 1:1 either in the TCZ + DXM arm or in the new defined standard of care (Soc) control arm containing DXM in a set of 60 patients in total (30 in the TCZ + DXM arm, 30 in the control DXM arm). Interim analysis will be carried out, and if the interim analysis indicates to continue the subtrial, a new set of 60 patients will be included on the same basis (30 in the TCZ + DXM arm, 30 in the control DXM arm). Inclusions of new sets will stop when statistical analyses conclude on futility or efficacy or by DSMB decision.
Duration of follow-up	90 days
Criteria for efficacy	**Measures** A core set of clinical measures will be recorded daily the first 2 weeks and then every week. The core measures include measures of WHO progression scale, oxygenation, mechanical and supportive ventilation (Optiflow and NIV). For patients who are eligible for an intervention trial (in both the intervention and control arms), this daily measurement will include trial-specific measures related to the trial outcomes of interest. **Primary endpoint:** Survival without needs of mechanical ventilation at day 14. A new DNR order will be considered as an event at the date of the DNR.

Secondary Endpoints:

1. WHO Ordinal Scale at day 7 and day 14

The 10-point scale is defined as follow:

OMS 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 OR SpO2/FIO2>=200	7
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

2. Overall survival at 14, 28, 60 and 90 days
3. Survival without needs of ventilator utilization (including non-invasive ventilation and Optiflow) at day 14.
4. Cumulative incidence of discharge alive at 14 and 28 days
5. Cumulative incidence of oxygen supply independency at 14 and 28 days

Exploratory outcomes;
Biological parameters improvement including CRP, neutrophil and lymphocytes counts

Criteria of safety

- Number of serious adverse events (SAEs)
- Number of Grade 3 and 4 AEs.

Investigational medication discontinuation (for any reason)

Statistical Method	Bayesian monitoring and analysis of the trial will be used. The primary outcome will be therefore analyzed using a Bayesian Cox model adjusted for age. The treatment effect will be summarized in terms of hazard ratio (HR) for the experimental vs. control arm. Every 60 patients randomized (30 in each arm) reach a minimal 7 days follow-up, an interim analysis is triggered, where several posterior probabilities will be calculated: 1) posterior probability of benefit $P_1 = P(HR < 1 \mid data)$; 2) posterior probability of at least a fair benefit $P_2 = P(HR < 0.8 \mid data)$, 3) posterior probability of inefficacy or harm $P_3 = P(HR > 1 \mid data)$. At the interim analysis, the trial can be stopped for futility if $P_2 < 0.10$ or $P_3 > 0.80$. At the first interim analysis, the trial can be stopped for efficacy if $P_1 > 0.99$, the. At the subsequent interim analyses, **efficacy boundaries are set to $P_1 > 0.95$ or $P_2 > 0.80$**. Decision boundaries are non-binding, and the DSMB can recommend continuing recruitment or stopping even if the aforementioned boundaries are not crossed. Overall, the trial is planned for a **total sample size of 660 participants** (330 per arm).

2 Major amendments to the protocol

An amendment has been submitted and approved on November 16, 2020, to modify the sample size (extending Bayesian interim analyses every other 60 patients randomized), and specifying survival without needs of invasive ventilation at day 14 as primary outcome.
In the original protocol, the maximum sample size was 180 patient, with a Bayesian interim analysis after 60 and 120 patients recruited, and the primary outcome was survival without needs of ventilator utilization (including non-invasive ventilation and Optiflow) at day 14.
Owing to the following reasons, the maximum sample size was increased up to 660, and the primary outcome changed to survival without needs of invasive ventilation at day 14:

1. Scientifically and medically, survival without mechanical ventilation better reflects the most severe worsening of patients' condition in the target population, and has been more often used in other trials compared to the original outcome.
2. Operationally, the experience of the CORIMUNO trials conducted until now has shown that the use of non-invasive ventilation and high-flow oxygen devices has been more difficult to record due to the large heterogeneity of the devices used.

Survival without needs of ventilator utilization (including non-invasive ventilation and Optiflow) was added to secondary outcomes.

The design was also modified as follows:
While we keep the core principles and stopping rules of the design, there are several issues to be considered, in the light of preceding CORIMUNO trials and other trials published or registered:

1. Targeting a very large effect as the current TOCIDEX trial (powered for HRs between 0.30 and 0.46) is likely too optimistic, and smaller effect sizes (e.g. HRs of 0.60–0.70) clearly represent important benefits for patients;
2. Small trials (total sample size < 200) could be justified in order to provide quick answers during the first wave of the pandemics, but are now not regarded as providing robust answers at the international level;
3. Switching the outcome to MV or death instead of any form of ventilator utilization or death implies a decrease in the expected event rate, thereby requiring a reassessment of the sample size.

Using the results of previous CORIMUNO trials (CORIMUNO-TOCI and CORIMUNO-ANA), we hypothesized that the proportion of patients with MV or death at day 14 would be 25%. A frequentist sample size calculation suggests that a total sample size of 634 patients is necessary to demonstrate a HR of 0.65 with power 80%, using a one-sided 5% type I error rate. Keeping the strategy of analysis every 60 patients of CORIMUNO trials, an indicative maximum sample size would be 660 (330 per arm).

The previous Bayesian design is then extended to analysis every 60 patients, triggered when data are available for at least 7 follow-up days for the first 60 patients.
At each analysis, the three previously planned posterior probabilities are computed:

- Posterior probability of benefit $P_1 = P(HR < 1 \mid data)$;
- Posterior probability of at least a fair benefit $P_2 = P(HR < 0.8 \mid data)$;
- Posterior probability of inefficacy or harm $P_3 = P(HR > 1 \mid data)$.

At each analysis, the following actions are triggered according to the thresholds given below, adapted from the Statistical Design and Analysis Plan for Sequential Parallel-Group RCF for COVID-19 (Harrell & Lindsell, 2020. http://hbiostat.org/proj/covid19/bayesplan.html):

- Stop with evidence for efficacy if $P_1 > 0.95$ ($P_1 > 0.99$ at the first analysis because we consider that there is a need for very convincing evidence if the sample size is very limited, and this was already planned in the previous version of the protocol);
- Stop for futility if $P_2 < 0.10$ or $P_3 > 0.80$;
- Stop with evidence for efficacy if $P_2 > 0.80$ (only actionable when at least 180 patients have been randomized, i.e. starting from the third analysis).

3 Trial termination

Recruitment to the CORIMUNO-TOCIDEX trial was stopped on May 18, 2021, following the recommendation of the DSMB, owing to slowed enrollment due to the decrease in the prevalence of COVID-19 and to evolving guidelines of care. Given the effect size at the interim analysis (May 10, 2021), it was considered that a statistically definitive result would not be expected without a sample size much larger than the planned 660. The DSMB did not identify any significant safety issues. At that date, 453 participants had been recruited to the trial.

4 Analysis population

4.1 Flow diagram

At the final analysis of trial, a flow chart will be constructed according to the CONSORT 2010 reporting guidelines. It will describe:

- The number of eligible patients, randomized patients and the number of patients who have actually followed the study;
- The intervention arm allocated per randomization;
- Early cessation of the intervention and their causes and drop-outs;
- The number of patients excluded from the analysis.

4.2 Definition of the analysis population

The final analysis will be carried out according to the intention to treat (ITT) principle, i.e. each randomised participant will be analysed in the group assigned to him/her by randomisation, regardless of the actual treatment received or other protocol deviations. In particular patients randomised while not meeting eligibility criteria will be kept in the analysis. For interim analyses, only patients with at least 7 days theoretical follow-up and data recorded for the first follow-up visit (7 days) will be analysed.
No data will be analysed for patients who have withdrawn their consent during the study and have expressed opposition to the analysis of their data. If necessary, the data concerning these patients that have been collected will be destroyed. The existence of these patients will nevertheless be documented in the study flow chart.

4.3 Sample size

The maximum sample size has been fixed for the whole trial at 660 (330 per arm), with a Bayesian analysis every 60 patients.
This sample size was based on considering that a frequentist sample size calculation would suggest that a total sample size of 634 patients is necessary to demonstrate a HR of 0.65 with power 80%, using a one-sided 5% type I error rate, and keeping the strategy of analysis every 60 patients of CORIMUNO trials.

5 Analysis principles

5.1 General principles for analysis of outcomes

The final results will be reported according to the recommendations of CONSORT 2010.

All outcomes will be analysed in superiority analyses, and the analyses will be adjusted for age and centre (randomisation stratification variable), the latter as a random effect.
The primary efficacy analyses will rely on computing the posterior distribution of the adjusted hazard ratio between the experimental and control arms for the primary outcome, using a proportional hazards model. These posterior distributions will be graphically displayed, and summarized by their medians and two-sided 90% credibility intervals (the Bayesian counterparts of confidence intervals).
For secondary efficacy and safety outcomes, frequentist (i.e. non-Bayesian) analyses will be used. No correction for multiplicity and no hierarchical testing procedures are planned in analysing secondary outcomes. These analyses will therefore be considered as exploratory in nature.

5.2 Participants' characteristics at inclusion

The characteristics of patients collected at inclusion will be described globally and by randomization group, using means, standard deviations, medians, interquartile intervals, minimum and maximum for quantitative variables and by their numbers and percentages by modality for qualitative variables.
The number of missing data for each variable will also be reported. No statistical tests for comparison between groups will be carried out.

5.3 Handling of missing or incoherent data

Given their nature and the trial settings, it is not be expected that primary outcome data would be missing. In the case of a follow-up shorter than 14 days, they will be naturally handled using methods for censored data. No imputation will be used for secondary efficacy and safety outcomes.

5.4 Statistical software

The analyses will be carried out using the R software version 4.0.1 or later (The R Foundation for Statistical Computing, Vienna, Austria), SAS version 9.4 or later (SAS Institute Cary, NC) and JAGS version 4.3.0 or later.

6 Primary outcome analysis

6.1 Definition of the primary outcome

The primary outcome is survival without needs of invasive ventilation at day 14.

6.2 Outcome analysis

6.2.1 Modelling

This section describes the Bayesian analysis of the primary outcome used for trial monitoring and final analysis.
A Bayesian Cox model will be estimated using Markov chain Monte Carlo (MCMC) methods, adjusted for age and centre (modelled as a random effect). The primary analysis will use a flat prior, and different sceptical or enthusiastic priors will be used as sensitivity analyses (see specification of the priors in the §5.2.2 below). In addition, a frequentist Cox model (adjusted for age and with random centre effect) will also be used.

6.2.2 Settings for Monte Carlo Markov Chain Bayesian analyses

The main analysis will use a Gaussian prior distribution with mean 0 and variance 10^2 for the log hazard ratio. The prior for the log hazard ratio for age will also be a Gaussian prior, with mean 0 and variance 10^2. Four different chains with different starting values will be run, with a burn-in of 10,000 iterations, and 100,000 additional iterations and a thinning interval of 10,

leading to keeping 10,000 values per chain, 40,000 in total. The convergence of the models will be assessed using the Gelman-Rubin statistic and by visual inspection of the trace of coefficients

As a sensitivity analysis, we will investigate different prior distributions, namely two sceptic priors centred on 0 with variance set so that a $P(HR < 0.2) = P(HR > 5) = 0.05$ (SD 0.975) or $P(HR < 0.2) = P(HR > 5) = 0.025$ (SD 0.82), and one enthusiastic informative prior centred on the targeted treatment effect (HR of 0.65) and informative with $\sigma = 0.975$.

6.2.3 Presentation of results

The posterior distribution of the hazard ratio will be displayed, and summarized by its median and two-sided 90% and 95% credibility intervals. Kaplan-Meier plots or cumulative incidence of the outcome will also be estimated in each arm, in a frequentist approach. Posterior probabilities of any benefit and at least a fair benefit will also be presented.

6.3 Stopping rules

At each interim analysis, the posterior distribution of the hazard ratio θ will be used to compute different posterior probabilities:

- Posterior probability of any benefit $P_1 = P(\theta < 1 \mid \text{data})$;
- Posterior probability of at least a fair benefit $P_2 = P(\theta < 0.8 \mid \text{data})$;
- Posterior probability of inefficacy or harm $P_3 = P(\theta > 1 \mid \text{data})$.

At each analysis, the following actions are triggered according to the thresholds given below, adapted from the Statistical Design and Analysis Plan for Sequential Parallel-Group RCF for COVID-19 (Harrell & Lindsell, 2020. http://hbiostat.org/proj/covid19/bayesplan.html):

- Stop with evidence for efficacy if $P_1 > 0.95$ ($P_1 > 0.99$ at the first analysis because we consider that there is a need for very convincing evidence if the sample size is very limited, and this was already planned in the previous version of the protocol);
- Stop for futility if $P_2 < 0.10$ or $P_3 > 0.80$;
- Stop with evidence for efficacy if $P_2 > 0.80$ (only actionable when at least 180 patients have been randomized, i.e. starting from the third analysis).

Any decision to stop or continue recruitment will be advised by the DSMB based of the aforementioned posterior probabilities, as well as safety data.

6.4 Calculation of the outcome

The day of randomisation will be counted as day 1. The time to mechanical ventilation or death will be computed starting from day one up to day 14 included.

A close data monitoring will be carried out to limit missing information on the use of ventilation as much as possible.

A new Do-Not-Resuscitate (DNR) order, i.e. a DNR order posterior to the date of randomisation and that has been noted as having been effectively used to limit care will be considered as an event for the primary outcome, at the date of limitation of care.

6.5 Subgroup analyses

The initial protocol specified that, at the end of the study, subgroup analyses would be performed according to antiviral therapies at baseline. Given the lack of efficacy of antivirals in the trials, no such analysis will be carried out.

A post-hoc subgroup analysis of the primary outcome is added according to CRP levels as randomisation (categorised as < 100 mg/L, ≥ 100 mg/L, 100 mg/L being close to the median value at randomisation). Interactions between the subgroup and treatment will be tested. Those analyses will use frequentist (i.e. non-Bayesian) models, adjusted for age and centre (as a random effect).

6.6 Conditional power analysis

Given the early trial termination without formally meeting the stopping rules, a conditional power analysis will be carried out using the approach developed in Andersen (Andersen PK. Conditional power calculations as an aid in the decision whether to continue a clinical trial. *Controlled Clinical Trials* 1987;8:67–74).
The conditional power will be computed as the probability of stopping for efficacy (using only $P_1 > 0.95$) after recruitment of an additional number of participants allowing to reach a total sample size between 660 (planned maximum sample size) and 2500. The conditional power will be computed for the initially planned hazard ratio (HR of 0.65), and for a less optimistic hazard ratio of 0.69, corresponding to the pooled odds ratio for progression to invasive mechanical ventilation, ECMO, or death at 28 days for the subgroup of participants receiving corticosteroids at baseline in the tocilizumab meta-analysis (WHO Rapid Evidence Appraisal for COVID-19 Therapies (REACT) Working Group, Shankar-Hari et al. Association between administration of IL-6 antagonists and mortality among patients hospitalized for COVID-19: A meta-analysis. JAMA. 2021 Jul 6:e2111330. doi: 10.1001/jama.2021.11330).

7 Secondary efficacy outcomes analysis

7.1 Definitions

- WHO progression scale at 7 and 14 days;
- Overall survival at 14, 28, 60 and 90 days;
- Time to discharge at 14 and 28 days;
- Time to oxygen supply independency at 14 and 28 days.

Biological parameters improvement including CRP, neutrophil and lymphocytes counts are not secondary outcomes, but exploratory outcomes.

7.2 Methods for analysis

7.2.1 Time-to-event outcomes

Time-to-event outcomes will be analysed using Cox or Fine-Gray regression models adjusted for the same variables as the primary outcome; results will be expressed as hazard ratios with 95% confidence interval. Competing risks analyses (Fine-Gray model) will be used for time to discharge, and time to oxygen supply independency, for which death will be considered as a competing event. When several timepoints are mentioned, separate models will be estimated at these timepoints. Point estimates of survival in each arm will be presented together with Kaplan-Meier survival curves.
For time to oxygen independency, when the precise date of oxygen independency is unknown, data will be censored at the last date when it was known that the patient was still under oxygen.

7.2.2 WHO ordinal scale

For the WHO ordinal scale, a proportional odds models will be used to compare the distribution of ordinal scores at day 7 and at day 14, adjusted for age and centre. The distribution of scores will be described at 7, and 14 days.
Given the possibility of missing data for the 10-point WHO Ordinal Scale for patients discharged alive, a 6-point WHO Ordinal Scale is added as a sensitivity analysis, which can be derived straightforwardly from the 10-point scale: 1: not in hospital; 2: hospitalized, no ventilation support; 3: hospitalized, with oxygen; 4: hospitalized, under NIV or HFO; 5: receiving MV: 6: death.

7.2.3 Biological and physiological outcomes

For biological outcomes, only descriptive analyses will be performed.

8 Safety analysis

8.1 Definitions

Adverse events are spontaneously declared on the CRF. For each adverse event, the following information is collected:

- Classification of the adverse event (AE) as a serious adverse event (SAE);
- Seriousness criteria for SAEs;
- Intensity (severity): mild, moderate or severe;
- Start/end dates;
- Investigator judgement on relationship with the study treatment, concomitant treatment, pre-existing disease and COVID-19;
- Modification of study treatment;
- Symptomatic treatment;
- Outcome.

Moreover, major safety endpoints are monitored: blood cells and platelets counts and liver transaminases, are monitored frequently, every three days systematically:

- Neutrophil count;
- Platelet count;
- Liver enzymes: ALT and AST;
- Occurrence of skin rashes;
- Systolic and diastolic blood pressure;
- Ventilator asynchronization.

8.2 Analysis

Adverse events and their characteristics will be described using numbers and percentages per treatment arm. The proportion of participants with each of the reported events, as well as the proportions of participants with at least one SAE will be compared using Fisher's exact tests. The total number of AE/SAEs and SAEs will also be described for each arm, and compared using Poisson models (with a robust error variance if necessary).

9 Summary of changes since previous versions

This Statistical Analysis Plan was developed from the Statistical Analysis Plan version 2.1 of previous CORIMUNO-19 trials.

9.1 Version 1.1 compared to version 1.0

- A version of the WHO ordinal scale on a 6-point scale was added for interim analyses, owing to a single score for patients discharged. Since scores for the 10-point scale were often missing for patients discharged, this was easier than introducing methods for left-censored data.
- A subgroup analysis of the primary outcome according to oxygen flow at randomisation (categorised as < 5 L, 5–9 L, ≥ 10 L) was planned for interim analyses, at the DSMB request.

9.2 Version 1.2 compared to version 1.1

- A paragraph (§3) on trial termination has been added. Subsequent paragraphs have been renumbered accordingly.
- The 6-point WHO scale was kept as a sensitivity analysis for the final analysis.

- A post-hoc subgroup analysis of the primary outcome was added according to CRP levels (categorised as < 100 mg/L, ≥ 100 mg/L).
- Added a precision on how missing data on the date of oxygen independency will be handled.
- Added a conditional power analysis.

Major amendments to the protocol

An amendment has been submitted and approved on November 16, 2020, to modify the sample size (extending Bayesian interim analyses every other 60 patients randomized), and specifying survival without needs of invasive ventilation at day 14 as primary outcome.
In the original protocol, the maximum sample size was 180 patient, with a Bayesian interim analysis after 60 and 120 patients recruited, and the primary outcome was survival without needs of ventilator utilization (including non-invasive ventilation and Optiflow) at day 14.
Owing to the following reasons, the maximum sample size was increased up to 660, and the primary outcome changed to survival without needs of invasive ventilation at day 14:

1. Scientifically and medically, survival without mechanical ventilation better reflects the most severe worsening of patients' condition in the target population, and has been more often used in other trials compared to the original outcome.
2. Operationally, the experience of the CORIMUNO trials conducted until now has shown that the use of non-invasive ventilation and high-flow oxygen devices has been more difficult to record due to the large heterogeneity of the devices used.

Survival without needs of ventilator utilization (including non-invasive ventilation and Optiflow) was added to secondary outcomes.

The design was also modified as follows:

While we keep the core principles and stopping rules of the design, there are several issues to be considered, in the light of preceding CORIMUNO trials and other trials published or registered:

1. Targeting a very large effect as the current TOCIDEX trial (powered for HRs between 0.30 and 0.46) is likely too optimistic, and smaller effect sizes (e.g. HRs of 0.60–0.70) clearly represent important benefits for patients;
2. Small trials (total sample size < 200) could be justified in order to provide quick answers during the first wave of the pandemics, but are now not regarded as providing robust answers at the international level;
3. Switching the outcome to MV or death instead of any form of ventilator utilization or death implies a decrease in the expected event rate, thereby requiring a reassessment of the sample size.

Using the results of previous CORIMUNO trials (CORIMUNO-TOCI and CORIMUNO-ANA), we hypothesized that the proportion of patients with MV or death at day 14 would be 25%. A frequentist sample size calculation suggests that a total sample size of 634 patients is necessary to demonstrate a HR of 0.65 with power 80%, using a one-sided 5% type I error rate. Keeping the strategy of analysis every 60 patients of CORIMUNO trials, an indicative maximum sample size would be 660 (330 per arm).

The previous Bayesian design is then extended to analysis every 60 patients, triggered when data are available for at least 7 follow-up days for the first 60 patients.
At each analysis, the three previously planned posterior probabilities are computed:

- Posterior probability of benefit $P_1 = P(HR < 1 \mid data)$;
- Posterior probability of at least a fair benefit $P_2 = P(HR < 0.8 \mid data)$;

- Posterior probability of inefficacy or harm $P_3 = P(HR > 1 \mid data)$.

  At each analysis, the following actions are triggered according to the thresholds given below, adapted from the Statistical Design and Analysis Plan for Sequential Parallel-Group RCF for COVID-19 (Harrell & Lindsell, 2020. http://hbiostat.org/proj/covid19/bayesplan.html):

- Stop with evidence for efficacy if $P_1 > 0.95$ ($P_1 > 0.99$ at the first analysis because we consider that there is a need for very convincing evidence if the sample size is very limited, and this was already planned in the previous version of the protocol);
- Stop for futility if $P_2 < 0.10$ or $P_3 > 0.80$;
- Stop with evidence for efficacy if $P_2 > 0.80$ (only actionable when at least 180 patients have been randomized, i.e. starting from the third analysis)