EPIDEMIOLOGICAL SCIENCE

Historically controlled comparison of glucocorticoids with or without tocilizumab versus supportive care only in patients with COVID-19-associated cytokine storm syndrome: results of the CHIC study

Sofia Ramiro ,[1,2] Rémy L M Mostard,[3] César Magro-Checa,[1] Christel M P van Dongen,[1] Tom Dormans,[4] Jacqueline Buijs,[5] Michiel Gronenschild,[3] Martijn D de Kruif,[3] Eric H J van Haren,[3] Tom van Kraaij,[3] Mathie P G Leers,[6] Ralph Peeters,[1] Dennis R Wong,[7] Robert B M Landewé [1,8]

Handling editor Josef S Smolen

► Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/annrheumdis-2020-218479).

For numbered affiliations see end of article.

Correspondence to
Dr Sofia Ramiro, Rheumatology, Zuyderland Medical Center, Heerlen, Limburg 6419, The Netherlands;
sofiaramiro@gmail.com

Received 1 July 2020
Revised 11 July 2020
Accepted 13 July 2020

To cite: Ramiro S, Mostard RLM, Magro-Checa C, *et al*. *Ann Rheum Dis* Epub ahead of print: [*please include* Day Month Year]. doi:10.1136/annrheumdis-2020-218479

ABSTRACT

Objectives To prospectively investigate in patients with severe COVID-19-associated cytokine storm syndrome (CSS) whether an intensive course of glucocorticoids with or without tocilizumab accelerates clinical improvement, reduces mortality and prevents invasive mechanical ventilation, in comparison with a historic control group of patients who received supportive care only.

Methods From 1 April 2020, patients with COVID-19-associated CSS, defined as rapid respiratory deterioration plus at least two out of three biomarkers with important elevations (C-reactive protein >100 mg/L; ferritin >900 µg/L; D-dimer >1500 µg/L), received high-dose intravenous methylprednisolone for 5 consecutive days (250 mg on day 1 followed by 80 mg on days 2–5). If the respiratory condition had not improved sufficiently (in 43%), the interleukin-6 receptor blocker tocilizumab (8 mg/kg body weight, single infusion) was added on or after day 2. Control patients with COVID-19-associated CSS (same definition) were retrospectively sampled from the pool of patients (n=350) admitted between 7 March and 31 March, and matched one to one to treated patients on sex and age. The primary outcome was ≥2 stages of improvement on a 7-item WHO-endorsed scale for trials in patients with severe influenza pneumonia, or discharge from the hospital. Secondary outcomes were hospital mortality and mechanical ventilation.

Results At baseline all patients with COVID-19 in the treatment group (n=86) and control group (n=86) had symptoms of CSS and faced acute respiratory failure. Treated patients had 79% higher likelihood on reaching the primary outcome (HR: 1.8; 95% CI 1.2 to 2.7) (7 days earlier), 65% less mortality (HR: 0.35; 95% CI 0.19 to 0.65) and 71% less invasive mechanical ventilation (HR: 0.29; 95% CI 0.14 to 0.65). Treatment effects remained constant in confounding and sensitivity analyses.

Conclusions A strategy involving a course of high-dose methylprednisolone, followed by tocilizumab if needed, may accelerate respiratory recovery, lower hospital mortality and reduce the likelihood of invasive mechanical ventilation in COVID-19-associated CSS.

Key messages

What is already known about this subject?

► COVID-19-associated cytokine storm syndrome (CSS) is an important complication of severe acute respiratory syndrome coronavirus 2 infection in up to 25% of the patients, often responsible for a fatal outcome.

What does this study add?

► A strategy involving a course of high-dose glucocorticoids, followed by tocilizumab if needed, has shown to accelerate respiratory recovery, lower hospital mortality and reduce the likelihood of invasive mechanical ventilation compared with supportive care only in COVID-19-associated CSS.

How might this impact on clinical practice or future developments?

► CSS should be recognised and considered as a treatable complication of COVID-19 and immunosuppressive treatment should be started timely.

► A treatment with high-dose glucocorticoids is a convenient choice since glucocorticoids are safe, widely available and inexpensive.

INTRODUCTION

Cytokine storm syndrome (CSS), a state of systemic hyperinflammation, is a rare and potentially lethal complication of various infections, malignancies and autoimmune diseases such as systemic juvenile idiopathic arthritis.[1 2] CSS is found in alarmingly high frequencies (10%–20%) in patients with COVID-19 pneumonia and may cause significant morbidity, including multiorgan failure, and mortality.[3 4] CSS can be suspected if patients experience rapid respiratory deterioration, in combination with high fever and disproportionally high C-reactive protein (CRP) and serum ferritin, among others.[4] While a curative therapy for COVID-19 is still lacking, intensive immunosuppressive treatment may ameliorate COVID-19-associated CSS and improve the outcome.[5] To date, information about immunosuppressive treatment of COVID-19-associated CSS is only anecdotal.[6–9]

The Zuyderland Medical Center (ZMC) is a large teaching hospital responsible for the care of 450.000 inhabitants in the South of the province of Limburg. South Limburg has an ageing population and a relatively poor health status[10 11] and the COVID-19 pandemic has hit the region hard. Between 7 March and 31 March 2020 more than 350 patients with severe COVID-19 were admitted for hospital care. Thirty per cent had symptoms of COVID-19-associated CSS, many of whom required intensive care treatment and over 40% died.

By the end of March 2020, ZMC physicians agreed to start an experimental treatment protocol for COVID-19-associated CSS and follow these patients meticulously in a prospective observational study. Rheumatologists were consulted because of their expertise in immunosuppressive treatment, as advised in recent European League Against Rheumatism recommendations.[12] The protocol responded to the broadly felt need among physicians to do more than 'only' providing supportive care to them. The clinicians refuted the seemingly obvious choice for a randomised controlled trial (RCT) with intensive immunosuppressive treatment versus supportive care alone in patients with CSS, with reference to the unacceptably high hospital mortality under supportive care only conditions, the reluctance to confront critically ill patients with an estimated 40% mortality and a one-to-one gamble of not receiving additional experimental treatment, and the overwhelming time pressure of the unprecedented pandemic.

Here we describe the results of 86 patients with COVID-associated CSS who have been treated according to the protocol (period 2) in comparison to 86 patients with COVID-associated CSS who had received supportive care before the protocol was in effect (period 1) (COVID High-intensity Immunosuppression in Cytokine storm syndrome (CHIC) study). Each patient in the treatment group was matched one to one to a control patient in order to create pseudorandomisation.

METHODS

Patients

In order to avoid exhaustion of the hospital care system, the ZMC had agreed upfront with local general practitioners and nursing home physicians to not refer (suspected) patients with COVID-19 to the hospital for diagnosis and supportive care if severe pre-existing clinical frailty was present, life expectancy was obviously limited or severe comorbidity in combination with COVID-19 was expected to have a very unfavourable outcome.

All patients admitted to ZMC for COVID-19 were registered in the ZuydErLand COVID-19 regiStry (ELVIS) from which demographic data, clinical signs and symptoms at presentation could be retrieved. All patients in the ELVIS received written information about the registry as well as an opt-out form in case they did not want to participate. None of the patients included in the CHIC study objected to participation.

Treatment group (period 2): Patients eligible to the CHIC treatment protocol had to have a diagnosis of COVID-19 and evidence for concomitant CSS. A diagnosis of COVID-19 involved the presence of clinical signs and symptoms suggestive of COVID-19 in combination with either a positive PCR test for severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) or a chest CT result of COVID-19 CT Classification (CO-RADS) 4 or 5.[13]

In order to meet the criteria for CSS in this study, patients had to have an oxygen saturation at rest ≤94% (ambient air) or tachypnoea (>30/min). In addition, they had to meet at least two out of the following three biomarker criteria: high CRP (>100 mg/L), high serum ferritin (>900 µg/L at one occasion, or a twofold increase of the level at admission within 48 hours) and high D-dimer level (>1500 µg/L). There were no pertinent exclusion criteria for the treatment protocol.

Every patient was fully informed about the off-label character of the treatment strategy and the potential side effects. Informed consent was obtained before the start of the treatment strategy. If the patient was incapable of executing permission, informed consent was given by the closest relative.

Control group (period 1): Patients potentially eligible to the retrospectively assembled control group had to be admitted between 7 March and 31 March 2020. Their data were collected retrospectively for the presence of COVID-19 pneumonia as described above. In addition, clinical data about daily respiratory status were retrieved, as well as CRP, serum ferritin and serum D-dimer levels. Missing laboratory tests were determined afterwards in stored serum samples, when available. The criterion for respiratory deterioration was checked as described above. CSS biomarker criteria were applied as described above. The eligibility of all control patients was independently checked in the patient's electronic file by two physicians (CD and CMC) and cases of disagreement were decided by consensus with a third physician (SR), without knowledge of the clinical course and outcome.

Outcomes

The primary outcome was discharge from the hospital or improvement of at least two stages (compared with baseline; whatever came first) on a WHO-endorsed 7-point ordinal scale, originally developed for trials with patients with influenza pneumonia, and used in several trials with patients with COVID-19.[14–16] The stages are: (1) non-hospitalised, able to resume normal activities; (2) non-hospitalised, but unable to resume normal activities; (3) hospitalised, not requiring oxygen therapy; (4) hospitalised, requiring additional oxygen therapy; (5) hospitalised, requiring high-flow nasal oxygen therapy, non-invasive mechanical ventilation or both; (6) hospitalised, requiring extracorporeal membrane oxygenation, mechanical ventilation or both; and (7) death. In our study we used the scale from 2 to 7, as we could not (yet) collect information on whether discharged patients are able to resume normal activities.

Key secondary outcomes were hospital mortality and the need to start invasive mechanical ventilation. Other secondary outcomes were improvement of one stage in the WHO score, WHO score at days 7 and 14, independence from oxygen therapy, duration of mechanical ventilation in the survivors and duration of hospitalisation in the survivors.

Baseline and time points of primary and secondary outcomes, if met, were determined prospectively for all treated patients and in retrospect by chart review for all control patients.

Treatment protocol

The treatment protocol included two steps: (1) immediate treatment with methylprednisolone (MP) 250 mg intravenously on day 1, followed by MP 80 mg intravenously on days 2–5, and an option for a 2-day extension if considered necessary and safe; (2) escalation of immunosuppressive treatment with a monoclonal antibody directed against the interleukin-6 receptor, tocilizumab (TCZ), between day 2 and day 5 (single-dose TCZ, 8 mg/kg body weight intravenous, max 800 mg). Criteria for escalation with TCZ were lack of clinical improvement or worsening in respiratory status (assessed on the WHO scale). Criteria for a 2-day extension of MP at day 5 were clear clinical improvement in

respiratory status (≥1 stage improvement on the WHO scale) but a partial decrease of biomarkers (CRP reduction less than 50%). Close multidisciplinary monitoring was an integral part of the strategy and was assured by daily meetings (RLMM, RBML, CMC, SR, CMPD, RP, MG, EHJvH, JB) in which all patients in the protocol were discussed and treatment was optimised. Discussions focused on immunosuppressive treatment decisions, treating secondary infections, thromboembolism and cardiac complications. Glucose levels were assessed twice daily during treatment with MP.

Cointerventions: All patients received ceftriaxone (2 g every 24 hours for 7 days) and up to 11 May 2020 in the presence of oxygen saturation <90% chloroquine 300 mg every 12 hours following a loading dose of 600 mg unless the corrected QT interval on an ECG was prolonged (>500 ms). Informed consent was obtained for this off-label therapy.

Complications

Complications during hospitalisation were closely monitored. Complications of special interest were well-known adverse events related to short-term high-dose MP and TCZ administration, and included bacterial or fungal infection, acute-onset congestive heart failure or aggravation of existing congestive heart failure, arrhythmia and gastrointestinal bleeding.

Matching procedure: After the first selection step, the two data files containing 92 patients from the treatment group and 106 patients from the control group were 1:1 matched on sex (M, F) and age (five age classes: <50, 50–59, 60–69, 70–79, ≥80 years) using the match command in Stata. The best matching result yielded two groups of 86 patients each (87% of potentially available patients could be matched).

Main analysis

Comparability at baseline of treatment group and control group was analysed descriptively for a wide range of variables and common univariable statistical tests for between-group differences were applied to test if the null hypothesis of no difference had to be rejected. Patients in the treatment group and control group were compared on a time-to-event basis, using proportional hazards regression analysis (Cox). Censoring of follow-up took place: (1) when the patient died; (2) when the patient was discharged; or (3) at the end of follow-up on 19 May 2020 (whatever came first). By convention, a patient who had died during the course of the study could not have improved ('zero improvement').

Because of the relatively small sample in relation to the relatively high number of variables at baseline, a prespecified analysis for effect mediation and confounding preceded the final selection of variables for multivariable adjustment. This analysis involved a two-step procedure, in which effect modification was excluded first by testing per baseline variable the interaction of that variable with treatment group, under adjustment for the main effects. Thereafter, confounding was checked per variable by investigating if the magnitude of the association between treatment group and outcome changed >10% by adding the variable to the model. It was decided upfront that—apart from treatment group, age and sex (default variables)—only variables with a clinically relevant interaction and a p value <0.1 were to be analysed in separate strata, and that only variables with true confounding potential that met the definition for confounding were to be included in the final multivariable models.

The proportional hazards assumption was checked by graphical diagnostics and statistical testing using Stata V.12. Kaplan-Meier survival plots were constructed and the survival curves for treated and control groups were compared using a log-rank test.

Sensitivity analysis

The effect size for treatment in the final multivariable models was challenged for robustness by several sensitivity analyses. The three sensitivity analyses were of the same type as the main analyses but on different patient selections: (1) all patients minus those who were already on mechanical ventilation at baseline; (2) all patients minus those who had received TCZ; or (3) all patients minus those in the lowest (<50 years) and highest (≥80 years) age groups (trimming).

Sample size considerations

At the start of the CHIC study there was only provisional information available about hospital mortality in patients with COVID-19-associated CSS under supportive care only conditions. We assumed 40% hospital mortality based on early experience. In order to declare an observed absolute difference of 20% or more statistically significant (50% mortality reduction, at alpha=0.05 and beta=0.20 (power: 80%)), at least 79 patients per group were required.

Role of the funding source

There was no funding source of this study. The corresponding author had full access to all the data in the study and had final responsibility for the decision to submit for publication.

RESULTS

After matching, 172 patients (86 per group) were available for analysis. All patients and controls were Caucasians. Mean age and sex distribution were similar. The age range was wide; younger and older patients were appropriately represented (online supplementary table S1). Table 1 shows a broad array of variables that describe the make-up of both groups. In general, the distribution of variables across groups was well balanced, but a few differences stood out. Mean body mass index (BMI) was high in both groups, but almost 2 units higher in the control group than in the treatment group. COVID-19 PCR positivity was slightly higher in the control group, which was the consequence of an imminent shortage in test capacity over time, in combination with increasing confidence in the diagnostic value of chest CT under high a priori probability conditions (the CO-RADS scores were balanced across groups). Differences in comorbidities at baseline were found in either direction. The control group contained more patients with diabetes (p<0.05). The treatment group had more patients with cardiovascular disease (p=0.2150) and arrhythmias (p=0.0290). The Charlson Comorbidity Index was low on average but slightly higher in the control group (p=0.1688).

The mean WHO score at baseline was similar in both groups, with the notion that the distribution across stages is skewed towards more patients requiring mechanical ventilation in the control group, offset by more patients requiring high-flow oxygen in the treatment group. Biomarkers of CSS showed very high mean levels for CRP, serum ferritin and D-dimer in both groups, but serum ferritin (p=0.0562) and D-dimer (p=0.3470) levels were slightly higher in the control group. Both ferritin and D-dimer had a non-normal distribution with outliers, and more than 50% of D-dimer levels in the control group were missing.

All patients in both groups (except for one in the treatment group) received antibiotic treatment by protocol, and almost

Table 1 Baseline demographic and clinical characteristics

	Treated group n=86	Control group n=86	P value*
Age (years)	67 (12)	67 (11)	0.8250
Male gender	68/86 (79%)	68/86 (79%)	
BMI (kg/m^2)	28.0 (4.9)	29.7 (5.6)	0.0418
Number of days sick	10.7 (5.5)	9.3 (5.7)	0.0948
Smoking			
Never smoker	35/83 (42%)	46/80 (58%)	0.1460
Ex-smoker	40/83 (48%)	28/80 (35%)	
Current smoker	8/83 (10%)	6/80 (8%)	
Positive COVID-19 PCR	76/85 (89%)	84/86 (98%)	0.0280
CO-RADS	4.7 (0.4)	4.8 (0.4)	0.7752
4	22/85 (26%)	15/63 (24%)	0.7730
5	63/85 (74%)	48/63 (76%)	
Hypertension	19/86 (22%)	27/86 (31%)	0.1680
Diabetes	9/86 (11%)	23/86 (27%)	0.0060
COPD	10/86 (12%)	7/86 (8%)	0.4430
Asthma	7/86 (8%)	5/86 (6%)	0.5490
Malignancy	4/86 (5%)	5/86 (6%)	0.7320
Cardiovascular disease	17/86 (20%)	11/86 (13%)	0.2150
Heart failure	2 (2%)	3/86 (3%)	0.6500
Arrhythmia	14/86 (16%)	5/86 (6%)	0.0290
Chronic kidney disease	4/86 (5%)	7/86 (8%)	0.3500
Cerebrovascular disease	4/86 (5%)	6/86 (7%)	0.5150
Peripheral vascular disease	5/86 (6%)	5/86 (6%)	1.0000
Autoimmune disease	9/86 (10%)	12/86 (14%)	0.4850
Peptic ulcer	1/86 (1%)	0/86 (0%)	0.3160
Moderate or severe liver disease	0/86 (0%)	1/86 (1%)	0.3160
Charlson Comorbidity Index (0–34)	0.90 (1.14)	1.16 (1.39)	0.1688
WHO score	4.3 (0.5)	4.4 (0.7)	0.1761
4: hospitalisation, requiring oxygen	65/86 (76%)	66/86 (77%)	<0.0001
5: hospitalisation requiring high-flow nasal oxygen therapy or non-invasive ventilation	20/86 (23%)	7/86 (8%)	
6: hospitalisation requiring ECMO, invasive mechanical ventilation or both	1/86 (1%)	13/86 (15%)	
Oxygen support at baseline			
Nasal oxygen	41/86 (48%)	41/86 (48%)	0.0010
OxyMask/NRM	24/86 (28%)	25/86 (29%)	
High-flow oxygen	20/86 (23%)	7/86 (8%)	
Mechanical ventilation	1/86 (1%)	13/86 (15%)	
Oxygen saturation	92 (10)	92 (6)	0.7432
CRP (mg/L)	160 (73)	167 (98)	0.5497
Ferritin (µg/L)	1493 (926)	1849 (1281)	0.0562
D-dimers (µg/L)	3935 (7350)	5633 (9408)	0.3470

Data are mean (SD) or n/N (%).
*Calculated with t-test or χ^2, as appropriate.
BMI, body mass index; COPD, chronic obstructive pulmonary disease; CO-RADS, COVID-19 CT Classification; CRP, C-reactive protein; ECMO, extracorporeal membrane oxygenation; NRM, non-rebreathing mask.

80% received chloroquine (table 2). None of the control patients received glucocorticoids (GC) or TCZ in any dose. Seven patients in the control group in the intensive care unit received antiviral treatment (lopinavir/ritonavir) before 20 March. From this date onward, antiviral treatment was removed from the protocol

Table 2 Treatments received during hospitalisation

	Treated group n=86	Control group n=86	P value*
Methylprednisolone	86/86 (100%)	0 (0%)	
Duration of methylprednisolone (days)	5.5 (1.6)	NA	
Second treatment with methylprednisolone	8/86 (9%)	0 (0%)	
Tocilizumab	37/86 (43%)	0 (0%)	
Days between start of methylprednisolone and start of tocilizumab	4.0 (4.4)	NA	
Anticoagulation			0.0160
None	0/86 (0%)	8/86 (9%)	
LMWH	78/86 (91%)	75/86 (87%)	
DOAC/NOAC	6/86 (7%)	2/86 (2%)	
Acenocoumarol	2/86 (2%)	1/86 (1%)	
Therapeutic anticoagulation	13/86 (15%)	7/86 (8%)	0.1540
Chloroquine	66/86 (77%)	68/86 (79%)	0.7130
Antibiotics	85/86 (99%)	86/86 (100%)	0.3160

Data are mean (SD) or n/N (%).
*Calculated with t-test or χ^2, as appropriate.
DOAC, direct oral anticoagulant; LMWH, low-molecular-weight heparin; NA, not applicable; NOAC, non-vitamin K antagonist oral anticoagulant.

in our intensive care unit due to the publication showing no benefit of lopinavir-ritonavir treatment beyond standard care in COVID-19.[14] Treated patients received MP slightly longer than the initially prescribed 5 days, due to treatment extensions in 25 patients (29%) and a second course of MP in eight patients (9%). Of all 86 treated patients, 37 received TCZ (43%) after a mean lag of 4 days, because of insufficient clinical improvement with MP. Two patients (2%) received a second dose of TCZ, both 5 days after the first administration.

Table 3 shows the primary and secondary outcomes. As compared with patients in the control group, patients in the treatment group had a 79% higher likelihood of reaching the primary outcome of two-stage improvement in respiratory status (HR: 1.79; 95% CI 1.20 to 2.67, table 3) and they reached it on average 7 days (median) earlier, all of them before discharge. Their WHO scores at days 7 and 14 (table 4) were consistently better (p<0.0001). The development of WHO clinical improvement per group is visualised in figure 1. Curves start to separate from 5 to 7 days of follow-up.

Hospital mortality was 65% lower in the treatment group than in the control group (HR: 0.35; 95% CI 0.19 to 0.65; table 3, figure 1). At hospital day 14, ten patients in the treatment group had deceased as compared with 33 in the control group (p<0.0001) (table 4).

The likelihood to evolve to mechanical ventilation due to respiratory deterioration was 71% lower in the treatment group (HR: 0.29; 95% CI 0.14 to 0.60; table 3, online supplementary figure S1). Among patients who were not mechanically ventilated at baseline (table 4), the daily incidence of mechanical ventilation (new start) was 1.3% vs 5.4% (p=0.0003). Once mechanically ventilated, the duration of mechanical ventilation was not different across groups.

The analyses of effect modification and confounding revealed one clinically relevant interaction (treatment group vs serum ferritin level at baseline) and six relevant confounders (BMI, smoking status, hypertension, diabetes, cardiovascular disease and arrhythmia). The effect size of treatment was higher in patients with serum ferritin levels above the median

Table 3 Effect of treatment (vs control) on outcomes, univariable and multivariable analyses

	Effect of treatment versus control Univariable analysis HR or coefficient (95% CI)	Effect of treatment versus control Multivariable analysis* HR or coefficient (95% CI)
Primary outcome		
Clinical improvement (2 points) in WHO score†	1.79 (1.20 to 2.67)	2.31 (1.45 to 3.68)
Key secondary outcomes		
Hospital mortality†	0.35 (0.19 to 0.65)	0.26 (0.13 to 0.52)
Mechanical ventilation†	0.29 (0.14 to 0.60)	0.22 (0.10 to 0.52)
Other key secondary outcomes		
Clinical improvement (1 point) in WHO score†	1.95 (1.33 to 2.87)	2.26 (1.44 to 3.54)
Independence from oxygen therapy†	1.80 (1.19 to 2.71)	2.36 (1.45 to 3.83)
Duration of mechanical ventilation in survivors‡	−2.57 (−12.08 to 6.93)	−6.83 (−21.45 to 7.79)
Duration of hospitalisation in survivors‡	−5.23 (−8.99 to −1.46)	−6.65 (−10.93 to −2.37)

*Adjusted for age, gender, body mass index, smoking status, hypertension, diabetes, cardiovascular disease and arrhythmia.
†Results from Cox regression models.
‡Results from linear regression models.
.

(>1419 µg/L) than in patients with serum ferritin levels below the median (HR for WHO 2-point clinical improvement: 2.7 vs 1.6) (online supplementary table S2).

The six relevant confounders were entered as covariates, together with age and sex, in the multivariable models for all outcomes (table 3). Adjustment for confounding increased rather than decreased the estimated treatment effect for all seven analysed outcomes. All models were checked for not violating the proportional hazards assumption (online supplementary figure S2).

All main effects remained constant and statistically significant in the three sensitivity analyses (table 5). Of note, in the sensitivity analysis that excluded the 43% patients who had received TCZ, the treatment effects for all outcomes increased and maintained statistical significance, suggesting that a clinically relevant treatment effect can be reached by high-dose GC alone (table 5, online supplementary table S3).

Patients tolerated the short but intensive immunosuppressive therapy well. Complications were balanced between groups (table 6). Bacterial infections were diagnosed during hospitalisation in 15 patients (8 in the treatment group vs 7 in the control group). There was a trend towards more pulmonary embolism in the treatment group (p=0.0590). Arrhythmias occurred in both groups, but slightly less frequently in the treatment group (p=0.265).

DISCUSSION

This historically controlled comparison of a strategy with intensive immunosuppression and close monitoring versus a strategy with supportive care only in patients with COVID-19-associated CSS suggests that clinically relevant improvement of respiratory status is 79% more likely, and can be accelerated by a median of 7 days, that hospital mortality can be reduced by 65% and that the need for mechanical ventilation during admission can be reduced by 71%. These outcomes were robust and, especially if confirmed in randomised trials later on, highly relevant from a medical and societal perspective.

It should be emphasised that this study cannot be read and interpreted as an RCT. Prognostic similarity at baseline cannot be assumed, in spite of several efforts to match the control patients as closely as possible to the treated patients. Residual confounding by unmeasured variables is likely. While patients were almost perfectly matched for age and gender and efforts were made to assure that control patients were only sampled if they had evidence of CSS, certain baseline differences remained, although in both directions. Diabetes and obesity were slightly more prevalent in the control patients and some biomarkers of CSS were slightly higher too. Cardiovascular comorbidity and arrhythmias, on the other hand, were more prevalent in the treated patients. A rigorous confounding analysis revealed that none of these potential confounders had a reducing influence on the magnitude of the treatment effect, which importantly adds to the credibility of the univariable results. Instead, HRs seemed to increase rather than decrease after adjusting for confounders, an observation that may point to statistical overfitting and therefore of limited relevance.

However, a potential (time) period effect may have affected the results.[17] Such an effect is inherent to the design of the study and cannot be adjusted for. All control patients were admitted at least 3 weeks earlier than the patients in the treatment group. These patients got sick in period 1, the initial phase of the pandemic, and it cannot be precluded that COVID-19-associated CSS was more severe in period 1 than in period 2, that patients received less than optimal supportive care during the first hectic weeks of the pandemic or that infected patients of the first hour simply had worst health. This argument, however, can also be reversed. During period 1 intensive care capacity still was relatively high and there was some consensus among experts to start low-threshold mechanical ventilation, while patients in period 2 had to 'compete with' those that already occupied high-care facilities and staff got exhausted. That more patients in the control group than in the treatment group already received mechanical ventilatory support at baseline is reflective of this situation. Adjustment for this imbalance did not affect the treatment effect. The argument that COVID-19-associated CSS was more severe in period 1 as compared with in period 2 also lacks substantiation. The start of our protocol coincided with the peak in COVID-19 admissions in the Netherlands (2 April) and preceded the peak in mortality by 5–7 days.[18] The incidence density of admissions and mortality increased during period 1 and decreased during period 2, but mortality as a fraction of number of hospital admissions in the Netherlands was actually higher in period 2 than in period 1. While this is an indirect argument, it argues against a better prognosis of patients with COVID-19-associated CSS in period 2. Still, we are dealing with a new disease and the standard of care is rapidly evolving. Changing policies with respect to the start of mechanical ventilation, diagnosing thrombosis and anticoagulation therapy occurred even within the time frame of our study. The trend of finding more pulmonary embolism in the treatment group, for instance, is a reflection of searching with more scrutiny for thrombosis over time. Such developments have an impact on the external validity (generalisability) of our study and of others to be published, and may be responsible for treatment effects in daily practice that seem less dramatic than the contrasts found in this early study.

Several experts, including the WHO,[19] warned against treating critically ill patients with a SARS-CoV-2 infection with GCs, an advice with potentially serious implications for many patients. The risk profile of such a short course of GC for treatment of CSS needs to be separated from pre-existing chronic use of GC for conditions like rheumatic and musculoskeletal diseases

Table 4 Outcomes

	Treated group n=86	Control group n=86	P value*
Primary outcome			
Clinical improvement (2 points) in WHO score	64	44	0.0025
Daily incidence rate (95% CI)	6.9% (5.4% to 8.8%)	4.0% (3.0% to 5.4%)	
25% improvement time (days)	6	8	
50% improvement time (days)	11	18	
Key secondary outcomes			
Hospital mortality	14	41	0.0004
Daily incidence rate (95% CI)	1.4% (0.9% to 2.4%)	3.5% (2.6% to 4.7%)	
25% survival time (days)	16	5	
50% survival time (days)	–	18	
Mechanical ventilation	10	24	0.0003
Daily incidence rate (95% CI)	1.3% (0.7% to 2.5%)	5.4% (3.6% to 8.0%)	
25% survival time (days)	–	3	
50% survival time (days)	–	-	
Other key secondary outcomes			
Clinical improvement (1 point) in WHO score	69	45	0.0003
Daily incidence rate (95% CI)	8.6% (6.8% to 10.9%)	4.6% (3.4% to 6.2%)	
25% improvement time (days)	5	7	
50% improvement time (days)	9	16	
Clinical improvement (1 point) in WHO score if no death takes later place	65	44	0.0009
Daily incidence rate (95% CI)	8.0% (6.3% to 10.2%)	4.5% (3.3% to 6.0%)	
25% improvement time (days)	6	7	
50% improvement time (days)	9	16	
WHO score at day 7			
2: no hospitalisation	18/84 (21%)	9/86 (11%)	<0.0001
3: hospitalisation, not requiring oxygen	6/84 (7%)	6/86 (7%)	
4: hospitalisation, requiring oxygen	29/84 (35%)	15/86 (17%)	
5: hospitalisation requiring high-flow nasal oxygen therapy or non-invasive ventilation	17/84 (20%)	2/86 (2%)	
6: hospitalisation requiring ECMO, invasive mechanical ventilation or both	7/84 (8%)	30/86 (35%)	
7: death	7/84 (8%)	24/86 (28%)	
WHO score at day 14			
2: no hospitalisation	48/83 (58%)	21/86 (24%)	<0.0001
3: hospitalisation, not requiring oxygen	6/83 (7%)	4/86 (5%)	
4: hospitalisation, requiring oxygen	10/83 (12%)	7/86 (8%)	
5: hospitalisation requiring high-flow nasal oxygen therapy or non-invasive ventilation	4/83 (5%)	2/86 (2%)	
6: hospitalisation requiring ECMO, invasive mechanical ventilation or both	5/83 (6%)	19/86 (22%)	
7: death	10/83 (12%)	33/86 (38%)	
Independence from oxygen therapy	61	40	0.0038
Daily incidence rate (95% CI)	5.9% (4.6% to 7.5%)	3.3% (2.4% to 4.4%)	
25% improvement time (days)	6	10	
50% improvement time (days)	11	19	
Duration of mechanical ventilation in survivors (days)	16.3 (11.9)	18.8 (10.2)	0.5809
Duration of hospitalisation in survivors and discharged (days)	10.8 (6.8)	15.9 (12.7)	0.0196

Data are number of cases, n/N (%), incidence rate with 95% CI, number of days or mean (SD), as indicated.
*Calculated with t-test, χ^2 or log-rank test, as appropriate.
ECMO, extracorporeal membrane oxygenation.

(RMD). They particularly feared impaired virus clearance and secondary bacterial infection. However, patients in the CHIC study tolerated the immunosuppressive therapy remarkably well and we did not find evidence for impaired viral clearance nor for bacterial superinfection. Longer follow-up, however, is needed to give final resolution about the safety and efficacy of the strategy.

Speculating about which component of the strategy yields most benefit is tempting but risky. We think it is the combination of early intervention (the 'window of opportunity hypothesis'), the intensive immunosuppression and the close monitoring by a multidisciplinary team that best explains the favourable results. The results of the CHIC study also suggest that the timely administration of high-dose GCs alone may provide significant benefit in more than half of the patients and that TCZ is only needed in those cases that had insufficient clinical improvement on MP alone. This is an important finding given the limited availability of TCZ in many countries and TCZ's high costs.

It is not unthinkable that treatment with other compounds than GC and TCZ that are often used by patients with RMDs,

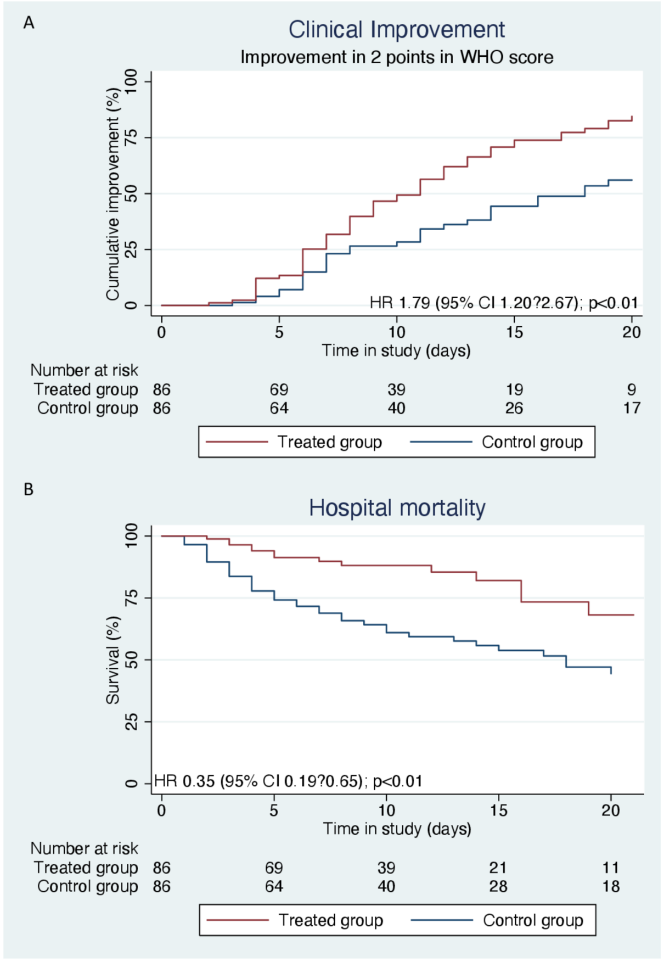

Figure 1 Clinical improvement and hospital mortality. Plots show clinical improvement (A) defined as a 2-point improvement in the 7-point WHO score and (B) hospital mortality in patients with COVID-19-associated cytokine release syndrome stratified for treatment (treated vs control group).

Table 5 Sensitivity analyses

	Effect of treatment versus control Univariable analysis HR or coefficient (95% CI)	Effect of treatment versus control Multivariable analysis* HR or coefficient (95% CI)
Primary outcome		
Clinical improvement (2 points) in WHO score†		
Excluding ventilated patients at baseline	1.63 (1.08 to 2.47)	2.03 (1.25 to 3.31)
Excluding patients receiving TCZ	2.53 (1.62 to 3.96)	3.33 (1.94 to 5.73)
Excluding extreme age groups‡	1.98 (1.28 to 3.08)	2.44 (1.45 to 4.13)
Key secondary outcomes		
Hospital mortality†		
Excluding ventilated patients at baseline	0.34 (0.18 to 0.64)	0.26 (0.13 to 0.54)
Excluding patients receiving TCZ	0.36 (0.16 to 0.81)	0.24 (0.10 to 0.62)
Excluding extreme age groups‡	0.31 (0.15 to 0.66)	0.33 (0.14 to 0.75)
Mechanical ventilation†		
Excluding patients receiving TCZ	0.11 (0.03 to 0.45)	0.09 (0.02 to 0.41)
Excluding extreme age groups	0.31 (0.15 to 0.66)	0.25 (0.10 to 0.59)
Other secondary outcomes		
Clinical improvement (1 point) in WHO score†		
Excluding ventilated patients at baseline	1.78 (1.19 to 2.67)	1.96 (1.22 to 3.12)
Excluding patients receiving TCZ	2.95 (1.89 to 4.60)	3.34 (1.97 to 5.66)
Excluding extreme age groups‡	2.04 (1.33 to 3.13)	2.44 (1.47 to 4.06)
Independence from oxygen therapy†		
Excluding ventilated patients at baseline	1.51 (0.99 to 2.30)	1.88 (1.15 to 3.08)
Excluding patients receiving TCZ	2.39 (1.52 to 3.78)	3.68 (2.09 to 6.48)
Excluding extreme age groups‡	2.04 (1.29 to 3.22)	2.71 (1.56 to 4.67)
Duration of mechanical ventilation in survivors§		
Excluding ventilated patients at baseline	−3.71 (−14.70 to 7.28)	−6.20 (−37.81 to 25.41)
Excluding patients receiving TCZ	3.18 (−12.98 to 19.34)	−1.72 (−24.92 to 21.49)
Excluding extreme age groups‡	−4.25 (−14.34 to 5.84)	−9.14 (−25.92 to 7.64)
Duration of hospitalisation in survivors§		
Excluding ventilated patients at baseline	−5.56 (−9.29 to −1.83)	−6.73 (−10.86 to −2.59)
Excluding patients receiving TCZ	−6.59 (−11.25 to −1.93)	−8.25 (−13.72 to −2.78)
Excluding extreme age groups‡	−5.80 (−10.23 to −1.37)	−7.42 (−12.60 to −2.24)

*Adjusted for age, sex, body mass index, smoking status, hypertension, diabetes, cardiovascular disease and arrhythmia.
†Results from Cox regression models, HR and 95% CI.
‡Excluding patients from the age category <50 and ≥80 years old.
§Results from linear regression models, regression coefficient and 95% CI.
TCZ, tocilizumab.

such as tumour necrosis factor alpha inhibitors, or interleukin-1 receptor antagonists, may have similar beneficial effects.[7 8 20] Our choice for high-dose MP, broadly available and with a well-known profile in severe systemic inflammatory diseases, and TCZ, with an existing niche indication for a rare form of iatrogenic CSS,[21] was to some extent serendipitous. Future RCTs may give resolution but will take time.

Our definition of CSS was rather pragmatic and arbitrary but sufficed to select the ~200 patients with COVID-19 and CSS from a total pool of approximately 800 (25%). When untreated, mortality in this severe subgroup was more than 40%. In addition to the clinical criterion of rapid deterioration of respiratory status, we used commonly available biomarkers for CSS at cut-off levels based on our experience with the early patients with COVID-19. That the treatment benefit was highest in patients with serum ferritin levels above the median value of 1419 µg/L can be seen as an important endorsement for the hypothesis that COVID-19-associated CSS determines the immediate prognosis of COVID-19. Fine-tuning these thresholds for CSS and introducing other biomarkers for diagnosing CSS may further improve the definition of CSS and optimise treatment effects.

The choice for the WHO-endorsed classification of outcomes designed for patients with severe influenza pneumonia, next to hospital mortality, appeared rational and workable. The definition of an improvement of at least two stages or hospital discharge indeed reflected clinical improvement properly. In fact, all discharged patients also met the criterion of two-stage improvement.

While we in principle advocate to further study strategies like ours in RCTs, it is an interesting philosophical question whether RCTs with 'supportive care only' in the control group will be justifiable in future. 'Supportive care only' for COVID-19-associated CSS is ethically arguable in light of the biological plausibility of CSS complicating COVID-19 and its high mortality, given the likely benefit of immunosuppressive therapy, such as reported here, even though formal evidence stemming

Table 6 Complications during the hospitalisation

	Treated group n=86	Control group n=86	P value*
Bacterial infection	8/86 (9%)	7/86 (8%)	0.7870
Pneumothorax	2/86 (2%)	0/86 (0%)	0.1550
Cerebrovascular accident	1/86 (1%)	3/86 (3%)	0.3120
Congestive heart failure	1/86 (1%)	1/86 (1%)	1.0000
Endocarditis/myocarditis	0/86 (0%)	0/86 (0%)	1.0000
Pericarditis	0/86 (0%)	1/86 (1%)	0.3160
Cardiac ischaemia	0/86 (0%)	0/86 (0%)	1.0000
New arrhythmia	5/86 (6%)	9/86 (10%)	0.265
Deep venous thrombosis	2/86 (2%)	2/86 (2%)	1.0000
Pulmonary embolism	18/86 (21%)	9/86 (11%)	0.0590
Renal failure	7/86 (8%)	12/86 (14%)	0.2240
Gastrointestinal bleeding	0/86 (0%)	0/86 (0%)	1.0000

Data are presented as n/N (%).
*Calculated with χ^2.

from RCTs is lacking. While the magnitude of the treatment effects found in the CHIC study may be somewhat downplayed by formal methodological reasoning, it is unlikely that the entire contrast is only due to residual confounding. The evaluation of the treatment of COVID-19-associated CSS may have been caught up by the crisis itself.

In conclusion, we have shown here that a strategy involving a course of high-dose MP, followed by TCZ in case of insufficient improvement, may accelerate respiratory recovery, lower hospital mortality and reduce the likelihood of invasive mechanical ventilation in COVID-19-associated CSS. Despite these promising results, further confirmation is still needed.

Author affiliations
[1]Rheumatology, Zuyderland Medical Center, Heerlen, Limburg, The Netherlands
[2]Rheumatology, Leiden University Medical Center, Leiden, Zuid-Holland, The Netherlands
[3]Department of Pulmonology, Zuyderland Medical Center, Heerlen, Limburg, The Netherlands
[4]Department of Intensive Care, Zuyderland Medical Center, Heerlen, Limburg, The Netherlands
[5]Department of Internal Medicine, Zuyderland Medical Center, Heerlen, Limburg, The Netherlands
[6]Department of Clinical Chemistry and Hematology, Zuyderland Medical Center, Heerlen, Limburg, The Netherlands
[7]Department of Clinical Pharmacy, Pharmacology and Toxicology, Zuyderland Medical Center, Heerlen, Limburg, The Netherlands
[8]Amsterdam Rheumatology Center, AMC, Amsterdam, The Netherlands

Acknowledgements The authors extensively thank the healthcare professionals from Zuyderland Medical Center who made this study possible, despite the pandemic situation, as well as the Board of ZMC for the trust in the study team and the support in the conduction of the study. Additionally, the authors would like to thank the medical students (Susan Voncken, Marlieke Elsendoorn, Jasper Broerse, Twan Feron, Mandy Jongbloed and Sophie Laven) involved in data entry and employees from the Research Department (Bureau Wetenschappelijk Onderzoek, BWO), namely Anke Linssen, Esther Bergman, Christel Jacquot, Marijke Lemmens and Audrey Merry. The authors also acknowledge Professor D van der Heijde for her critical reading of the manuscript.

Contributors SR, RLMM, CMC and RBML designed the study. All authors contributed to data collection. SR and RBML analysed the data. SR, RLMM and RBML critically interpreted the results and drafted the first version of the manuscript. All coauthors discussed the findings together, critically reviewed the manuscript and approved its final version.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests SR reports personal fees from AbbVie, personal fees from Eli Lilly, grants and personal fees from MSD, personal fees from Novartis, personal fees from UCB, personal fees from Sanofi, outside the submitted work. RLMM reports personal fees from Boehringer Ingelheim, personal fees from Roche, personal fees from Galapagos, outside the submitted work. CMC is a clinical trial investigator for a study sponsored by Lilly and was a subinvestigator for a study sponsored by GSK. CvD reports personal fees from Novartis, personal fees from Roche, outside the submitted work. TD reports grants from Adrenomed, grants from Inotrem, grants from Roche, grants from Shionogi and Co, other from CASTOR, outside the submitted work. MG reports personal fees from Roche, personal fees from MSD, outside the submitted work. MdK reports personal fees from ALK, personal fees from AstraZeneca, personal fees from Boehringer Ingelheim, personal fees from Sanofi Genzyme, outside the submitted work. ML reports grants from AstraZeneca, grants from Pfizer, personal fees from Roche, outside the submitted work. RP reports grants and personal fees from Pfizer, grants and personal fees from AbbVie, outside the submitted work. RL reports personal fees from AbbVie, personal fees from BMS, personal fees from Galapagos, personal fees from Gilead, personal fees from Jansen, personal fees from Novartis, personal fees from Pfizer, personal fees from Roche, personal fees from UCB, outside the submitted work; and owner and director of Rheumatology Consultancy, a company that provides consultancy and read services for clinical trials.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Obtained.

Ethics approval The Medical Ethics Committee and the Board of ZMC approved the study protocol and the study started enrolling patients on 1 April 2020.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data may be obtained from a third party and are not publicly available.

ORCID iDs
Sofia Ramiro http://orcid.org/0000-0002-8899-9087
Robert B M Landewé http://orcid.org/0000-0002-0577-6620

REFERENCES

1 Janka GE. Familial and acquired hemophagocytic lymphohistiocytosis. *Annu Rev Med* 2012;63:233–46.
2 Henderson LA, Canna SW, Schulert GS, *et al*. On the alert for cytokine storm: immunopathology in COVID-19. *Arthritis Rheumatol* 2020;72:1059–63.
3 Moore JB, June CH. Cytokine release syndrome in severe COVID-19. *Science* 2020;368:473–4.
4 Zhou F, Yu T, Du R, *et al*. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. *Lancet* 2020;395:1054–62.
5 Mehta P, McAuley DF, Brown M, *et al*. COVID-19: consider cytokine storm syndromes and immunosuppression. *Lancet* 2020;395:1033–4.
6 Wu C, Chen X, Cai Y, *et al*. Risk factors associated with acute respiratory distress syndrome and death in patients with coronavirus disease 2019 pneumonia in Wuhan, China. *JAMA Intern Med* 2020. doi:10.1001/jamainternmed.2020.0994. [Epub ahead of print: 13 Mar 2020].
7 Cavalli G, De Luca G, Campochiaro C, *et al*. Interleukin-1 blockade with high-dose anakinra in patients with COVID-19, acute respiratory distress syndrome, and hyperinflammation: a retrospective cohort study. *Lancet Rheumatol* 2020;2:e325–31.
8 Huet T, Beaussier H, Voisin O, *et al*. Anakinra for severe forms of COVID-19: a cohort study. *Lancet Rheumatol* 2020;2:e393–400.
9 Li H, Chen C, Hu F, *et al*. Impact of corticosteroid therapy on outcomes of persons with SARS-CoV-2, SARS-CoV, or MERS-CoV infection: a systematic review and meta-analysis. *Leukemia* 2020;34:1503–11.
10 Bosma H, Jansen M, Groffen D, *et al*. Een nadere analyse van de gezondheidsachterstand in stadsregio Parkstad. *Tijdschrift voor gezondheidswetenschappen* 2013;91:318–23.
11 Health atlas 2020 district health services Zuid-Limburg. Available: https://www.gezondheidsatlaszl.nl/ [Accessed 4 Jun 2020].
12 Landewé RB, Machado PM, Kroon F, *et al*. EULAR provisional recommendations for the management of rheumatic and musculoskeletal diseases in the context of SARS-CoV-2. *Ann Rheum Dis* 2020;79:851–8.
13 Prokop M, van Everdingen W, van Rees Vellinga T, *et al*. CO-RADS - A categorical CT assessment scheme for patients with suspected COVID-19: definition and evaluation. *Radiology* 2020;201473:201473.
14 Cao B, Wang Y, Wen D, *et al*. A trial of Lopinavir-Ritonavir in adults hospitalized with severe Covid-19. *N Engl J Med* 2020;382:1787–99.
15 Wang Y, Zhang D, Du G, *et al*. Remdesivir in adults with severe COVID-19: a randomised, double-blind, placebo-controlled, multicentre trial. *Lancet* 2020;395:1569–78.

16 WHO R&D Blueprint. Novel coronavirus, outline of designs for experimental therapeutics, 2020. Available: https://apps.who.int/iris/bitstream/handle/10665/330695/WHO-HEO-RDBlueprintnCoV-2020.5-eng.pdf?sequence=1&isAllowed=y [Accessed 4 Jun 2020].
17 Holford TR. Understanding the effects of age, period, and cohort on incidence and mortality rates. *Annu Rev Public Health* 1991;12:425–57.
18 . Available: https://www.rivm.nl/coronavirus-covid-19/grafieken [Accessed 4 Jun 2020].
19 World Health Organization. Clinical management of severe acute respiratory infection (SARI) when COVID-19 disease is suspected, 2020. Available: https://apps.who.int/iris/handle/10665/331446?search-result=true&query=10665%2F331446&scope=&rpp=10&sort_by=score&order=desc [Accessed 30 Jun 2020].
20 Feldmann M, Maini RN, Woody JN, *et al*. Trials of anti-tumour necrosis factor therapy for COVID-19 are urgently needed. *Lancet* 2020;395:1407–9.
21 . Available: https://www.ema.europa.eu/en/documents/smop/chmp-post-authorisation-summary-positive-opinion-roactemra_en-3.pdf [Accessed 4 Jun 2020].