The Journal of Infectious Diseases

MAJOR ARTICLE

The Effect of IL-6 Inhibitors on Mortality Among Hospitalized COVID-19 Patients: A Multicenter Study

Pranay Sinha,[1,a] S. Reza Jafarzadeh,[2,a] Sabrina A. Assoumou,[1] Catherine G. Bielick,[3] Bethanne Carpenter,[4] Shivani Garg,[5] Sahni Harleen,[6] Tuhina Neogi,[2] Midori Jane Nishio,[7] Manish Sagar,[1] Veronika Sharp,[8] and Eugene Y. Kissin[2]

[1]Section of Infectious Diseases, Department of Medicine, Boston University School of Medicine, Boston, Massachusetts, USA, [2]Section of Rheumatology, Department of Medicine, Boston University School of Medicine, Boston, Massachusetts, USA, [3]Department of Internal Medicine, University School of Medicine, Boston, Massachusetts, USA, [4]Department of Pharmacy, Santa Clara Valley Medical Center, San Jose, California, USA, [5]Department of Medicine, Rheumatology Division, University of Wisconsin, Madison, Wisconsin, USA, [6]Division of Infectious Diseases, Department of Medicine, Santa Clara Valley Medical Center, San Jose, California, USA, [7]Department of Rheumatology, John Muir Specialty Group, Walnut Creek, California, USA, and [8]Division of Rheumatology, Department of Medicine, Santa Clara Valley Medical Center, San Jose, California, USA

Background. The effectiveness of interleukin-6 inhibitors (IL-6i) in ameliorating coronavirus disease 2019 (COVID-19) remains uncertain.

Methods. We analyzed data for patients aged ≥18 years admitted with a positive severe acute respiratory syndrome coronavirus 2 polymerase chain reaction test at 4 safety-net hospital systems with diverse populations and high rates of medical comorbidities in 3 US regions. We used inverse probability of treatment weighting via machine learning for confounding adjustment by demographics, comorbidities, and disease severity markers. We estimated the average treatment effect, the odds of IL-6i effect on in-hospital mortality from COVID-19, using a logistic marginal structural model.

Results. Of 516 patients, 104 (20.1%) received IL-6i. Estimate of the average treatment effect adjusted for confounders suggested a 37% reduction in odds of in-hospital mortality in those who received IL-6i compared with those who did not, although the confidence interval included the null value of 1 (odds ratio = 0.63; 95% confidence interval, .29–1.38). A sensitivity analysis suggested that potential unmeasured confounding would require a minimum odds ratio of 2.55 to nullify our estimated IL-6i effect size.

Conclusions. Despite low precision, our findings suggested a relatively large effect size of IL-6i in reducing the odds of COVID-19–related in-hospital mortality.

Keywords. COVID-19; interleukin 6 inhibitors; cytokine release syndrome.

The coronavirus disease-19 (COVID-19) pandemic, which started in December 2019, has now spread worldwide causing more than 20 million cases and more than 750 000 deaths to date [1]. While the disease is asymptomatic in an estimated 18.5% to 32% of infected individuals, it produces severe disease in 1% to 27% of cases [2, 3]. Among hospitalized patients, a 24% mortality rate was reported in New York City [4].

Although the early phase of disease is associated with viral pathology, the individuals who progress to severe disease demonstrate features of a dysregulated immune response and cytokine release syndrome [5]. Interleukin (IL)-6 elevation is a prominent component of COVID-19 immunopathology and is associated with severe COVID-19 disease [6]. One study reported individuals with IL-6 levels >80 pg/mL were at 22-fold greater risk of mechanical ventilation [7]. Numerous inflammatory markers are elevated among severely ill COVID-19 patients, including C-reactive protein (CRP) [8]. CRP production is stimulated by IL-6 and their levels are well correlated in COVID-19 patients [7]. Hyperinflammation and macrophage activation syndrome associated with COVID-19 have also been hypothesized [9]. In addition, autopsies of patients who died from COVID-19 revealed findings consistent with macrophage activation syndrome in their mediastinal lymph nodes [10]. Thus, there may be a role for targeted IL-6 inhibitor therapies [11].

Tocilizumab, an IL-6 inhibitor (IL-6i), showed promise in a small observational study where it produced a rapid decrease in inflammatory markers and hypoxemia in severely ill COVID-19 patients [12]. Larger observational studies in Italy and the United States have found that use of IL-6i may be associated with mortality benefit [13–16]. However, other studies did not find significant benefits [17, 18]. In the absence of available data from randomized trials and due to the biological plausibility of IL-6 blockade, several hospitals in the United States opted to use IL-6i in severely ill COVID-19 patients guided by multidisciplinary teams.

After observing a lower than expected mortality in patients treated with IL-6i outside the structure of a randomized controlled trial, we sought to determine whether the

Received 21 August 2020; editorial decision 10 November 2020; accepted 16 November 2020; published online November 20, 2020.

[a]P. S. and S. R. J. contributed equally.

Correspondence: E. Y. Kissin, MD, 72 E. Newton St, E501, Boston, MA 02118 (eukissin@bu.edu).

The Journal of Infectious Diseases® 2021;223:581–8

DOI: 10.1093/infdis/jiaa717

effect was in fact due to the use of IL-6i. However, because patients with greater disease severity may have been considered for biologics therapy more often than those with milder COVID-19 symptoms, the possible effect of IL-6i on COVID-19–related mortality can be difficult to assess due to confounding by indication. Therefore, we aimed to emulate a target clinical trial of IL-6i on COVID-19–related in-hospital mortality by analyzing data from 4 hospitals systems in the United States with different utilization of 2 IL-6 inhibitors, tocilizumab and sarilumab, for COVID-19 disease using causal inference-based machine learning techniques to address potential confounding.

METHODS

Study Design, Setting, and Participants

We conducted a retrospective study of hospitalized patients aged ≥18 years with a positive polymerase chain reaction (PCR) test for severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). This multisite study included 4 participating hospital systems in 3 different regions in the United States, with differing utilization of IL-6i for COVID-19. Hospitals where biologics were used sparingly were included to help reduce the effects of confounding by indication.

We manually abstracted data from the electronic medical records of patients who met inclusion criteria between the dates of 1 March 2020 and 15 April 2020 including the following: (1) demographics and baseline clinical and laboratory data for factors that had previously been associated with disease morbidity [19, 20] (all laboratory testing was done at the discretion of the clinical teams); (2) medications administered for COVID-19, including hydroxychloroquine, glucocorticoids, remdesivir, lopinavir/ritonavir, IL-6 inhibitors, or IL-1 inhibitors (anakinra), which were prescribed according to the protocols of the various hospitals and/or decision by treating medical teams (we collected information on medications for COVID-19 administered both prior to IL-6i as well as during the rest of the hospitalization course); (3) the primary outcome of in-hospital mortality; and (4) other events during the hospitalization: transfer to the intensive care unit (ICU), intubation, occurrence of superinfection, and discharge from the hospital.

Institutional Procedures

At the hospital in which IL-6i was prescribed more frequently and in a protocolized manner (see below), a multidisciplinary group of physicians and pharmacists from the departments of infectious diseases, rheumatology, and pulmonary/critical care formed a committee to design a treatment algorithm for use of biologic agents. The World Health Organization-approved monitored emergency use of unregistered investigational interventions (MEURI) framework [21] was adopted and the committee closely monitored the safety of off-label use as well as clinical outcomes by performing iterative reviews of clinical data. A multidisciplinary ethics panel also reviewed the treatment protocol considerations using the MEURI framework.

The protocol for the use of the IL-6i tocilizumab and sarilumab at the high-utilization hospital required the following: ≥4 liters nasal cannula oxygen to maintain oxygen saturation >92% as well as CRP > 100 mg/L or lactate dehydrogenase (LDH) > 450 U/L, while not having a neutrophil count <100 cells/μL, platelet count <50 000 cells/μL, alanine transaminase > 200 IU/L, or history of diverticulitis or gastrointestinal bleed or ulceration within 90 days. Patients with documented bacterial infection, older than 90 years, with stage IV cancer not in remission, severe acute respiratory distress syndrome, shock, or sequential organ failure assessment score >11 were also excluded from treatment.

Patients meeting the inclusion criteria were given tocilizumab 400 mg fixed dose or sarilumab 200 mg fixed dose as an intravenous infusion over 60 minutes. Patients with the above oxygen requirements and ferritin >5000 ng/mL were instead given anakinra 200 mg twice daily or 100 mg every 6 hours for 3 days (dose adjusted for renal impairment when necessary). Individuals eligible for IL-6 inhibitors but with contraindications, such as gastrointestinal bleed, elevated aminotransferases, or absolute neutrophils <1000 cells/μL, were also given anakinra.

In the other 3 hospital systems, biologics were restricted to individuals who were critically or severely ill due to COVID-19 disease, and IL-6i was administered to COVID-19–infected patients guided by suspicion of COVID-19–induced cytokine storm on an ad hoc basis by a multidisciplinary team, and were administered much less frequently than at the other hospital. Severe illness indicating IL-6i use was defined if any of the following conditions were met: (1) respiratory rate ≥30 breaths/min; (2) $Spo_2 \leq 93\%$ while breathing room air; or (3) $Pao_2/Fio_2 \leq 300$ mmHg. Critical illness was defined if any of the following conditions were met: (1) respiratory failure requiring mechanical ventilation; (2) shock; or (3) respiratory combined with other organ failure, with ICU admission.

Patients in this study who received IL-6i did so off label, and not as part of another study. Ten of the 104 patients who received IL-6i also received a second course or treatment on an ad hoc basis.

Statistical Methods

We compared the in-hospital mortality of patients treated with IL-6i versus those who were not, using causal inference-based machine learning techniques to help address potential confounders. We identified potential confounders at baseline (ie, time point closest to admission) prior to any IL-6i exposure; these variables were identified based upon previous reports of association with disease morbidity and mortality [4, 19, 20]. We then used propensity score modeling and implemented an inverse probability weighting (IPW) approach. IPW is a method to adjust for measured confounders and involves calculating

propensity scores, that is an individual's probability of receiving a treatment based on observed covariates. Specifically, we used IPW [22] to fit a logistic marginal structural model (MSM) to estimate the effect of IL-6 inhibitors on in-hospital mortality [23]. To enhance robustness and reduce the possibility of bias in estimating propensity scores [24], we used a data-driven approach known as super learning [25], which is an ensemble of several parametric (such as logistic regression) and nonparametric machine learning prediction algorithms. Use of super learning reduces bias by avoiding unrealistic assumptions concerning the true functional form between covariates and treatment assignment mechanism, which may not be well represented if logistic regression alone were used, especially when a relatively large number of confounders are considered, which may result in data sparsity. Our super learner [26] included logistic regression, Bayesian logistic regression, Bayesian additive regression trees, stepwise regression based on Akaike information criterion, least absolute shrinkage and selection operator regression, random forest, recursive partitioning and regression trees, and extreme gradient boosting. Based on propensity scores, the IPW method assigns weights to observations such that a sample mimicking a target population is created, in which the distribution of confounders between treated and untreated groups is balanced (ie, adjusted for pretreatment confounders). We assessed the degree to which the groups were balanced by computing standardized mean differences (SMD) [27] in the original data before and after propensity score modeling and weighting. A standardized difference of 0.1 or lower implies negligible imbalances in mean or prevalence of covariates between treatment groups. Appropriate consideration of whether a covariate could be a potential confounder (ie, occurred pretreatment) or a potential mediator (ie, occurred posttreatment) was exercised such that no intermediate variable was included in the analysis [28]. Similarly, if a laboratory value was only available for a patient after the administration of treatment, the patient's pretreatment value for that variable was considered to be missing and was addressed by imputation because posttreatment variables could not have a confounding effect on treatment-outcome relationship. Average treatment effect estimates, expressed as odds ratios and corresponding confidence intervals, were then obtained from a logistic MSM by fitting a weighted logistic regression to the weighted data.

Imputation of missing data was done using multiple imputation by chained equations methodology [29] based on predictive mean matching to generate 250 completed data sets. Propensity score estimation by super learning and logistic MSM model fitting were repeated 250 times and the final results were pooled across 250 imputed data sets using Rubin's method [30]. Finally, to assess robustness of our effect measure estimate to unmeasured or uncontrolled confounding, we carried out a sensitivity analysis where we estimated the minimum strength of an association on odds ratio scale that potential unmeasured confounding should collectively have to fully explain away our observed treatment effect. The magnitude of such potential unmeasured confounding is represented by the *E* value statistic, where a large value (on odds ratio scale in our study) implies that a substantially strong unmeasured confounding needs to exist (which is less probable) to nullify the observed treatment effect [31]. All analyses in this study were performed in R statistical software package (version 4.0.2, R Foundation for Statistical Computing; https://www.R-project.org) [32].

We also descriptively documented medications received and other clinical events during the hospitalization course after IL-6i was provided.

All activities associated with this project were approved by the Institutional Review Boards of Boston University Medical Center, Jon Muir Health, Santa Clara Valley Medical Center, and the University of Wisconsin Medical Center.

RESULTS

The characteristics of the 4 hospital systems are shown in Table 1. The hospital with the greatest use of IL-6i had 318 COVID-19 patients included in the analysis, and the hospitals with lesser IL-6i use had 95, 48, and 55, respectively (Table 1).

Table 1. Hospital Characteristics

Hospital Information	High Utilization IL-6 Inhibitor Hospital (n = 318)	1st Low Utilization IL-6 Inhibitor Hospital (n = 95)	2nd Low Utilization IL-6 Inhibitor Hospital (n = 48)	3rd Low Utilization IL-6 Inhibitor Hospital (n = 55)
Geographic location	East coast	West coast	West coast	Midwest
Data collection date range	17 March 2020–8 April 2020	1 March 2020–11 April 2020	8 March 2020–10 April 2020	1 March 2020–15 April 2020
Patients hospitalized at end of follow-up, %	1	1	0	2
Medicaid population, %	50	60	4	13
No. of hospital beds pre-COVID	514	1182	799	953
No. ICU beds pre-COVID	51	114	83	150
Proning in ICU, date started	13 March 2020	4 April 2020	27 March 2020	15 March 2020
Patients treated with IL-6, No. (%)	87 (27.4)	10 (10.5)	4 (8.3)	3 (5.5)
Death, No. (%)	31 (9.7)	14 (14.7)	6 (12.5)	8 (14.5)

Of the 516 patients included in this study, 104 (20.1%) were exposed to IL-6i (Table 2). Compared to unexposed patients, the initial admission values for patients exposed to IL-6i showed significantly lower Pao_2:Fio_2 ratio (316.82 vs 373.72; $P < .001$), higher mean CRP (129.44 vs 98.23 mg/L; $P = .003$), higher mean LDH (453.83 vs 387.16 U/L; $P = .021$), and higher mean prothrombin time (16.69 vs 13.27 s; $P = .013$). Other baseline inflammatory markers were similar between the 2 groups (Table 2). In addition, patients exposed to IL-6i were slightly older, with higher body mass index, and had higher proportion of men, and more patients with cardiac and chronic kidney disease.

Potential confounders included in the propensity score model are shown in Table 2, which demonstrates that there were imbalances between exposed and unexposed groups in almost half of covariates in the original data, where estimates of SMDs were higher than 0.1 (indicating imbalance). Most SMDs reached <0.1, based upon the average SMDs across 250 imputed data sets that were balanced by propensity score modeling and weighting, indicating a negligible difference or imbalance in the final analytic datasets.

In-hospital mortality occurred in 13 (12.5%) and 46 (11.2%) of exposed and unexposed patients, respectively, in terms of unadjusted frequency. However, the estimate of average treatment effect by the logistic MSM indicated a 37% reduction in the odds of in-hospital mortality in those who were treated with IL-6i, compared with those who were not, although the confidence interval included the null value of 1 (odds ratio = 0.63; 95% CI, .29–1.38). In a sensitivity analysis, we obtained an *E* value of 2.55, which indicated the minimum strength required for potentially unmeasured confounding to nullify our 37% estimated reduction in the odds of in-hospital mortality in treated versus untreated patients. There was no interaction between admission to high utilization/low utilization hospitals and IL-6i on in-hospital mortality (exponentiated coefficient for interaction = 0.38; 95% CI, .06–2.43).

Additional COVID-19 treatments received in the IL-6i exposed versus unexposed group during the hospitalization are presented in Table 3, with the IL-6i patients' data illustrating treatments received prior to and after IL-6i. Remdesivir was dosed at 200 mg on the first day of administration and 100 mg per day for the next 4 days. Corticosteroid doses varied widely from 5 mg prednisone to 500 mg methylprednisolone per day as they were administered for many disparate reasons including asthma exacerbation and comorbid inflammatory arthritis as well as specifically for COVID-19. On average, patients received IL-6i on hospital day 3 (SD 1.9). Of the 104 IL-6i–exposed patients, 16 (15.4%) were already in the ICU or on mechanical ventilation when they received IL-6i, while 33 (24.6%) and 23 (22.1%) were later admitted to ICU and were put on mechanical ventilation, respectively. Of the unexposed patients, 73 (17.8%) required mechanical ventilation. Exposed patients were discharged alive 86% of the time, while this occurred in 88% of unexposed patients. Superinfection occurred in 14 (13.5%) and 50 (13.8%) of treated and untreated patients, respectively ($P = .84$). There were 6 cases of fungal infection, all in the untreated group (3 pulmonary aspergillosis, 1 central nervous system fluid with *Cryptococcus*, and 2 candidemia). There were no cases of mycobacterial infection. Overall, 3 (2.9%) patients who received IL-6i had a superinfection and died, while this occurred in 18 (4.3%) unexposed patients.

DISCUSSION

In this analysis, we found that IL-6 inhibitors lowered the odds of mortality in patients hospitalized with COVID-19 but the confidence interval around our estimated treatment effect included the null. While our study did not have enough precision to exclude the null, partly due to our study sample size, the magnitude of the estimated effect was substantial, that is a 37% reduction in odds of in-hospital mortality. This is similar to the effect size demonstrated in another recent study in the United States [33]. Our data are also concordant with the 40% and 45% risk reduction associated with IL-6i on combined endpoints of death and/or mechanical ventilation previously reported [34, 35]. In comparison, remdesivir, the only medication approved for treatment of COVID-19 in the United States, showed a hazard ratio for 14-day mortality of 0.70 (95% CI, .47–1.04) [36]. Dexamethasone, to date the only medication with a proven mortality benefit in COVID-19, produced a 17% reduction in odds of 28-day mortality (95% CI, .75–.93) [37]. However, there are preliminary reports that randomized controlled trials of tocilizumab (COVACTA) and sarilumab have not shown a survival benefit [38, 39]. The actual trial results are pending publication and thus issues regarding medication dosing, timing of treatment, and concomitant therapies or supportive care, patient-related factors, and comorbidities that might influence outcomes results remain to be seen. Two additional randomized controlled studies reported lack of benefit from tocilizumab on COVID-19–related mortality [40, 41]. However, one of these trials reported an adjusted tocilizumab hazards ratio for intubation or death of 0.66 (95% CI, .28–1.52), with less than half the number of patients enrolled compared to our study [40]. Furthermore, almost a third of patients also received remdesivir, which could have affected results, especially because only 3% of patient who received placebo died. The other trial also found a reduced hazards ratio of 0.58 for mechanical ventilation or death in the treatment group (consistent credible intervals .30–1.09), while enrolling an even smaller number of patients, one-fourth of our total [41]. A larger study, EMPACTA, with 389 more severely ill patients of whom 19.3% progressed to mechanical ventilation and/or death in the placebo arm, did find that tocilizumab reduced this outcome to 12.2% with a hazards ration of 0.56 (95% CI, .32–.97) [42].

Table 2. Baseline Characteristics of the Study Population

Characteristics	Unexposed (n = 412)	IL-6i Exposed (n = 104)	*P* Value	SMD for Raw Data	Average Absolute SMD Across 250 Imputed Datasets After Weighting
Age, mean (SD)	58.94 (17.09)	55.86 (16.23)	.097	0.185	0.178
Female sex	190 (46.1)	36 (34.6)	.045	0.236	0.096
White race[a]	236 (57.3)	64 (61.5)	.500	0.087	0.077
Medical Comorbidities					
BMI, mean (SD)	30.74 (7.47)	32.38 (9.39)	.060	0.193	0.004
Lung disease	66 (16.0)	17 (16.3)	1.000	0.009	0.005
Cardiac disease	85 (20.6)	15 (14.4)	.196	0.164	0.063
Diabetes	161 (39.1)	44 (42.3)	.625	0.066	0.026
Hemoglobin A1c, mean (SD)	8.13 (2.55)	8.35 (2.90)	.617	0.082	0.066
Hypertension	223 (54.1)	57 (54.8)	.988	0.014	0.043
Chronic kidney disease	58 (14.1)	7 (6.7)	.064	0.242	0.059
Transplant	8 (1.9)	0 (0.0)	.323	0.199	0.018
Immunodeficiency	10 (2.4)	3 (2.9)	1.000	0.009	0.007
Autoimmune disease	21 (5.1)	4 (3.8)	.783	0.061	0.023
Sickle cell disease	3 (0.7)	0 (0.0)	.880	0.121	0.007
Liver disease	15 (3.6)	3 (2.9)	.939	0.043	0.014
Systemic cancer	12 (2.9)	3 (2.9)	1.000	0.002	0.006
Obstructive sleep apnea	26 (6.3)	7 (6.7)	1.000	0.017	0.008
Pregnancy	6 (1.5)	2 (1.9)	1.000	0.036	0.011
Medications administered					
Colchicine	2 (0.5)	0 (0.0)	1.000	0.099	0.005
Steroids prehospital	23 (5.6)	3 (2.9)	.383	0.134	0.014
Steroids in hospital	15 (3.6)	3 (2.9)	.939	0.043	0.041
NSAID	59 (14.3)	9 (8.7)	.172	0.178	0.055
Hydroxychloroquine prehospital	13 (3.2)	1 (1.0)	.372	0.155	0.025
Hydroxychloroquine in-hospital pre-treatment	314 (76.2)	93 (89.4)	.005	0.356	0.018
Anakinra	21 (5.1)	0 (0.0)	.038		0.048
Immunosuppressants, no IL-6 or IL-1 inhibitors	23 (5.6)	1 (1.0)	.082	0.262	0.044
Disease Severity					
Pao_2: Fio_2 ratio, mean (SD)	373.72 (147.25)	316.82 (102.29)	<.001	0.449	0.133
Ferritin, ng/mL, mean (SD)	1148.85 (2793.09)	986.23 (998.50)	.574	0.078	0.137
C-reactive protein, mg/L, mean (SD)	98.23 (90.97)	129.44 (90.91)	.003	0.343	0.200
Procalcitonin, ng/mL, mean (SD)	1.00 (5.62)	0.63 (1.99)	.524	0.088	0.138
White blood cell count, thousands, mean (SD)	7.62 (5.45)	8.10 (6.84)	.451	0.077	0.050
Absolute neutrophils, thousands, mean (SD)	5.44 (3.23)	5.67 (3.06)	.515	0.073	0.071
Absolute lymphocytes, mean (SD)	1.29 (1.20)	1.88 (6.06)	.066	0.136	0.054
D-dimer, ng/mL FEU, mean (SD)	417.12 (1451.85)	330.21 (583.88)	.564	0.079	0.154
Prothrombin time, s, mean (SD)	13.27 (4.21)	16.69 (20.72)	.013	0.229	0.026
Troponin I, ng/mL, mean (SD)	0.06 (0.19)	0.02 (0.04)	.100	0.245	0.079
Lactate dehydrogenase, U/L, mean (SD)	387.06 (250.77)	453.83 (233.92)	.021	0.275	0.119
Pretreatment ICU or mechanical ventilation	0 (0.0)	16 (15.4)	<.001	0.603	0.061

Data are No. (%) except where indicated.

Abbreviations: BMI, body mass index; FEU, fibrinogen-equivalent units; ICU, intensive care unit; NSAID, nonsteroidal anti-inflammatory drug; SMD, standardized mean difference.

[a]Other races assessed were African American, Latinx, Native American, and Asian.

Although crossing of unity on a ratio scale may commonly be interpreted as the absence of a real effect, the American Statistical Association [43] and leading methodology experts [44, 45] strongly oppose a binary interpretation of confidence intervals based on an arbitrary threshold for drawing inference from a study's findings. While all values represented in a confidence interval (including the null) are compatible with data, the point estimate represents the most compatible value given the data. The precision of our provided effect estimate (ie, width of the corresponding confidence interval) is a function of the study sample size. Future studies can use our estimates and report a precision

Table 3. Treatments Received During Hospitalization

COVID-19 Treatments	Untreated (n = 412), No. (%)	IL-6i Treated (n = 104), No. (%)
Tocilizumab	0 (0.0)	79 (76)
Sarilumab	0 (0.0)	25 (24)
Glucocorticoids	27 (7)	13 (13)[a]
Hydroxychloroquine	314 (76)	101 (97)[b]
Remdesivir	2 (0.005)	2 (2)[c]
Lopinavir/ritonavir	1 (0.002)	0 (0)
Anakinra	21 (5.1)	0 (0)[d]

[a]Pre–IL-6i treatment, 5 and post–IL-6i treatment, 8.

[b]Pre–IL-6i treatment, 92 and post–IL-6i treatment 9.

[c]All post–IL-6i treatment.

[d]All in untreated group.

weighting calculation via meta-analysis to see if the findings remain consistent.

This study has several strengths. It is a multicenter study that includes a diverse patient population with a relatively high prevalence of preexisting comorbidities, which are often under-represented in most randomized controlled trials and in the studies that have so far been published on outcomes for using IL-6i in COVID-19–infected patients [46]. Our findings are therefore more generalizable. The inclusion of hospitals with varying utilization of IL-6i for off-label COVID-19 therapy allowed us to further enhance generalizability. Further, we used causal inference-based state-of-the-art statistical methods, including super learning for propensity scores estimation and IPW implementation [47]. Super learning does not rely on arbitrary and untestable assumptions about the relationship between covariates and outcome in data and allows incorporation of almost all available information, thus minimizing potential bias in fitting the propensity scores model. Using super learning in IPW implementation allowed us to fit a MSM, which is a family of causal models that is often used to emulate a target trial [48].

As in any observational study, our findings are subject to bias by unmeasured or uncontrolled confounding; however, we included data from hospital systems with varying utilization of IL-6i to limit confounding by indication (thus patients' exposure to IL-6i related not only to the degree of systemic inflammation and hypoxia but also to the hospital where they received care), and we provided a sensitivity analysis to estimate a magnitude of residual confounding. Our estimated *E* value is relatively large on the odds ratio scale, suggesting that considerable unmeasured confounding would be needed to nullify the estimated average treatment effect. The clinical information that we were not able to collect included date from onset of symptoms, and potentially detailed hospital-specific practice patterns and protocol differences. Importantly, it was difficult to control for the timing of IL-6i use in our observational study. While we appropriately adjusted for pretreatment confounding without improperly including any posttreatment intermediate variable, the timing of IL-6i with regard to the severity of disease may impact the effectiveness of therapy. For example, it is suggested that treatment administration in critical illness may not reverse the cytokine-mediated injury that has already occurred [16]. Additionally, although we considered tocilizumab and sarilumab to be equivalent in this study based on internal data that suggested similar rates in CRP reduction and similar reduction in intubation and in-hospital mortality (unpublished data), they may not be equally effective. Further, there may have been some secular changes in management of COVID-19 over the time period of observation that could impact outcomes such as in-hospital mortality.

In conclusion, we found a signal for the beneficial effect of IL-6i therapy on reduction of in-hospital mortality, albeit with low precision. Our findings can inform clinical care and research while we await further evidence from ongoing randomized control trials.

Notes

Acknowledgment. The authors thank the following hospital systems staff for assistance in data collection and/or review of IL6-i treatment protocols: Archana Asundi, Tamar F. Barlam, Joshua A. Barocas, Alison L. Blackman, Nahid Bhadelia, Diana F. Clarke, Dylan J. Clemens, Ellen R. Cooper, Megan R. Davis, Lauren M. Dzera, Muhammad Dhanani, Rachel L. Epstein, Maura A. Fagan, Davidson H. Hamer Leah Harvey, Finn J. Hawkins, Natasha S. Hochberg, Michael H. Ieong, Robin R. Ingalls, Raagini Jawa, Andrew J. Jenkins, Ghulam K. Khan, Simmeon Kimmel, David R. Leeds, Nina H. Lin, Benjamin P. Linas, Jai G. Marathe, Angela McLaughlin, Vagishwari Murugesan, Alison L. Nelson, Stephen I. Pelton, Bhavesh B. Shah, Amy B. Triche, Megan Victoria Alexander, Allan J. Walkey, Judy J. Wang, Lee M. Wetzler, A. Halem, A. Pop-Vicas, Brock M. Knapp, Sarah Ferree, Davuluri Srijana, and Alexander Miller.

Disclaimer. The funders had no role in the design, execution, or interpretation of this study.

Potential conflicts of interest. All authors: No reported conflicts of interest. All authors have submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Conflicts that the editors consider relevant to the content of the manuscript have been disclosed.

References

1. Johns Hopkins Coronairus Resource Center. COVID-19 data in motion, **2020**. https://coronavirus.jhu.edu/. Accessed 12 August 2020.
2. Payne DC, Smith-Jeffcoat SE, Nowak G, et al; CDC COVID-19 Surge Laboratory Group. SARS-CoV-2 infections and serologic responses from a sample of U.S. Navy service members–USS Theodore Roosevelt, April 2020. MMWR Morb Mortal Wkly Rep **2020**; 69:714–21.

3. Tabata S, Imai K, Kawano S, et al. Clinical characteristics of COVID-19 in 104 people with SARS-CoV-2 infection on the Diamond Princess cruise ship: a retrospective analysis. Lancet Infect Dis **2020**; 20:1043–50.
4. Petrilli CM, Jones SA, Yang J, et al. Factors associated with hospital admission and critical illness among 5279 people with coronavirus disease 2019 in New York City: prospective cohort study. BMJ **2020**; 369:m1966.
5. Siddiqi HK, Mehra MR. COVID-19 illness in native and immunosuppressed states: A clinical-therapeutic staging proposal. J Heart Lung Transplant **2020**; 39:405–7.
6. Gao Y, Li T, Han M, et al. Diagnostic utility of clinical laboratory data determinations for patients with the severe COVID-19. J Med Virol **2020**; 92:791–6.
7. Herold T, Jurinovic V, Arnreich C, et al. Elevated levels of IL-6 and CRP predict the need for mechanical ventilation in COVID-19. J Allergy Clin Immunol **2020**; 146:128–36.e4.
8. Liu J, Li S, Liu J, et al. Longitudinal characteristics of lymphocyte responses and cytokine profiles in the peripheral blood of SARS-CoV-2 infected patients. EBioMedicine **2020**; 55:102763.
9. Banu N, Panikar SS, Leal LR, Leal AR. Protective role of ACE2 and its downregulation in SARS-CoV-2 infection leading to macrophage activation syndrome: therapeutic implications. Life Sci **2020**; 256:117905.
10. Prilutskiy A, Kritselis M, Shevtsov A, et al. SARS-CoV-2 infection–associated hemophagocytic lymphohistiocytosis. Am J Clin Pathol **2020**; 154:466–74.
11. Fu B, Xu X, Wei H. Why tocilizumab could be an effective treatment for severe COVID-19? J Transl Med **2020**; 18:164.
12. Xu X, Han M, Li T, et al. Effective treatment of severe COVID-19 patients with tocilizumab. Proc Natl Acad Sci U S A **2020**; 117:10970–5.
13. Guaraldi G, Meschiari M, Cozzi-Lepri A, et al. Tocilizumab in patients with severe COVID-19: a retrospective cohort study. Lancet Rheumatol **2020**; 2:E474–84.
14. Ramiro S, Mostard RLM, Magro-Checa C, et al. 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. Ann Rheum Dis **2020**; 79:1143–51.
15. Price CC, Altice FL, Shyr Y, et al. Tocilizumab treatment for cytokine release syndrome in hospitalized patients with coronavirus disease-2019: survival and clinical outcomes. Chest **2020**; 158:1397–408.
16. Sinha P, Mostaghim A, Bielick CG, et al. Early administration of interleukin-6 inhibitors for patients with severe COVID-19 disease is associated with decreased intubation, reduced mortality, and increased discharge. Int J Infect Dis **2020**; 99:28–33.
17. Morena V, Milazzo L, Oreni L, et al. Off-label use of tocilizumab for the treatment of SARS-CoV-2 pneumonia in Milan, Italy. Eur J Intern Med **2020**; 76:36–42.
18. Campochiaro C, Della-Torre E, Cavalli G, et al; TOCI-RAF Study Group. Efficacy and safety of tocilizumab in severe COVID-19 patients: a single-centre retrospective cohort study. Eur J Intern Med **2020**; 76:43–9.
19. Cummings MJ, Baldwin MR, Abrams D, et al. Epidemiology, clinical course, and outcomes of critically ill adults with COVID-19 in New York City: a prospective cohort study. Lancet **2020**; 395:1763–70.
20. Wang K, Zuo P, Liu Y, et al. Clinical and laboratory predictors of in-hospital mortality in patients with COVID-19: a cohort study in Wuhan, China. Clin Infect Dis **2020**; 71:2079–88.
21. World Health Organization (WHO). Monitored emergency use of unregistered and experimental interventions (MEURI). Geneva, Switzerland: WHO, **2014**.
22. Cole SR, Hernán MA. Constructing inverse probability weights for marginal structural models. Am J Epidemiol **2008**; 168:656–64.
23. Hernán MA, Brumback B, Robins JM. Marginal structural models to estimate the causal effect of zidovudine on the survival of HIV-positive men. Epidemiology **2000**; 11:561–70.
24. Gruber S, Logan RW, Jarrín I, Monge S, Hernán MA. Ensemble learning of inverse probability weights for marginal structural modeling in large observational datasets. Stat Med **2015**; 34:106–17.
25. Pirracchio R, Petersen ML, van der Laan M. Improving propensity score estimators' robustness to model misspecification using super learner. Am J Epidemiol **2015**; 181:108–19.
26. van der Laan MJ, Polley EC, Hubbard AE. Super learner. Stat Appl Genet Mol Biol **2007**; 6:Article25.
27. Austin PC. An introduction to propensity score methods for reducing the effects of confounding in observational studies. Multivariate Behav Res **2011**; 46:399–424.
28. Hernán MA, Hernández-Díaz S, Robins JM. A structural approach to selection bias. Epidemiology **2004**; 15:615–25.
29. Schafer JL. Multiple imputation: a primer. Stat Methods Med Res **1999**; 8:3–15.
30. Rubin DB, Schenker N. Multiple imputation in health-care databases: an overview and some applications. Stat Med **1991**; 10:585–98.
31. VanderWeele TJ, Ding P. Sensitivity analysis in observational research: introducing the e-value. Ann Intern Med **2017**; 167:268–74.
32. Team RC. R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing, **2020**.

33. Biran N, Ip A, Ahn J, et al. Tocilizumab among patients with COVID-19 in the intensive care unit: a multicentre observational study. Lancet Rheumatol **2020**; 2:e603–12.
34. Guaraldi G, Meschiari M, Cozzi-Lepri A, et al. Tocilizumab in patients with severe COVID-19: a retrospective cohort study. Lancet Rheumatol **2020**; 2:e474–84.
35. Somers EC, Eschenauer GA, Troost JP, et al. Tocilizumab for treatment of mechanically ventilated patients with COVID-19 [published online ahead of print 11 July 2020]. Clin Infect Dis doi: 10.1093/cid/ciaa954.
36. Beigel JH, Tomashek KM, Dodd LE, et al. Remdesivir for the treatment of Covid-19—final report. N Engl J Med **2020**; 383:1813–26.
37. RECOVERY Collaborative Group, Horby P, Lim WS, et al. Dexamethasone in hospitalized patients with COVID-19–preliminary report [published online ahead of print 17 July 2020]. N Engl J Med doi:10.1056/NEJMoa2021436.
38. Roche. Roche provides an update on the phase III COVACTA trial of Actemra/RoActemra in hospitalised patients with severe COVID-19 associated pneumonia. https://www.roche.com/investors/updates/inv-update-2020-07-29.htm. Accessed 12 August 2020.
39. Sanofi. Sanofi and Regeneron provide update on Kevzara® (sarilumab) phase 3 U.S. trial in COVID-19 patients. https://www.sanofi.com/en/media-room/press-releases/2020/2020-07-02-22-30-00. Accessed 12 August 2020.
40. Stone JH, Frigault MJ, Serling-Boyd NJ, et al. Efficacy of tocilizumab in patients hospitalized with COVID-19 [published online ahead of print 21 October 2020]. N Engl J Med doi:10.1056/NEJMoa2028836.
41. Hermine O, Mariette X, Tharaux P et al. Effect of tocilizumab vs usual care in adults hospitalized with COVID-19 and moderate or severe pneumonia [published online ahead of print 20 October 2020]. JAMA Intern Med doi:10.1001/jamainternmed.2020.6820.
42. Roche. Roche's phase III EMPACTA study showed Actemra/RoActemra reduced the likelihood of needing mechanical ventilation in hospitalised patients with COVID-19 associated pneumonia, **18** September **2020**. https://www.roche.com/investors/updates/inv-update-2020-09-18.htm. Accessed 5 November 2020.
43. Wasserstein RL, Lazar NA. The ASA's statement on p-values: context, process, and purpose. Am Stat **2016**; 70: 129–33.
44. Greenland S, Senn SJ, Rothman KJ, et al. Statistical tests, *P* values, confidence intervals, and power: a guide to misinterpretations. Eur J Epidemiol **2016**; 31:337–50.
45. Greenland S. A serious misinterpretation of a consistent inverse association of statin use with glioma across 3 case-control studies. Eur J Epidemiol **2017**; 32:87–8.
46. Clark LT, Watkins L, Piña IL, et al. Increasing diversity in clinical trials: overcoming critical barriers. Curr Probl Cardiol **2019**; 44:148–72.
47. Hernán MA, Alonso A, Logan R, et al. Observational studies analyzed like randomized experiments: an application to postmenopausal hormone therapy and coronary heart disease. Epidemiology **2008**; 19:766–79.
48. Hernán MA, Robins JM. Using big data to emulate a target trial when a randomized trial is not available. Am J Epidemiol **2016**; 183:758–64.