Comparing the length of stay of COVID-19 patients receiving and not receiving hydroxychloroquine treatment using discrete-time Markov chain

Chunhui Gu
Department of Biostatistics, School of Public Health
University of Texas Health Science Center

Introduction

The pandemics of COVID-19 in the United States had led to about 30 million infections and 540 thousand deaths (Centers for Disease Control and Prevention, 2021) and still hasn't shown a sign of slowing down. It imposed an urgent need for an effective treatment for COVID-19 than any time. Since the emergence of COVID-19, a variety of treatments have been proposed to use off-label to treat severe COVID cases in hope of having some benefits. Among those treatments, hydroxychloroquine (HCQ) was considered as the most prospective one because its underlying pharmacological mechanism could block the pathway that coronavirus used to invade cells (Schrezenmeier & Dörner, 2020).

Many studies had been done in exploring the potential benefits of HCQ in treating COVID by evaluating whether there is a difference in the probability of encountering an adverse event or whether the time-to-event of adverse events changes (Jorge, 2020; Kashour et al., 2021). However, very few studies focused on the pattern of stay in different care areas after a COVID patient's admission, and there no study about how HCQ could change the pattern of transiting between different care areas.

The intensive Care Unit (ICU), considered as the last resort and the most valuable resource in a hospital, had experienced an unprecedented shortage after the break of COVID in many hospitals across the United States (Leatherby et al., 2020). Since it is very expensive to build new ICU capacity and could be a waste of resource if there is much less ICU need than expected, it is extremely crucial to know how many ICU beds are needed in the next few days for those newly admitted COVID patients.

Some studies have suggested that HCQ could shorten the length of stay in a hospital in COVID patients (Abdelrahman et al., 2021). But the analysis methods used in those studies could not give how patients bounce between different care areas.

In this study, we were interested in how the HCQ change the transition pattern among different care areas in COVID patients, especially between ICU and other care areas.

Method

Data extraction

The subjects we included in this study were retrieved from OptumCOVID, which was a massive electronic health record (EHR) database containing necessary information from COVID patients across the nation. The version which we used in this study stored records up to 2021 Jan 22. All queries were done MongoDB, a non-relational database management software through its Python interface.

Records of different kinds were recorded chronologically and stored in different collections of the database. A variety of constraints were enforced to retrieve variables needed as accurately as possible. (For example, the COVID diagnosis date was defined as the date of the earliest diagnosis after 2020/01/01. The completed list of definitions of variables in this study can be found in Supplementary document 1.)

Since reinfection rarely happened in a patient already infected by COVID-19 at the time that this study was done (Centers for Disease Control and Prevention, 2020). In this study, we assumed that each patient only got COVID infection one time and the earliest recorded date of COVID positive diagnosis records was treated as the COVID diagnosis time of that patient.

Eligibility of patients

Include Criterion: Patients who were diagnosed as COVID either by tests or professionals' judge between 2020/01/01 and 2021/01/22.

Exclude Criterion: 1) age < 18 and age >= 90. 2) Patients who have missing information in gender, ethnicity, division, age, smoking status, and BMI.

375460 patients were included as a result.

HCQ treatment and control groups

A patient was assigned to the HCQ treatment group if it had a record of HCQ administration within 7 days after being diagnosed. The indications that HCQ was used were unclear for patients in the database and regular contraindications of HCQ were assumed. Patients in the treatment group and control group were called cases and controls respectively for short.

Propensity score matching of control patients to treatment patients

Propensity score matching (PSM) was used to adjust other possible confounders to the treatment effect of the HCQ. PSM was performed by R 4.0.2 and package MatchIt 4.0.1. Logistic regression was fit to predict the potential of being assigned with HCQ treatment based on

characteristics of patients that can be classified into two categories: 1) baseline demographic characteristics, including age, gender, ethnicity, region division. 2) clinical data about the baseline severity that close to the day of usage of HCQ (records within one week after being diagnosed as COVID, including comorbidity, usage of antibiotics, or Immunosuppressive drugs. Another special covariate balanced by matching is an indictor variable whether a patient first diagnosis date was later than 06/15/2020, which was the date that FDA revoked the emergency use authorization (EUA) to use hydroxychloroquine and chloroquine to treat COVID-19 that was not related to clinical trials (Food and Drug Administration, 2020). For the completed list of variables and how they were defined please check the supplementary table 1.

The matched data were generated based on a fixed ratio for cases versus controls using the nearest neighbor matching algorithm without replacement. 1:1, 1:2, and 1:4 ratios were tested and the one with a good matching result and included more patients were used in the final analysis. The average treatment effect on the population who received it as a result.

Analysis method

A homogenous discrete-time Markov chain (DTMC) model was used to estimate the transition probabilities between different care areas for the treatment group and control group respectively by R 4.0.2 and package markovchain 0.8.5-4. Maximum likelihood estimator "MLE" was used in the "method" parameter and kept the left parameters as default.

Expected lengths of stay in those care areas from day 1 to day 30 for patients in the two treatment groups and with different initial states were estimated by using the estimated transition probability matrix from the fitted model (Pérez et al., 2006).

Model description:

States of the Markov Chain model

The original care areas were classified into three main states: Intensive Care Unit (ICU), high-dependency units, and other hospital services or self-care. The high-dependency units included "CARDIAC CATHETERIZATION", "URGENT CARE UNIT", "TELEMETRY UNIT", "DELIVERY ROOM", "EMERGENCY DEPARTMENT", "OPERATING ROOM". Other hospital services or self-care included "AMBULATORY CARE", "AMBULATORY SURGERY", "MEDICAL OR SURGICAL UNIT", "PEDIATRIC UNIT", "PSYCHIATRIC UNIT", "REHABILITATION UNIT", "HOSPICE CARE AREA", "LABORATORY", "DAY CARE UNIT", "POST ANESTHESIA CARE UNIT (PACU) / RECOVERY ROOM", "ANCILLARY SERVICES", and "OTHER CARE AREA". The ICU, high-dependency units, and other hospital services or self-care were called state 1, state 2, and state 3 respectively in the DTMC model.

When a patient had an event during a hospital visiting, a corresponding record was generated in the database with the date and care area it happened. For the treatment group, only records after HCQ administration were extracted. For the interval time between two records, we filled it with the care area of the last available record of that patient to create a sequence of each patient's stay. The unit of time was one day for those sequences. For each patient, only records

within 30 days after the patient was diagnosed with COVID positive were extracted. The actual length of the sequence of a patient could less than 30 units if the patient entered the database within less than 30 days from the updated date of the database.

The transition probability matrix

Two different three-state transition probability matrices, P1 and P2, were assumed for the treatment group and control group respectively. There is no restriction for any transition probability between those states other than the row sum must be 1 for each row of the transition probability matrices (Table 1).

Table 1: The transition probability matrix structure

	ICU (state 1)	High-dependency units (state 2)	Other hospital services or self-care (state 3)
ICU (state 1)	P11	P12	P13
High-dependency units (state 2)	P21	P22	P23
Other hospital services or self-care (state 3)	P31	P32	P33

Assumption of the model

Two key assumptions were made for the DTMC model: the Markov property and the stationary property.

The Markov property assumes that the stay of the next time unit only depends on the stay of the current time unit. For example, the probability of transferring from ICU to the high-dependency units is the same no matter the previous care areas she or he stayed before current ICU stay.

Its mathematical form is:
$$P(X_n = x_n | X_{n-1} = x_{n-1}, X_{n-2} = x_{n-2}, \ldots. X_1 = x_1) = P(X_n = x_n | X_{n-1} = x_{n-1})$$

Where X_n is the state in the n-th time unit.

The stationary property assumes that the transition probabilities between states are independent of time. For example, the probability of a patient's transferring from the ICU to the high-dependency units after one day on the first day after its diagnosis is the same as the probability of the patient's transferring from the ICU to the high-dependency units in the 20th day after its diagnosis.

Its mathematical form is:
$$P(X_n = x_n | X_{n-1} = x_{n-1}) = P(X_m = x_m | X_{m-1} = x_{m-1})$$
Where X_n is the state in the n-th time unit, and m is another time point that is different than n.

Simulation:
The simulation was used to check whether the model was appropriate for the distribution of the data and the purpose of the study.

The following transition probability matrix was used as the true underlying true transition probability matrix **P** in simulation (Table 2). This probability transition matrix reflected our belief that it was more possible to transfer to high-dependency units from ICU rather than transfer to low dependent units directly from ICU.

Table 2. transition probability matrix for simulation

	ICU (state 1)	High-dependency units (state 2)	Other hospital services or self-care (state 3)
ICU (state 1)	0.90	0.08	0.02
High-dependency units (state 2)	0.02	0.90	0.08
Other hospital services or self-care (state 3)	0.01	0.01	0.98

The "rmarkovchain" function in the "markovchain" package was used to generate simulation data. First, we generated the initial states of 3000 virtual "patients" with probabilities of 0.1, 0.65, and 0.25 to be in "ICU", "High-dependency units", and "Other hospital services or self-care" respectively. Then for each of those "patients", we generated a sequence of 30 units based on its initial state and the transition probability matrix **P**. Then we fitted the simulated data with the proposed model with a 95% confidence interval (95% CI) and check whether the 95% CIs covered each of the true underlying parameters (the 9 true underlying transition probabilities in the cell of the transition matrix in here). We repeated this process for 1000 times and calculated the percentage that the 95% CIs covered the true parameters as the coverage probabilities (CPs). The CPs from the simulation for the transition probabilities was 1.000 (state1 to state1), 0.948 (state2 to state1), 0.952 (state3 to state1), 0.966(state1 to state2), 1.000 (state2 to state2), 0.959 (state2 to state3), 0.962(state3 to state1), 0.956(state 3 to state2), and 1.000 (state3 to state3). The CPs were close to 95%, which indicated that the model was appropriate for the data.

Table3: Coverage probabilities for the transition probabilities between states

	ICU (state 1)	High-dependency units (state 2)	Other hospital services or self-care (state 3)
ICU (state 1)	1.000	0.966	0.962
High-dependency units (state 2)	0.948	1.000	0.956
Other hospital services or self-care (state 3)	0.952	0.959	1.000

Analysis results:

A ratio of 1:2 was used for matching cases versus controls. Good balance was achieved between the case and control group with the absolute standard difference (ASD) of all baseline variables were less than 0.1 (Figure 1) and covariates were balanced between the treatment and control group (Table 4). After matching, there were 4439 patients in the treatment group and 8878 patients in the control group. 429674 unmatched patients in the control group were discarded.
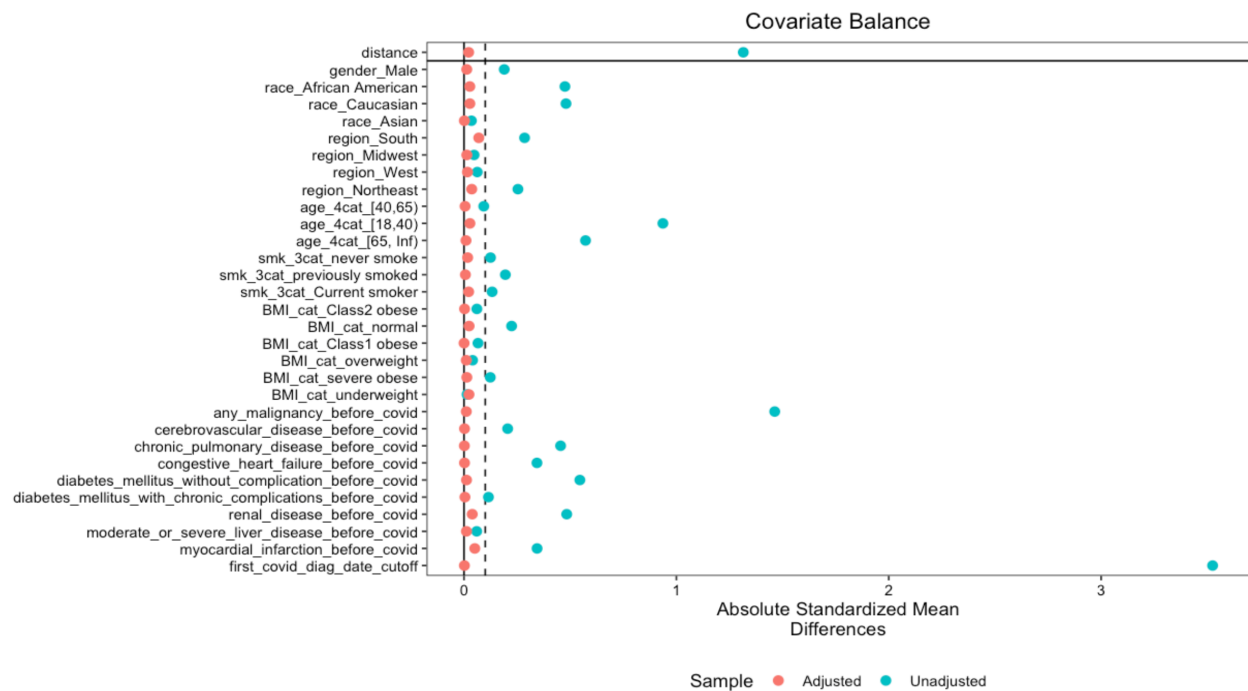
Figure1. Absolute standard mean difference (ASD) of covariates
Blue dots: ASD of covariates before matching. Red dots: ASD of covariates after matching.

Table 4. Descriptive table for covariates

	0 (N = 372198)	1 (N = 3659)
age		
[18,40)	115,762 (31)	259 (7)
[40,65)	168,455 (45)	1,489 (41)
[65, Inf)	87,981 (24)	1,911 (52)
gender		
Female	216,627 (58)	1,784 (49)
Male	155,571 (42)	1,875 (51)
race		
Caucasian	312,747 (84)	2,216 (61)
African American	50,933 (14)	1,338 (37)
Asian	8,518 (2)	105 (3)
region		
South	68,832 (18)	364 (10)
Midwest	198,987 (53)	1,869 (51)
West	19,034 (5)	143 (4)
Northeast	85,345 (23)	1,283 (35)
BMI		
underweight	4,298 (1)	37 (1)
normal	79,873 (21)	503 (14)
overweight	109,013 (29)	1,006 (27)
Class1 obese	85,515 (23)	946 (26)
Class2 obese	49,246 (13)	564 (15)
severe obese	44,253 (12)	603 (16)
Any malignancy before COVID		
Yes	44,271 (12)	2,749 (75)
No	327,927 (88)	910 (25)
Cerebrovascular disease before COVID		
Yes	2,965 (1)	200 (5)
No	369,233 (99)	3,459 (95)
Chronic pulmonary disease before COVID		
Yes	15,494 (4)	857 (23)
No	356,704 (96)	2,802 (77)
Congestive heart failure before COVID		
Yes	9,072 (2)	532 (15)
No	363,126 (98)	3,127 (85)
Diabetes mellitus without complication before COVID		
Yes	16,797 (5)	1,074 (29)
No	355,401 (95)	2,585 (71)
Diabetes mellitus with chronic complications before COVID		
Yes	549 (0)	58 (2)
No	371,649 (100)	3,601 (98)
Renal disease before COVID		
Yes	10,540 (3)	851 (23)
No	361,658 (97)	2,808 (77)
Moderate or severe liver disease before COVID		
Yes	339 (0)	19 (1)
No	371,859 (100)	3,640 (99)
Myocardial_infarction before COVID		
Yes	4,411 (1)	461 (13)
No	367,787 (99)	3,198 (87)
COVID diagnosis date after 2020-06-15		
Yes	317,779 (85)	199 (5)
No	54,419 (15)	3,460 (95)

	0 (N = 7318)	1 (N = 3659)
age		
[18,40)	478 (7)	259 (7)
[40,65)	2,979 (41)	1,489 (41)
[65, Inf)	3,861 (53)	1,911 (52)
gender		
Female	3,622 (49)	1,784 (49)
Male	3,696 (51)	1,875 (51)
race		
Caucasian	4,520 (62)	2,216 (61)
African American	2,596 (35)	1,338 (37)
Asian	202 (3)	105 (3)
region		
South	919 (13)	364 (10)
Midwest	3,671 (50)	1,869 (51)
West	296 (4)	143 (4)
Northeast	2,432 (33)	1,283 (35)
BMI		
underweight	75 (1)	37 (1)
normal	1,047 (14)	503 (14)
overweight	2,031 (28)	1,006 (27)
Class1 obese	1,828 (25)	946 (26)
Class2 obese	1,153 (16)	564 (15)
severe obese	1,184 (16)	603 (16)
Any malignancy before COVID		
Yes	5,516 (75)	2,749 (75)
No	1,802 (25)	910 (25)
Cerebrovascular disease before COVID		
Yes	390 (5)	200 (5)
No	6,928 (95)	3,459 (95)
Chronic pulmonary disease before COVID		
Yes	1,728 (24)	857 (23)
No	5,590 (76)	2,802 (77)
Congestive heart failure before COVID		
Yes	1,047 (14)	532 (15)
No	6,271 (86)	3,127 (85)
Diabetes mellitus without complication before COVID		
Yes	2,093 (29)	1,074 (29)
No	5,225 (71)	2,585 (71)
Diabetes mellitus with chronic complications before COVID		
Yes	106 (1)	58 (2)
No	7,212 (99)	3,601 (98)
Renal disease before COVID		
Yes	1,609 (22)	851 (23)
No	5,709 (78)	2,808 (77)
Moderate or severe liver disease before COVID		
Yes	38 (1)	19 (1)
No	7,280 (99)	3,640 (99)
Myocardial_infarction before COVID		
Yes	788 (11)	461 (13)
No	6,530 (89)	3,198 (87)
COVID diagnosis date after 2020-06-15		
Yes	392 (5)	199 (5)
No	6,926 (95)	3,460 (95)

Left: distribution of covariates before matching. Right: distribution of covariates after matching.
Before matching, we could see an obvious imbalance of each of those covariates. After matching the distribution of those covariates are similar.

There was no obvious difference in the transition probabilities between the treatment group and the control group (Table 2 and Table 3). The 95% CI of the transition probability of each of those state-wise transitions from the treatment and control group were overlapped with each other.

Table 5. Estimated transition probability matrix for the treatment group

	ICU (state 1)	High-dependency units (state 2)	Other hospital services or self-care (state 3)
ICU (state 1)	0.941	0.002	0.057
High-dependency units (state 2)	0.010	0.946	0.044
Other hospital services or self-care (state 3)	0.008	0.006	0.986

Table 6. Estimated transition probability matrix for the control group

	ICU (state 1)	High-dependency units (state 2)	Other hospital services or self-care (state 3)
ICU (state 1)	0.929	0.006	0.065
High-dependency units (state 2)	0.007	0.946	0.046
Other hospital services or self-care (state 3)	0.006	0.012	0.982

From the expected lengths of stay in those three care areas, we found that when a patient was in ICU on her or his COVID diagnosis date, HCQ didn't help to reduce the length of stay in the high-dependency units or ICU. But if a patient was original in other hospital services or self-care on her or his COVID diagnosis date, HCQ could reduce the length of stay in the high-dependency units or ICU by about one day during a 30-day care period (Table 7-9).

Table 7: Estimated stay in care areas when the initial stay is in ICU

day	ICU		high-dependency units		other hospital units or self-care	
	Control	Treatment	Control	Treatment	Control	Treatment
1	1	1	0	0	0	0
2	1.92873423	1.94111623	0.0057742	0.00221881	0.06549157	0.05666496
3	2.79169032	2.82727756	0.01739283	0.00673617	0.19091685	0.16598627
4	3.59393432	3.66215116	0.03488685	0.01361291	0.37117883	0.32423593
5	4.34014308	4.44915991	0.05825322	0.02289301	0.6016037	0.52794708
6	5.03463404	5.19149866	0.08745955	0.03460546	0.87790641	0.77389588
7	5.68139283	5.89214937	0.12244832	0.04876605	1.19615885	1.05908458
8	6.28409868	6.55389531	0.16314065	0.06537883	1.55276066	1.38072586
9	6.84614797	7.17933424	0.2094396	0.08443762	1.94441243	1.73622814
10	7.37067587	7.77089079	0.26123317	0.10592721	2.36809096	2.123182

11	7.86057644	8.33082794	0.31839692	0.12982458	2.82102665	2.53934747
12	8.31852109	8.86125783	0.38079633	0.15609993	3.30068259	2.98264223
13	8.7469757	9.36415176	0.44828885	0.18471762	3.80473545	3.45113062
14	9.14821637	9.84134956	0.52072576	0.21563703	4.33105787	3.94301341
15	9.52434402	10.2945684	0.59795372	0.24881333	4.87770226	4.45661829
16	9.87729783	10.7254108	0.67981623	0.28419819	5.44288594	4.99039098
17	10.2088676	11.1353726	0.76615483	0.32174039	6.02497754	5.542887
18	10.5207055	11.5258497	0.85681017	0.36138636	6.62248438	6.11276399
19	10.814336	11.8981447	0.95162296	0.40308072	7.23404101	6.69877456
20	11.0911666	12.2534736	1.05043477	0.44676669	7.85839861	7.29975967
21	11.3524961	12.5929711	1.15308869	0.4923865	8.49441526	7.91464238
22	11.599523	12.9176961	1.25942991	0.53988173	9.14104706	8.54242214
23	11.8333538	13.2286369	1.36930627	0.58919364	9.79733989	9.18216945
24	12.0550095	13.5267157	1.4825686	0.64026344	10.4624219	9.83302081
25	12.2654321	13.8127933	1.59907112	0.69303253	11.1354968	10.4941741
26	12.4654914	14.0876729	1.71867166	0.74744269	11.815837	11.1648844
27	12.6559896	14.352104	1.84123193	0.80343629	12.5027784	11.8444597
28	12.8376674	14.6067861	1.96661766	0.86095644	13.1957149	12.5322574
29	13.0112081	14.8523721	2.09469873	0.91994712	13.8940932	13.2276808
30	13.1772422	15.089471	2.22534928	0.98035332	14.5974085	13.9301757

Table 8: Estimated stay in care areas when the initial stay is in the high-dependency units

day	ICU		high-dependency units		other hospital units or self-care	
	Control	Treatment	Control	Treatment	Control	Treatment
1	0	0	1	1	0	0
2	0.00754227	0.00974077	1.94632709	1.94611155	0.04613064	0.04414768
3	0.02194177	0.02846659	2.84246322	2.8415173	0.13559501	0.13001611
4	0.04257293	0.0554784	3.69166569	3.68921259	0.26576137	0.25530901
5	0.06886467	0.09012965	4.49698156	4.49201642	0.43415377	0.41785394
6	0.10029576	0.13182245	5.2612611	5.25258179	0.63844314	0.61559576
7	0.13639064	0.18000401	5.98717044	5.97340545	0.87643892	0.84659054
8	0.17671553	0.2341633	6.67720339	6.65683706	1.14608108	1.10899964
9	0.22087488	0.29382803	7.3336925	7.30508779	1.44543262	1.40108418
10	0.26850812	0.35856172	7.95881943	7.92023852	1.77267245	1.72119976
11	0.31928667	0.42796111	8.55462466	8.50424743	2.12608867	2.06779146

12	0.37291122	0.50165368	9.12301653	9.05895722	2.50407224	2.4393891
13	0.4291092	0.57929539	9.66577981	9.58610193	2.90511098	2.83460268
14	0.4876325	0.66056855	10.1845836	10.0873133	3.32778389	3.25211819
15	0.54825535	0.74517988	10.6809888	10.5641267	3.77075582	3.69069347
16	0.6107724	0.83285866	11.1564552	11.0179869	4.23277244	4.14915442
17	0.67499694	0.92335509	11.6123476	11.4502536	4.71265545	4.62639135
18	0.7407593	1.01643868	12.0499426	11.8622058	5.20929811	5.12135549
19	0.80790535	1.11189687	12.4704336	12.2550474	5.72166101	5.63305575
20	0.87629516	1.20953361	12.8749368	12.6299108	6.24876805	6.1605556
21	0.94580173	1.30916817	13.2644956	12.9878617	6.78970269	6.70297011
22	1.01630987	1.41063397	13.6400857	13.3299028	7.34360443	7.25946319
23	1.08771518	1.51377752	14.0026194	13.6569775	7.9096654	7.82924495
24	1.15992305	1.61845741	14.3529497	13.9699734	8.48712728	8.41156916
25	1.23284783	1.72454342	14.6918739	14.2697257	9.07527827	9.00573091
26	1.30641203	1.83191565	15.0201376	14.55702	9.67345033	9.61106437
27	1.38054554	1.94046377	15.3384379	14.8325956	10.2810166	10.2269406
28	1.45518505	2.05008623	15.6474263	15.097148	10.8973886	10.8527658
29	1.53027338	2.16068964	15.947712	15.3513316	11.5220146	11.4879788
30	1.60575893	2.27218813	16.2398645	15.5957619	12.1543766	12.1320499

Table 9: Estimated stay in care areas when the initial state is in the other hospital units or self-care

	ICU		high-dependency units		other hospital units or self-care	
day	Control	Treatment	Control	Treatment	Control	Treatment
1	0	0	0	0	1	1
2	0.00557712	0.00776414	0.01208786	0.0058231	1.98233502	1.98641276
3	0.01632656	0.02278646	0.03543346	0.01709362	2.94823999	2.96011992
4	0.03187297	0.04459399	0.06926382	0.03346011	3.89886321	3.92194589
5	0.05186836	0.07274459	0.11285917	0.05459384	4.83527246	4.87266157
6	0.07599004	0.10682492	0.16554928	0.08018729	5.75846068	5.8129878
7	0.1039388	0.14644859	0.22671	0.10995277	6.66935119	6.74359864
8	0.13543722	0.1912544	0.29576015	0.1436212	7.56880263	7.6651244
9	0.17022802	0.24090468	0.3721585	0.18094085	8.45761349	8.57815447

10	0.20807264	0.29508379	0.45540103	0.22167623	9.33652633	9.48323999
11	0.24874989	0.35349662	0.54501839	0.26560706	10.2062317	10.3808963
12	0.29205461	0.41586734	0.64057348	0.31252728	11.0673719	11.2716054
13	0.33779657	0.48193806	0.74165924	0.36224413	11.9205442	12.1558178
14	0.38579936	0.55146772	0.84789658	0.41457727	12.7663041	13.033955
15	0.43589937	0.62423095	0.95893245	0.46935801	13.6051682	13.906411
16	0.48794487	0.70001708	1.07443807	0.52642855	14.4376171	14.7735544
17	0.54179516	0.77862917	1.19410721	0.58564127	15.2640976	15.6357296
18	0.59731972	0.85988311	1.31765468	0.64685807	16.0850256	16.4932588
19	0.65439752	0.94360679	1.44481486	0.70994976	16.9007876	17.3464434
20	0.71291632	1.0296393	1.57534039	0.77479549	17.7117433	18.1955652
21	0.77277198	1.11783025	1.70900084	0.84128219	18.5182272	19.0408876
22	0.83386794	1.208039	1.84558163	0.90930411	19.3205504	19.8826569
23	0.89611464	1.30013413	1.98488289	0.97876229	20.1190025	20.7211036
24	0.959429	1.39399274	2.12671843	1.04956419	20.9138526	21.5564431
25	1.02373394	1.48949996	2.27091488	1.1216232	21.7053512	22.3888768
26	1.088958	1.58654839	2.41731072	1.19485833	22.4937313	23.2185933
27	1.15503486	1.68503765	2.5657555	1.26919378	23.2792096	24.0457686
28	1.22190302	1.78487387	2.71610912	1.34455867	24.0619879	24.8705675
29	1.28950543	1.88596933	2.86824104	1.42088667	24.8422535	25.693144
30	1.35778916	1.988242	3.02202969	1.49811573	25.6201811	26.5136423

Discussion:

This study had some advantages over other studies that focused on the treatment effect of HCQ. First, we had a large size of HCQ treatment cases, strictly matched controls from a huge database, as well as the number of confounders taken into consideration during matching. Second, the usage of the DTMC model allowed us not only to estimate the length of stay in ICU within a period but for the other two care areas at the same time given whether a patient received HCQ as treatment and the care area she or he was on the COVID diagnosis date.

For the Markov property and stationary property assumption, there is no effective way to test that. Since the transition probabilities between care areas could change along with time after diagnosed as COVID, we only extracted records up to 30 days after diagnosis for the sake of the stationary property assumption.

Due to the large sample size, we could get a relatively concise result from the DTMC model. Although there was a reduced length of stay in ICU and high-dependency units after HCQ administration when the initial stay was in other hospital services and home, the difference was not clinically significant. The evidence from this study still suggested not to use HCQ as a standard treatment for COVID.

The tables that estimated expected lengths of stay in those three care areas can be used to estimate how long a patient will stay in ICU and high-dependency units given her or his initial care area when diagnosed as COVID. It will help administers of hospitals to decide whether they need new ICU capacity and if yes how many new ICU beds will be needed for the next 30 days for those newly diagnosed COVID patients.

Reference:

Abdelrahman, Z., Liu, Q., Jiang, S., Li, M., Sun, Q., Zhang, Y., & Wang, X. (2021). Evaluation of the Current Therapeutic Approaches for COVID-19: A Systematic Review and a Meta-analysis. *Frontiers in Pharmacology*, *12*. https://doi.org/10.3389/fphar.2021.607408

Centers for Disease Control and Prevention. (2020). *Reinfection with COVID-19*. https://www.cdc.gov/coronavirus/2019-ncov/your-health/reinfection.html

Centers for Disease Control and Prevention. (2021). *COVID Data Tracker*. United States COVID-19 Cases and Deaths by State. https://covid.cdc.gov/covid-data-tracker/#datatracker-home

Food and Drug Administration. (2020). *FDA cautions against use of hydroxychloroquine or chloroquine for COVID-19 outside of the hospital setting or a clinical trial due to risk of heart rhythm problems*. https://www.fda.gov/drugs/drug-safety-and-availability/fda-cautions-against-use-hydroxychloroquine-or-chloroquine-covid-19-outside-hospital-setting-or

Jorge, A. (2020). Hydroxychloroquine in the prevention of COVID-19 mortality. *The Lancet Rheumatology*. https://doi.org/10.1016/s2665-9913(20)30390-8

Kashour, Z., Riaz, M., Garbati, M. A., Aldosary, O., Tlayjeh, H., Gerberi, D., Murad, M. H., Sohail, M. R., Kashour, T., & Tleyjeh, I. M. (2021). Efficacy of chloroquine or hydroxychloroquine in COVID-19 patients: a systematic review and meta-analysis. *Journal of Antimicrobial Chemotherapy*, *76*(1), 30–42. https://doi.org/10.1093/jac/dkaa403

Leatherby, L., Keefe, J., Tompkins, L., Smart, C., & Conlen, M. (2020). *There's No Place for Them to Go': I.C.U. Beds Near Capacity Across U.S.* https://www.nytimes.com/interactive/2020/12/09/us/covid-hospitals-icu-capacity.html

Pérez, A., Chan, W., & Dennis, R. J. (2006). Predicting the length of stay of patients admitted for intensive care using a first step analysis. *Health Services and Outcomes Research Methodology*, *6*(3–4), 127–138. https://doi.org/10.1007/s10742-006-0009-9

Schrezenmeier, E., & Dörner, T. (2020). Mechanisms of action of hydroxychloroquine and chloroquine: implications for rheumatology. *Nature Reviews Rheumatology*, *16*(3), 155–

166. https://doi.org/10.1038/s41584-020-0372-x

Supplementary document1:

Variable definitions

Variable name:
first_covid_diag_date
Description:
The earliest date that a patient was diagnosed as COVID
Definition:
There two conditions:
For each patient, the earliest date of a document in the "diagnosis" collection that had the
"**diagnosis_cd**" field equal to 'U071' after the **COVID_START (2020/01/01)**.

For each patient, the earliest date of documents in the **"labs"** collection with one of
the **TEST_CODES** in the **"test_code"** field **OR** with one of the **TEST_NAMES** in
the **"test_name"** field, then if the corresponding **"test_result"** fields are one of the values in
the **TEST_RESULTS**. (Please check the supplementary files for possible values of
TEST_CODES, TEST_NAME, and TEST_RESULTS.

The earliest date of the COVID diagnosis date of those two approaches was used as the
first_covid_diag_date.

medadm_drug_name_hcq_use
Description:
Whether a patient has Hydroxychloroquine administration
Definition:
For a patient, if there was a record of Hydroxychloroquine administration within 7 days after COVID
positive diagnosis date then she or he is in the HCQ treatment group. The Hydroxychloroquine
administration records were retrieved from the "cov_proc_pt_timeline" collection with the
"DRUG_NAME" field matched "Hydroxychloroquine" by the regular expression with the case-
insensitive option, or "GENERIC_DESC" matched "'HYDROXYCHLOROQUINE SULFATE" by the
regular expression with case-insensitive option.

hcq_usage_date
Description:
The date that patients start HCQ administration after diagnosed as COVID positive.
Definition:
The date of the earliest record within 7 days after COVID positive diagnosis date

Regular demographic variables: **age, gender, race, region,**
Description:
Demographic variables that describe the socioeconomic status
Definition:

Those variables are extracted from the fields in a document that is unique for each patient in the "patients" collection. Age was the original continuous variable and was divided into 3 categories 18-40, 40-65, and 65-90 (left bound inclusive and right bound exclusive).

BMI
Description:
A 5-level categorical variable of the body-mass-index
Definition:
The latest document with "obs_type" as "BMI" before the first_covid_diag_date. The original value is continuous and was divided into 5 categories. Categorization was based on the criterion on the BMI page of the CDC website. Evidence about the correlation between obesity and COVID severity could be found at https://www.cdc.gov/obesity/data/obesity-and-covid-19.html.

Smoke status
The latest document with "obs_type" as "smoke" before the first_covid_diag_date. The original variable had 7 levels and was collapsed into 3 levels. "Previously smoked" and "Not currently smoking" were treated as "Previously smoked". "Unknown smoking status", "Other smoking status", and "Not recorded" were treated as missing.

Comorbidity-related variables:
**any_malignancy_before_covid,
cerebrovascular_disease_before_covid,
chronic_pulmonary_disease_before_covid,
congestive_heart_failure_before_covid,
diabetes_mellitus_without_complication_before_covid,
diabetes_mellitus_with_chronic_complications_before_covid,
renal_disease_before_covid,
moderate_or_severe_liver_disease_before_covid,
myocardial_infarction_before_covid**

Description:
Whether a patient has a certain kind of comorbidity before HCQ administration.

Definition:
If a patient had records of a certain group of ICD-10 codes of a kind of comorbidity within one week after the COVID diagnosis date, then the corresponding comorbidity variable was 1 otherwise 0.

Take "any_malignancy_before_covid" as example, "**diagnosis_cd**" field need to match one of the ICD10 codes (ICD-10 *-C26.x, C30.x-C34.x, C37.x-C41.x, C43.x, C45.x-C58.x, C60.x-C76.x, C81.x-C85.x, C88.x, C90.x-C97.x) by regular expression. And **"diag_date"** is less than the **"first_covid_diag_date"**. The full list of corresponding ICD10 codes of each of those variables can be found in the supplementary files.