American Journal of Epidemiology
Published by Oxford University Press on behalf of the Johns Hopkins Bloomberg School of Public Health 2021.

Vol. 190, No. 11
https://doi.org/10.1093/aje/kwab183
Advance Access publication:
June 24, 2021

Original Contribution

Pharmacoepidemiology, Machine Learning, and COVID-19: An Intent-to-Treat Analysis of Hydroxychloroquine, With or Without Azithromycin, and COVID-19 Outcomes Among Hospitalized US Veterans

Hanna Gerlovin*, Daniel C. Posner, Yuk-Lam Ho, Christopher T. Rentsch, Janet P. Tate, Joseph T. King, Jr., Katherine E. Kurgansky, Ioana Danciu, Lauren Costa, Franciel A. Linares, Ian D. Goethert, Daniel A. Jacobson, Matthew S. Freiberg, Edmon Begoli, Sumitra Muralidhar, Rachel B. Ramoni, Georgia Tourassi, J. Michael Gaziano, Amy C. Justice, David R. Gagnon, and Kelly Cho

* Correspondence to Dr. Hanna Gerlovin, VA Boston Healthcare System, 150 South Huntington Avenue (151-MAV), Boston, MA 02130 (e-mail: Hanna.Gerlovin@va.gov).

Initially submitted December 15, 2020; accepted for publication June 17, 2021.

Hydroxychloroquine (HCQ) was proposed as an early therapy for coronavirus disease 2019 (COVID-19) after in vitro studies indicated possible benefit. Previous in vivo observational studies have presented conflicting results, though recent randomized clinical trials have reported no benefit from HCQ among patients hospitalized with COVID-19. We examined the effects of HCQ alone and in combination with azithromycin in a hospitalized population of US veterans with COVID-19, using a propensity score–adjusted survival analysis with imputation of missing data. According to electronic health record data from the US Department of Veterans Affairs health care system, 64,055 US Veterans were tested for the virus that causes COVID-19 between March 1, 2020 and April 30, 2020. Of the 7,193 veterans who tested positive, 2,809 were hospitalized, and 657 individuals were prescribed HCQ within the first 48-hours of hospitalization for the treatment of COVID-19. There was no apparent benefit associated with HCQ receipt, alone or in combination with azithromycin, and there was an increased risk of intubation when HCQ was used in combination with azithromycin (hazard ratio = 1.55; 95% confidence interval: 1.07, 2.24). In conclusion, we assessed the effectiveness of HCQ with or without azithromycin in treatment of patients hospitalized with COVID-19, using a national sample of the US veteran population. Using rigorous study design and analytic methods to reduce confounding and bias, we found no evidence of a survival benefit from the administration of HCQ.

COVID-19; gradient boosting; hydroxychloroquine; pharmacoepidemiology; propensity score; survival analysis; treatment outcome

Abbreviations: CI, confidence interval; GBM, gradient boosting machine; HCQ, hydroxychloroquine; HR, hazard ratio; SARS-CoV-2, severe acute respiratory syndrome coronavirus 2; VA, Veterans Affairs.

In the swell of the coronavirus disease 2019 (COVID-19) pandemic, the world rushed to find therapies and prophylactic treatments, and hydroxychloroquine (HCQ) became an early front-runner (1, 2). HCQ is a common antimalarial and antirheumatologic drug with immunosuppressive functions. Results of early in vitro studies suggested HCQ might be repurposed to treat infections with a strong immune component (1, 3, 4), such as COVID-19. This was appealing, considering its low cost and widespread availability. The US Food and Drug Administration issued an emergency use authorization for HCQ on March 28, 2020 (5), prior to the completion of a randomized controlled trial, only to revoke it less than 3 months later, following concerns about HCQ-associated adverse events reported from observational studies (6, 7).

At about the same time as the US Food and Drug Administration's retraction, 3 randomized controlled trials—Outcomes Related to COVID-19 Treated With Hydroxychloroquine Among Inpatients With Symptomatic Disease

(ORCHID), Randomized Evaluation of COVID-19 Therapy (RECOVERY), and World Health Organization Public Health Emergency "Solidarity" Clinical Trial for COVID-19 Treatments (SOLIDARITY)—discontinued their HCQ arms because interim analyses showed no benefit in reducing deaths among inpatients with COVID-19 (8–10). The results of these trials recently were made public (11–13). Although randomized controlled trials are the gold standard for evaluating the effectiveness of a drug (14), none of the trials investigating HCQ treatment explored the combination with azithromycin in their study design. Azithromycin has also been given to patients with COVID-19, and in 1 small observational study, researchers hypothesized that the combination of the 2 drugs reduced viral load (1).

Results from observational studies of HCQ in treating COVID-19 have been inconsistent and subject to bias (15–19). Early studies claiming a benefit were from small samples with limited data and little control of potential confounders. Timing of treatment during hospitalization was often poorly defined, no studies appeared to control for secular trends in the timing of treatment, and several studies used data from HCQ use before the US Food and Drug Administration's initial emergency use authorization (20). Particularly, the study design and analytic techniques may not have been able to account for the various sources of potential and residual confounding (21–23).

In a recent meta-analysis of HCQ and death in patients hospitalized with COVID-19 (16), observational data were used in 25 of the 29 studies, and 10 of these peer-reviewed and preprint publications used some form of propensity adjustment. One main goal of propensity analysis is to balance confounding factors to emulate a randomized controlled trial setting (24). Recent studies on propensity scoring have shown that machine learning methods can achieve better balance than traditional regression methods in observational studies (25–29). Gradient-boosted modeling using decision trees allows for interactions among the variables used in propensity score calculation and makes no assumptions about the shape of the relationship between the confounder and treatment received (25).

In this article, we apply careful study design and statistical analytic approaches, using machine learning methods, to evaluate the effectiveness of HCQ, with or without azithromycin, in the treatment of COVID-19 in the US veteran population. We empirically assess the bias of the results by considering a priori–defined clinical confounders and a range of sensitivity analyses. Finally, we compare our analytic results with the existing literature on HCQ effectiveness for COVID-19 and draw conclusions about the implications for confounding in the context of an evolving pandemic.

METHODS

Veteran Affairs health care cohort

The US Department of Veterans Affairs (VA) is the largest single-payer US health care system, with 6 million veterans under care in the past 2 years. Structured electronic health records in the VA's Corporate Data Warehouse include all clinical encounters. Record domains include demographics, laboratory results, vital signs, health factors, pharmacy prescription fills, hospitalizations, and outpatient visits. A COVID-19–specific research database was constructed in the Knowledge, Discovery, and Innovation computing environment at the Department of Energy's Oak Ridge National Laboratory. The work for this analysis under the US Food and Drug Administration–led COVID-19 Insights Partnership projects was approved by both Department of Energy and VA institutional review boards and is a joint activity involving VA and Department of Energy investigators.

Study design

We designed our study cohort to mimic criteria that might be expected in a clinical trial setting (Figure 1). Key variables, index date, and exposure criteria follow a template developed for communicating reproducible, observational study designs in pharmacoepidemiology (30). Day 0, or the index date, was classified as the day of first hospitalization on the same day or after first positive diagnosis for severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), the virus that causes COVID-19.

Identification of COVID-19 cases

The VA COVID-19 research cohort includes individuals who were tested for SARS-CoV-2 inside or outside of a VA facility. We used the VA's National Surveillance Tool, the authoritative data source for defining positive and negative SARS-CoV-2 cases (31), to identify COVID-19 case patients as veterans who had a positive diagnosis.

Inclusion criteria

We restricted the sample to individuals hospitalized within the VA only, because of limited HCQ use and outcomes data for patients outside of VA hospitals. The base cohort included case patients with COVID-19 before April 30, 2020, after which HCQ use dramatically decreased (Figure 2). We included case patients for whom onset of infection was no later than hospital admission or June 1, 2020. We excluded patients who had received HCQ or azithromycin for non–COVID-19 illnesses (i.e., anyone using HCQ in the year prior to or using azithromycin within 14 days before the index date). We also excluded patients who were discharged, intubated, or died within 48 hours of admission, to avoid immortal time bias. We removed patients who received care at hospitals that were not prescribing HCQ, to ensure all individuals had a nonzero probability of receiving treatment.

Exposure assessment

Initiation was defined as the date of first inpatient prescription fill from index date until the end of follow-up. For an intention-to-treat analysis, we classified individuals into the following 4 groups on the basis of those initiating 1 or both of the drugs within the first 48 hours of hospitalization: HCQ plus azithromycin, HCQ alone, azithromycin alone, and neither drug. For example, any individuals who received only HCQ within the 48-hour window but later were

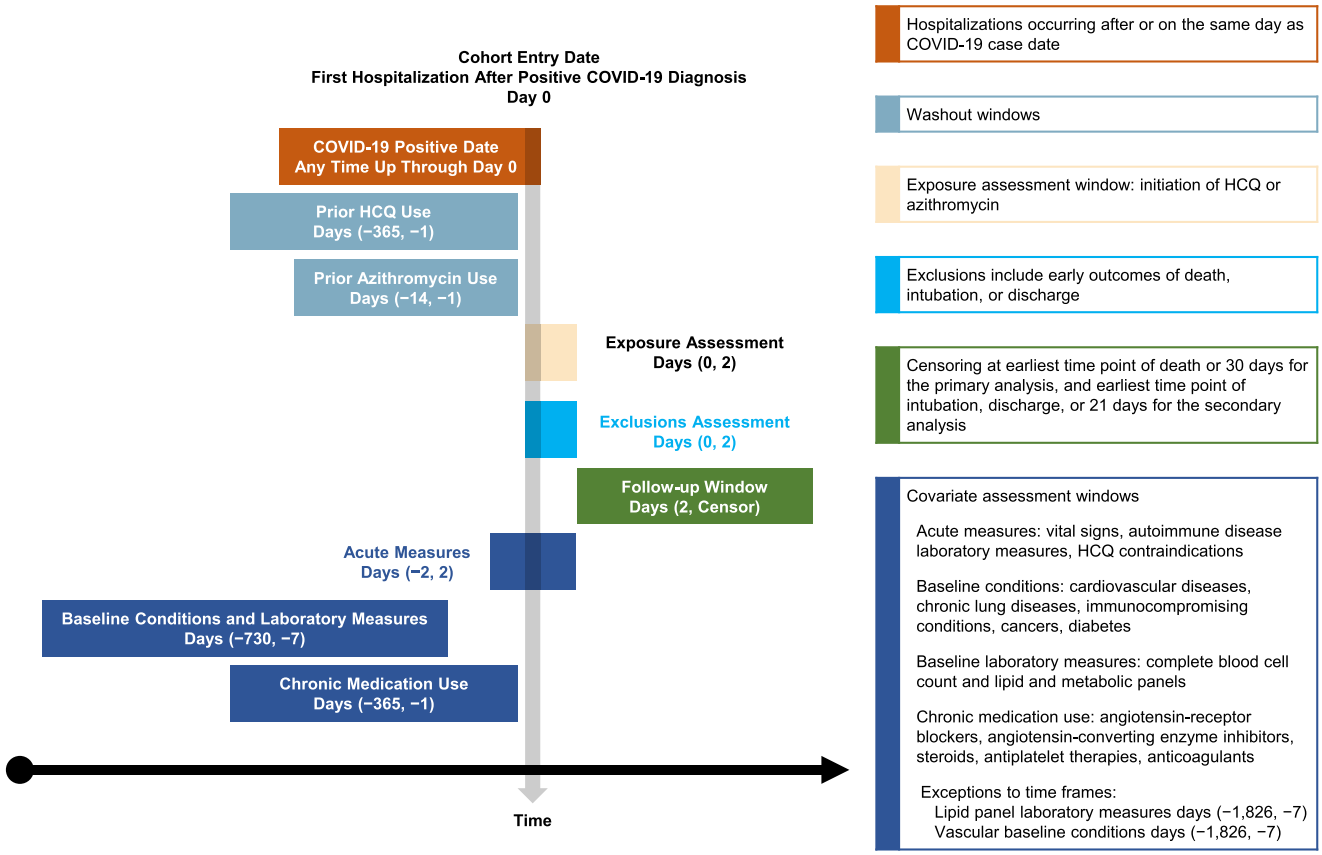

Figure 1. Study design diagram of a national sample of hospitalized US veterans with coronavirus disease 2019 (COVID-19) (n = 1,769) between March 1, 2020 and April 30, 2020. HCQ, hydroxychloroquine.

prescribed azithromycin after 48 hours were categorized in the HCQ-alone group.

Outcomes assessment

Outcome: death VA all-cause mortality information was based on the Beneficiary Identification Records Locator Subsystem, clinical records, and Social Security Death Index data (32). Time to death was measured from the index date and censored anyone who remained alive at 30 days.

Outcome: invasive ventilation/intubation We considered only invasive ventilation in our outcomes assessment, using diagnostic *International Classification of Diseases, Tenth Edition*, codes (0BH13EZ, 0BH17EZ, 0BH18EZ, 5A1935Z, 5A1945Z, 5A1955Z) and current procedural terminology code 31500. Among hospitalized patients with COVID-19, more than 95% of the intubations occurred within 21 days of admission, thus we analyzed the outcome using time to intubation during this 21-day period, with censoring at death or discharge.

Covariates and confounders

Potential confounders were assembled in clinically meaningful identification time frames (Figure 1). For uncommon laboratory tests that were measured acutely (e.g., lactate dehydrogenase, C-reactive protein, d-dimer, ferritin), we used evidence of measurement as the covariate of interest. Patient demographics (i.e., age, sex, region of the United States, urbanicity), height and weight, smoking status, alcohol use disorder, and evidence of recent long-term care were taken from data prior to the index date. Additional variables considered potential confounders of treatment and both primary outcomes were chronic medications, concurrent inpatient treatments (for COVID-19 or HCQ contraindications), chronic conditions (based on diagnostic codes and including a frailty score (33)), and acute laboratory results and vital signs (those related to acute illness). All potential confounders were included in the propensity model. Complete descriptions of diagnostic and medication codes are provided in Web Tables 1 and 2 and Web Appendix 1 (available at https://doi.org/10.1093/aje/kwab183).

Statistical analyses

All analyses were performed using R software (R Foundation for Statistical Computing, Vienna, Austria) (34) and publicly available packages.

Missing data Missing covariate information was imputed using the multiple imputation from the chained equations

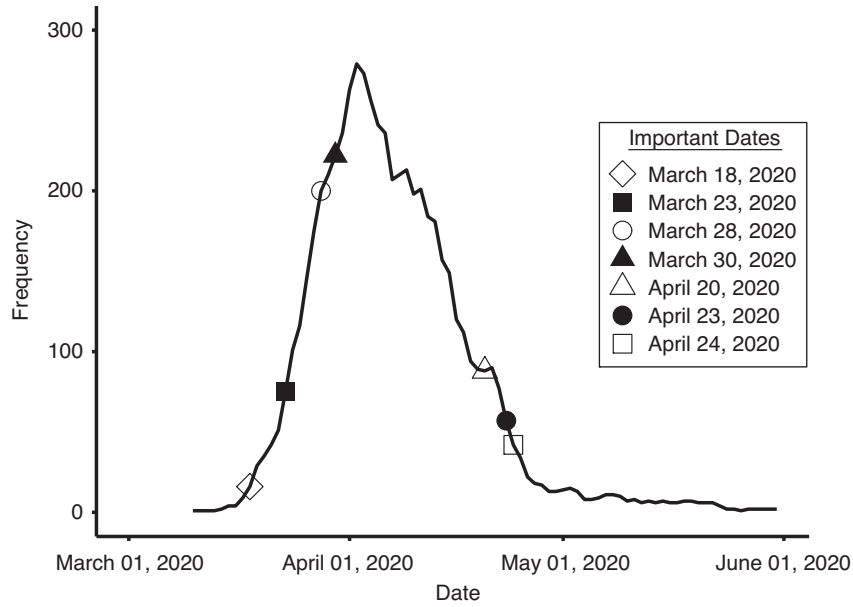

Figure 2. New-user hydroxychloroquine (HCQ) prescriptions over time in a national sample of hospitalized US veterans with coronavirus disease 2019 (COVID-19) (n = 1,769) between March 1, 2020, and April 30, 2020. Important dates, as annotated on the plot, are: A) March 18, 2020: pharmacy benefits management (PBM) literature summary of off-label COVID-19 therapeutic options posted; B) March 23, 2020: PBM HCQ prioritization criteria posted; C) March 28, 2020: US Food and Drug Administration (FDA) approves HCQ under emergency use authorization; D) March 30, 2020: PBM HCQ safety document posted; E) April 20, 2020: PBM HCQ prioritization criteria archived (lack of effectiveness data); F) April 23, 2020: PBM HCQ safety document updated; G) April 24, 2020: FDA drug safety communication.

"mice" package in R (35, 36). Ten imputed data sets were generated, analyzed separately, and the final results were subsequently combined using Rubin's rules to determine final effect sizes and confidence intervals (CIs) (37).

Propensity score calculation Propensity scores for each treatment were estimated from a gradient boosting machine (GBM) (38), an ensemble of models that takes baseline measures and characteristics as inputs and outputs the patient's predicted probability (or propensity score) for receiving each treatment (both HCQ and azithromycin, HCQ alone, azithromycin alone, or neither). We used decision trees as base learners for GBM, using the "gbm" and "WeightIt" R packages to fit our models (39, 40). The hyperparameters were set as an interaction depth of 4, a maximum of 5,000 trees, and shrinkage of 0.1. We optimized the maximum of standardized mean differences between potential confounders across the treatment arms.

For each patient, the propensity score was converted to a stabilized inverse probability of treatment weight. We evaluated the propensity scores using the "cobalt" package in R to look at the distributions of average standardized mean differences between each pair of treatments (41). The relative influence (42) was calculated as the normalized amount of change in the balance metric for each variable when it was used to split a node.

Outcome models The stabilized inverse probability of treatment weights from the propensity modeling steps were included as subject-level weights in Cox proportional hazards multivariable models for estimating treatment effects on death and intubation using the "survival" package (43) in R. An α level of 0.05 was used.

Sensitivity analyses

We assessed design assumptions and data restrictions with a series of sensitivity analyses to address questions regarding timing, analytic design, and methods. To consider whether timing of treatment initiation made a difference in survival, we considered a shorter, 24-hour exposure window, with corresponding adjustments in exclusions and outcomes. We explored the effect of the secular prescribing trend(s) by limiting analyses to time windows framed by regulatory guidelines and patterns of use within the VA. The final set of sensitivity analyses focused on the statistical and machine-learning methods and assumptions. We also considered a set of doubly robust models, with select confounders included in both the propensity and outcome models (44, 45). Complete details about cohort restrictions and sensitivity analyses performed are provided in Web Table 3 and Web Appendix 2.

RESULTS

Characteristics of users

As of April 30, 2020, there were 7,193 SARS-CoV-2–positive cases out of 64,055 individuals tested overall (46),

Table 1. Select Baseline Characteristics of a National Sample of Hospitalized US Veterans With Coronavirus Disease 2019 Between March 1, 2020, and April 30, 2020 (n = 1,769), by 48-Hour Treatment Exposure

Baseline Characteristic	Neither Drug (n = 770)			Azithromycin Alone (n = 342)			HCQ Alone (n = 228)			HCQ + Azithromycin (n = 429)			*P* Value
	No.	%	Median (IQR)	No.	%	Median (IQR)	No.	%	Median (IQR)	No.	%	Median (IQR)	
Demographics and lifestyle													
Age, years[a]	71.47 (13.12)			67.64 (13.55)			70.24 (12.80)			67.81 (13.22)			<0.001
Male sex	736	95.6		322	94.2		219	96.1		413	96.3		0.527
Race/ethnicity													<0.001
Non-Hispanic White	321	41.7		107	31.3		65	28.5		139	32.4		
Non-Hispanic Black	341	44.3		190	55.6		138	60.5		225	52.4		
Hispanic	77	10.0		21	6.1		15	6.6		42	9.8		
Other	31	4.0		24	7.0		10	4.4		23	5.4		
Days to admission[a,b]	2.6 (8.1)			0.8 (4.4)			1.0 (2.3)			0.6 (2.3)			<0.001
Week of admission[a]	15.3 (2.1)			14.0 (1.9)			14.5 (1.4)			14.0 (1.2)			<0.001
Total station size[a]	61,485 (31,509)			64,346 (3,1385)			49,902 (20,603)			57,189 (28,718)			<0.001
Urban	707	91.8		310	90.6		217	95.2		408	95.1		0.031
Coming from LTC facility[c]	130	16.9		18	5.3		23	10.1		20	4.7		<0.001
In ICU at 48 hours	134	17.4		77	22.5		41	18.0		124	28.9		<0.001
Smoking status[d,e]													0.042
Never	167	30.1		90	35.3		50	35.2		111	39.5		
Current	239	43.1		85	33.3		54	38.0		104	37.0		
Former	148	26.7		80	31.4		38	26.8		66	23.5		
Prior laboratory measures[f]													
Hemoglobin, g/dL			12.8 (11.1, 14.1)			13.5 (12.0, 14.6)			13.2 (11.4, 14.3)			13.4 (12.2, 14.6)	<0.001
HbA1C,[g] %			6.2 (5.6, 7.3)			6.1 (5.6, 7.1)			6.2 (5.7, 7.4)			6.2 (5.7, 7.1)	0.733
LDL-C, mg/dL			78.0 (58.0, 102.5)			90.3 (69.0, 121.4)			81.0 (59.0, 108.6)			87.0 (64.0, 110.0)	<0.001
Lymphocyte count, K/mm^3			1.6 (1.3, 2.2)			1.7 (1.4, 2.2)			1.7 (1.4, 2.1)			1.7 (1.4, 2.2)	0.617

Table continues

Table 1. Continued

Baseline Characteristic	Neither Drug (*n* = 770)			Azithromycin Alone (*n* = 342)			HCQ Alone (*n* = 228)			HCQ + Azithromycin (*n* = 429)			*P* Value
	No.	%	Median (IQR)	No.	%	Median (IQR)	No.	%	Median (IQR)	No.	%	Median (IQR)	
Acute laboratory measures[h]													
eGFR, mL/min			63.1 (38.5, 85.1)			61.6 (39.7, 83.8)			55.2 (32.2, 81.8)			62.7 (41.8, 82.6)	0.156
WBC count, K/mm^3			6.0 (4.6, 7.9)			6.1 (4.9, 8.4)			6.1 (4.5, 8.1)			6.3 (4.9, 8.2)	0.383
ALT, U/L			24.0 (16.0, 39.0)			28.0 (18.0, 44.0)			30.0 (20.0, 44.0)			33.0 (22.5, 49.0)	<0.001
C-reactive protein,[i] mg/dL			19.0 (6.3, 71.7)			11.4 (5.5, 44.2)			21.6 (8.3, 75.3)			15.4 (8.5, 42.4)	0.009
Missing data	251	32.6		108	31.6		61	26.8		81	18.9		
D-dimer,[i] µg/mL			13 (3, 2,199)			6 (3, 3,600)			1,004 (4, 2,195)			1,411 (3, 3,071)	0.555
Missing data	664	86.2		296	86.5		190	83.3		348	81.1		
Acute vital signs[j]													
Body mass index[k,l] ≥30	282	37.0		152	44.6		107	46.9		209	48.7		<0.001
Oxygen saturation[m] ≤93%	131	18.1		67	20.8		52	24.0		116	28.2		0.001
Respiratory rate[k] > 22/min	82	11.0		45	13.2		37	16.6		76	17.9		0.006
Temperature[m] ≥ 38 °C (100.4 °F)	117	15.6		64	18.8		43	19.2		97	22.8		0.023
Prior medications[n]													
Any ACE or ARB[o]	309	40.1		137	40.1		88	38.6		182	42.4		0.784
Any anticoagulant	107	13.9		38	11.1		26	11.4		43	10.0		0.213
In-hospital medications[p]													
Dexamethasone	2	0.3		0	0.0		1	0.4		7	1.6		0.007
Methylprednisolone	10	1.3		5	1.5		9	3.9		19	4.4		0.002
Remdesivir	12	1.6		1	0.3		1	0.4		0	0.0		0.014
Comorbidity scores[a,q]													
Charlson Comorbidity Index[g]	4.84 (3.42)			4.13 (2.90)			4.61 (3.24)			4.10 (2.84)			0.005
Frailty Index	0.31 (0.17)			0.24 (0.16)			0.27 (0.17)			0.24 (0.15)			<0.001
5-Year cardiovascular diseases													
Coronary heart disease	301	39.1		107	31.3		80	35.1		119	27.7		0.001
Cerebrovascular accident	212	27.5		78	22.8		58	25.4		74	17.2		0.001
Peripheral vascular disease	212	27.5		64	18.7		59	25.9		100	23.3		0.014

Table continues

Table 1. Continued

Baseline Characteristic	Neither Drug (*n* = 770)			Azithromycin Alone (*n* = 342)			HCQ Alone (*n* = 228)			HCQ + Azithromycin (*n* = 429)			*P* Value
	No.	%	Median (IQR)	No.	%	Median (IQR)	No.	%	Median (IQR)	No.	%	Median (IQR)	
Prior conditions[r]													
Diabetes	395	51.3		151	44.2		123	53.9		205	47.8		0.065
Hypertension	616	80.0		246	71.9		179	78.5		312	72.7		0.004
Any lung disease[s]	269	34.9		104	30.4		68	29.8		124	28.9		0.12
Dementia	176	22.9		40	11.7		30	13.2		44	10.3		<0.001

Abbreviations: ACE, angiotensin-converting enzyme inhibitor; ALT, alanine aminotransferase; ARB, angiotensin-receptor blocker; eGFR, estimated glomerular filtration rate; HbA1C, glycosylated hemoglobin; HCQ, hydroxychloroquine; ICU, intensive care unit; LDL-C, low-density lipoprotein cholesterol; LTC, long-term care; PS, propensity score; WBC, white blood cell.

[a] Values are expressed as mean (standard deviation).

[b] Days between a severe acute respiratory syndrome coronavirus 2–positive laboratory result and hospital admission.

[c] Any prior admissions to or from a long-term care, skilled nursing, or community housing facility up to 6 months before hospitalization.

[d] Smoking taken as mode from health factors data. Number missing omitted from table.

[e] Proportions may not sum to 100 due to rounding

[f] Prior laboratory measures timing: 2 years up to 7 days before hospitalization: HbA1C, hemoglobin, and lymphocytes; 5 years up to 7 days before hospitalization: LDL-C.

[g] Variable not used in propensity score model(s).

[h] Acute laboratory measures timing: 7 days before hospitalization up to date of first medication or 48 hours, whichever came first: ALT, eGFR, and WBC count; any measure 48 hours before up through 48 hours after hospital admission: C-reactive protein, D-dimer.

[i] Rare laboratory measures fed into PS model using indicator of collection (C-reactive protein, lactate dehydrogenase, ferritin, and d-dimer)

[j] Vital signs timing was within 2 days of index date, except for height and weight, which were from the closest measure before index.

[k] Variable included in PS model as a continuous measure only.

[l] Weight (kg)/height $(m)^2$.

[m] Variable included as both indicator and continuous measure in PS model.

[n] Prior medications: prescribed in the year before index through outpatient only.

[o] Indicators for any ACE and any ARB were included separately in the PS model.

[p] In-hospital medications: received at any point in first 48 hours of hospitalization, based on inpatient pharmacy data.

[q] Comorbidity scores timing: Charlson Comorbidity Index score during the 2 years before hospitalization, Frailty Index in 3 years before hospitalization.

[r] Prior conditions timing: any 1 inpatient code or 2 outpatient codes in the 2 years up to 7 days before hospitalization.

[s] Asthma, bronchitis, and chronic obstructive pulmonary disease were entered into the PS model as separate indicators.

yielding an analytic cohort of 1,769 individuals (Figure 3). The mean number of days between a positive laboratory result and hospitalization was 2.1 (Web Figure 1). In the first 48 hours of hospitalization, 429 individuals (24%) initiated treatment with HCQ and azithromycin, 228 (13%) received HCQ alone, and 342 (19%) received azithromycin alone; 770 patients (44%) were not prescribed either of these 2 treatment strategies (Table 1, Web Table 4).

Those who initiated azithromycin alone or in combination with HCQ in the first 48 hours of hospitalization were younger (mean ages, 67.6 and 67.8 years, respectively) than those initiating HCQ alone (mean age, 70.2 years) or neither treatment (mean age, 71.5 years) in the same exposure time frame. Non-Hispanic Black patients were more likely to receive at least 1 treatment than patients of other race/ethnicity groups, and those in urban settings were more likely to be prescribed some form of HCQ. Patients coming from long-term care or nursing facilities were less likely to have either treatment initiated within the first 48 hours of admission. Acute laboratory measurements, such as lactate dehydrogenase and C-reactive protein levels, were more commonly available for patients those initiating both treatments.

Propensity model

Before weighting, the exposure groups differed with respect to multiple covariates (Table 1). Overall, the GBM was able to balance a large majority of the variables in the primary analysis model. Complete love and balance plots can be found in Web Figures 2–15. The week-of-admission variable did not achieve the recommended threshold of 0.1 (47) or even 0.2 for the average standardized mean difference comparing those initiating both treatments with any of the other 3 groups. Similarly, the average standardized mean difference for week of admission comparing HCQ alone with the neither-treatment or azithromycin-alone groups was approximately 0.2. Figure 4 displays the average relative importance or influence of a given predictor in the primary propensity model. Notably, total station size and week of admission were the most important factors across all imputations and sensitivity analyses (Web Figures 16–18).

Primary analysis

Of the 429 individuals initiating both HCQ and azithromycin in the first 48 hours after VA hospital admission, 90 (21%) died within 30 days after admission and 64 (15%) were intubated within 21 days of admission (Table 2, Web Table 5). After weighting, patients for whom both treatments were initiated had a 22% increased hazard of death (hazard ratio (HR) = 1.22, 95% CI: 0.91, 1.63) and 55% increased hazard of intubation (HR = 1.55, 95% CI: 1.07, 2.24) compared with patients receiving neither treatment within the first 48 hours after hospitalization.

Comparing those exposed to HCQ alone with patients who received neither treatment in the 48 hours following admission, there were nonstatistically significant increased risks of both death within 30 days of index (HR = 1.21, 95% CI: 0.82, 1.76) and intubation within 21 days of index (HR = 1.33, 95% CI: 0.82, 2.15). Meanwhile, patients initiating azithromycin alone in the first 48 hours had similar hazards for death (HR = 0.90, 95% CI: 0.64, 1.27) and intubation (HR = 1.03, 95% CI: 0.66, 1.61) compared with patients receiving neither treatment. None of these analyses indicated a benefit of HCQ or azithromycin.

Sensitivity analyses

There were few measurable changes in the effect estimates and CIs of the 2 comparisons (i.e., both vs. neither; HCQ alone vs. neither) for many of the sensitivity analyses. Figure 5 summarizes the average treatment effect HR (95% CI) for those initiating any combination of HCQ compared with those receiving neither treatment in the 48 hours after admission. Complete results, including event counts and number exposed, from all sensitivity analyses can be found in Web Tables 6–7.

Censoring at change in treatment (adding either azithromycin or HCQ after 48 hours after hospitalization) produced substantially different results for death (for HCQ vs. neither treatment, HR = 1.42, 95% CI: 0.92, 2.18; for both treatments vs. neither treatment, HR = 1.63, 95% CI: 1.18,2.25; Figure 5A). This corresponded to 75 fewer "cases," mostly from the neither-treatment group. A similar pattern of inflated HRs and fewer cases can be seen for the intubation outcome (Figure 5B).

Dropping the index dates that occurred before the Pharmacy Benefits Management's guidelines for HCQ emergency-use authorization posted on March 30, 2020, left slightly more than two-thirds of the total sample (n = 1,218). This change did not affect the intent-to-treat HR for HCQ alone versus neither drug in terms of death, but it did shift the final estimate for both treatments versus neither treatment away from the null (HR = 1.41, 95% CI: 0.98, 2.03), indicating greater harm. For the intubation outcome, the HR shifted again; however, these data may not be interpretable, due to the small number of cases.

DISCUSSION

Key findings

We found no benefit in COVID-19 death and intubation from treatment with HCQ alone or in combination with azithromycin when administered shortly after hospital admission. The direction of the effect was consistent across all models and comparable to findings of recent studies of HCQ for the treatment of SARS-CoV-2 infection in the inpatient hospital setting (11, 12, 48–50).

Research in context

A previous analysis of HCQ effectiveness among veterans demonstrated no evidence of benefit for those prescribed HCQ with or without azithromycin, with indication of harm from HCQ alone (20). The sample size was small (n = 807), with a restricted follow-up window for certain individuals.

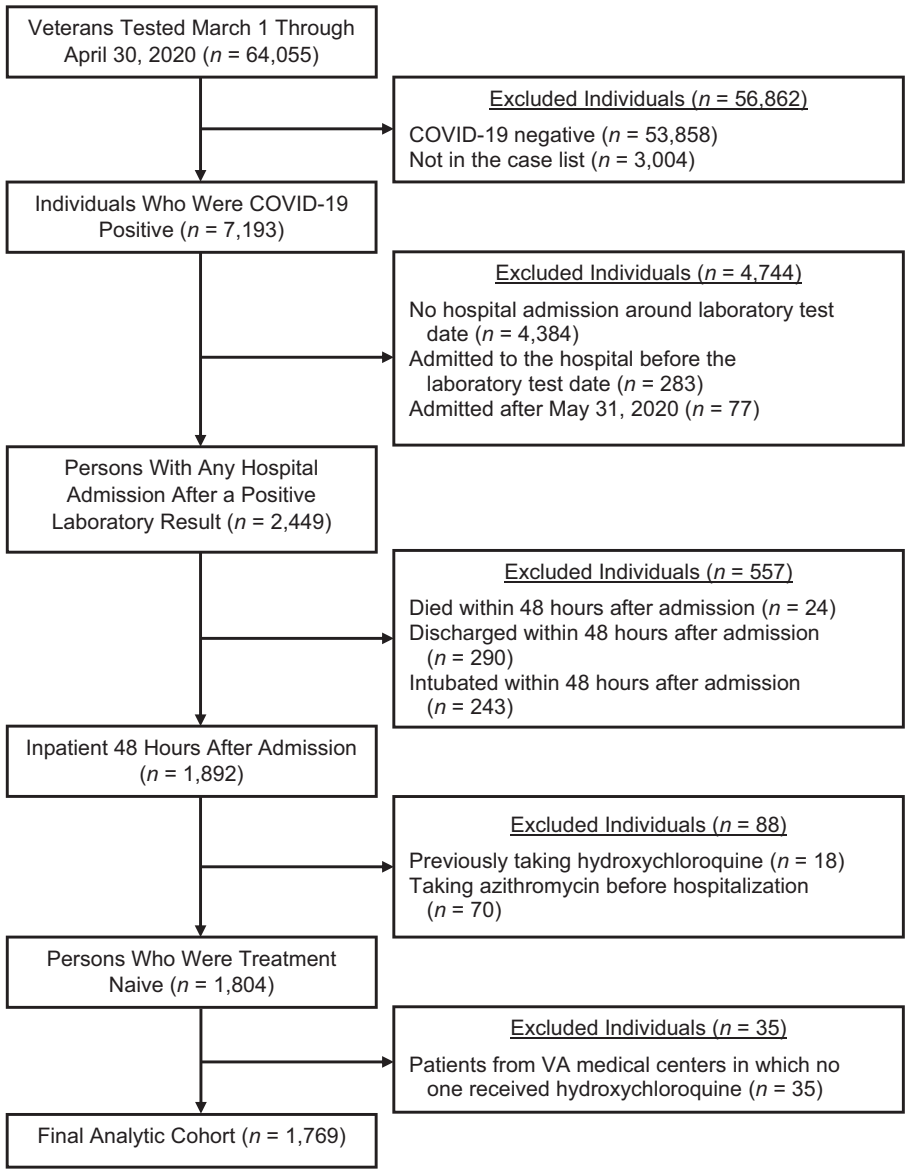

Figure 3. Study inclusion and exclusion criteria for the primary analysis of a national sample of hospitalized US veterans with coronavirus disease 2019 (n = 1,769) between March 1, 2020, and April 30, 2020.

In contrast, authors of a study from the Henry Ford Health System estimated that any form of HCQ led to significant reductions of in-hospital deaths (for HCQ vs. neither treatment, HR = 0.66; for both treatments vs. neither treatment, HR = 0.71) (51). The study differed in population demographics, size, and by the use of a multivariate modeling approach that included a limited number of confounders in the models. This study was criticized for insufficiently controlling for confounding by indication (i.e., sicker patients were less likely to receive HCQ) (52). In addition, the Henry Ford study did not account for secular trends, which we demonstrate here are an important factor to include in analyses.

Methodological differences

Using a sample twice the size of the prior VA study (n = 1,769 vs. 807), we found similar average treatment effects of HCQ with or without azithromycin compared with neither treatment. The differences that exist in our findings can likely be explained by our use of an adjudicated, algorithm-based case definition (the VA's National Surveillance Tool) that captures laboratory-identified cases as well as those not in the VA system. However, given the 95% agreement between COVID-19 case definitions based on VA laboratory test only and National Surveillance Tool positive definitions, it is also possible that the search terms used in

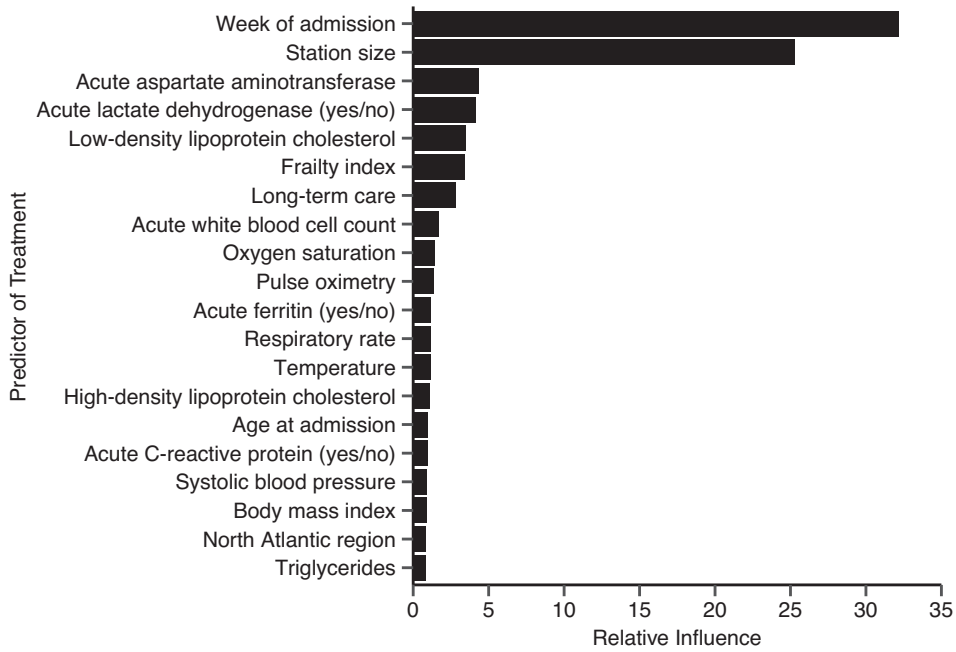

Figure 4. Relative influence plot of variables included in propensity model (primary analysis).

the prior study did not capture all SARS-CoV-2 positive cases. Magagnoli et al. (20) additionally restricted follow-up through April 29, 2020, meaning that the outcomes of patients hospitalized toward the end of April would not have had enough time to be observed. In a sensitivity analysis using a similar enrollment restriction (i.e., hospitalized on or before April 30), but with adequate follow-up time for all individuals, we saw no change in our results or conclusions.

Strengths

Relative to other observational cohorts in the United States, the VA has more longitudinal data and limited loss to follow-up. These factors allow for a more complete assessment of patients' comorbidities and outcomes.

Chronological bias (53) is a challenging feature of research related to HCQ. Chronological bias can be introduced by variable prescribing patterns for the drug (54) in conjunction with the geographic spread of the disease (55) and a constantly evolving knowledge base about the disease and treatments for it (52). We explored multiple sensitivity analyses that demonstrated consistent results when timing of hospitalization and hospital size and capacity were accounted for in the models.

We considered the importance of timing of treatment with a sensitivity analysis setting the exposure window to 24 hours, as in other studies (48, 56). This resulted in similar

Table 2. Effect of HCQ With or Without Azithromycin on Death Over 30 Days and Intubation Over 21 Days in a National Sample of Hospitalized US Veterans With Coronavirus Disease 2019 (n = 1,769) Between March 1, 2020, and April 30, 2020

Drug	No. Exposed	Death				Intubation			
		No. of Cases	No. of Person-Days	HR	95% CI	No. of Cases	No. of Person-Days	HR	95% CI
Neither drug	770	141	20,376	1.00	Referent	69	7,241	1.00	Referent
Azithromycin alone	342	56	9,174	0.90	0.64, 1.27	39	2,625	1.03	0.66, 1.61
HCQ alone	228	49	5,853	1.21	0.82, 1.76	32	1,897	1.33	0.82, 2.15
HCQ and azithromycin	429	90	11,153	1.22	0.91, 1.63	64	3,370	1.55	1.07, 2.24

Abbreviations: CI, confidence interval; HCQ, hydroxychloroquine; HR, hazard ratio.

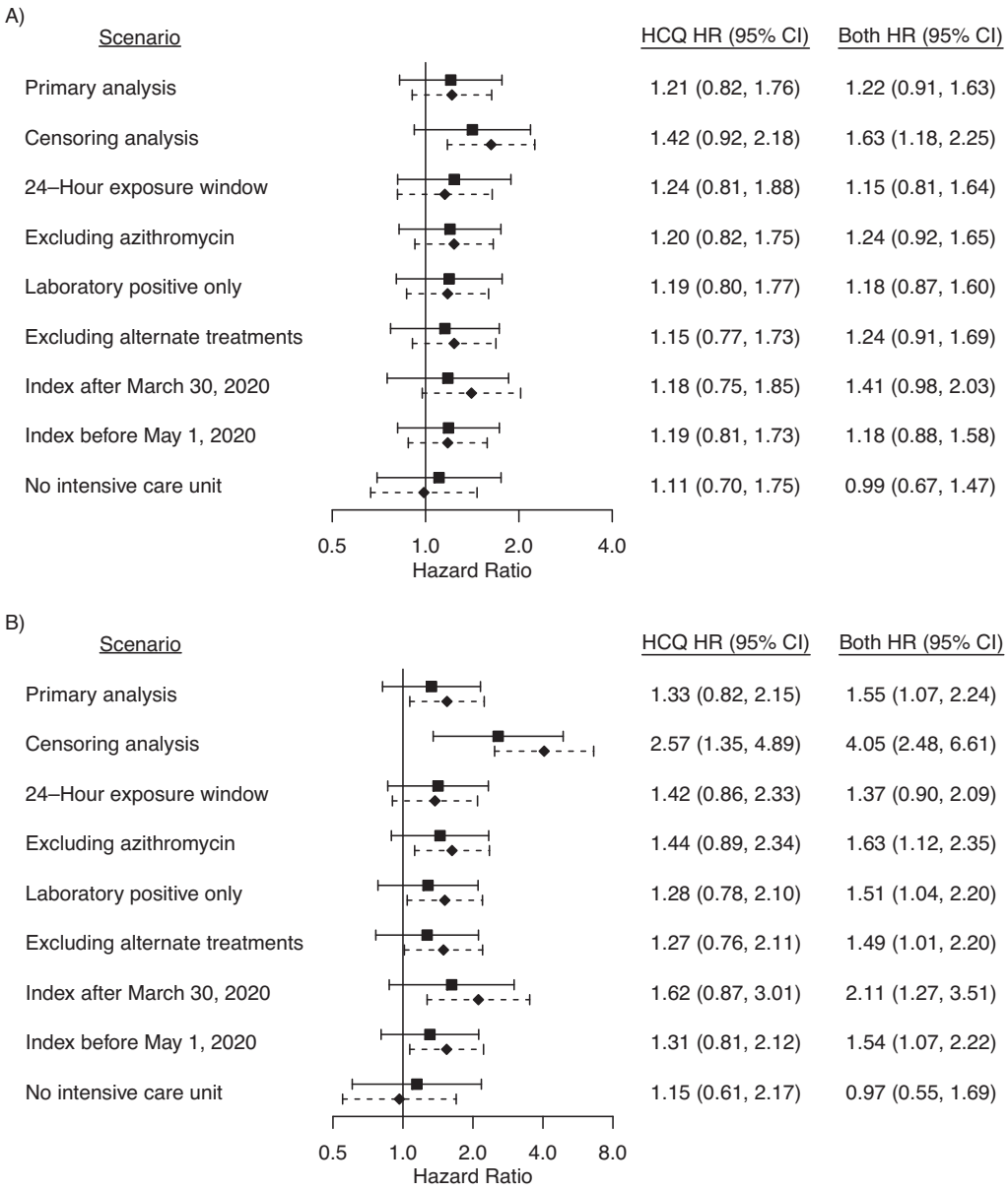

Figure 5. Forest plots comparing A) death and B) intubation hazard ratios (HRs) across sensitivity analyses of a national sample of hospitalized US veterans with coronavirus disease 2019 (N = 1,769) between March 1, 2020, and April 30, 2020. Squares with solid lines and diamonds with dashed lines represent the HRs and 95% confidence intervals (Cis) corresponding to the average treatment effects of hydroxychloroquine (HCQ) and HCQ with azithromycin (both) compared with neither treatment, respectively. Sensitivity analyses were conducted to determine differences from primary analysis. The terms under "Scenario" in panels A and B refer to the following: censoring, when HCQ or azithromycin was added to the patient's treatment after the 48-hour exposure window; 24-hour exposure window, using a 24-hour window (from hospitalization) for exposure definition and corresponding exclusions; excluding azithromycin, removing the azithromycin-alone group before propensity modeling; laboratory positive only, restricting the cohort to only those with a positive laboratory test result in the Veterans Affairs laboratory records; excluding alternate treatments, removing individuals taking dexamethasone, lopinavir-ritonavir, remdesivir, or tocilizumab in the 48-hour exposure window; index after March 30, 2020, restricting index dates to after March 30, 2020 (or after the emergency use authorization was issued for HCQ use); index before May 1, 2020, restricting index dates to April 30, 2020, or earlier; no intensive care unit, removing any individuals admitted to an intensive care unit within the 48-hour exposure assessment window.

estimates to the primary analysis for death (for HCQ vs. neither treatment, HR = 1.24; for both treatments vs. neither treatment, HR = 1.15). Given that 11% of the sample added 1 or more of the treatments in the 24- to 48-hour window, our use of the 48-hour window may be preferable because it more effectively avoids misclassification. We excluded individuals who died, were intubated, or discharged within the 48 hours, because the patients would not have had enough

time to experience benefit or harm from HCQ. Although this ensured the circumvention of immortal time bias when defining the group with neither treatment, we recognize this approach may present additional limitations.

Limitations

Our results may not be generalizable to those intubated before receiving treatment or to those with less severe illness who are discharged almost immediately. Compared with the overall US population, VA users are older, mostly male, with more comorbidities and lower socioeconomic status (55). Our results may differ from studies of younger and healthier populations with a higher proportion of women. However, older, male, and sicker individuals are at higher risk for severe COVID-19, which warranted the study of this drug early on, despite historical data indicating these might also be the individuals most at risk for adverse events from HCQ (3).

Propensity weighting did not completely eliminate covariate imbalance across the treatment groups. To address this limitation, we performed a series of doubly robust models (44, 45) (as described in Web Appendix 1) in which covariates were included in both the propensity and outcome models. The estimated HRs and CIs were similar to those reported from the primary analysis, further confirming the lack of benefit from HCQ.

Our analysis did not account for any changes in HCQ or azithromycin status after the 48-hour exposure assessment, such as new prescriptions or treatment discontinuation. We attempted to address the change in treatment after 48 hours through a sensitivity-analysis censoring at the addition of another treatment. This per-protocol, on-treatment analysis has been shown to confer bias in the clinical trial setting (57) and thus is not preferred over the intention-to-treat method used. In fact, we observed this bias in the shifted HRs and CIs that made HCQ (with and without azithromycin) appear harmful compared with receiving neither treatment.

After 48 hours from the index date, approximately 25% of the patients in the combination-treatment group were in the intensive care unit, compared with 5% of patients in the neither-treatment group, 19% in the azithromycin alone group, and 13% for those receiving HCQ alone. We did not look at this particular outcome or adjust for it as a confounder in the propensity models. However, in a sensitivity analysis from which these individuals were removed, the HRs for both death and intubation of the combined treatment group, relative to neither treatment, shifted completely to the null, indicating that HCQ may have been seen as a "rescue" therapy for patients in the intensive care unit. Of note, even with this restriction, we found no evidence of benefit.

Despite our array of sensitivity analyses, we acknowledge there is still a possibility of some unmeasured and residual confounding we were unable to account for. However, the GBM approach allowed us to control for many variables, and any remaining unmeasured confounders would likely require strong associations with both the treatment assignment and outcomes, to explain away the null relationship observed in the data.

Implications

In the early months of the pandemic, there was much uncertainty about risk factors of COVID-19 and subsequent deaths, which translated to inconsistent results and conclusions from studies with moderate to severe levels of bias (16). With our best attempts to adjust for possible confounding, we found confirmatory evidence for an increased risk of intubation for those who were treated with the combination of HCQ and azithromycin for COVID-19 in a hospital setting. We found no inpatient survival benefit to the administration HCQ, with or without concomitant azithromycin.

Our study reflects the challenges of modeling effectiveness during the start of a pandemic and demonstrates that consistent data over time are critical for disentangling the effects of confounding by indication. Although we cannot account for compassionate use of HCQ, we do show that sensitivity analyses in both study design and modeling can allow researchers to account for many potential confounders using electronic health record data when a priori relationships are not well established.

ACKNOWLEDGMENTS

Author affiliations: Massachusetts Veterans Epidemiology, Research and Information Center, VA Boston Healthcare System, US Department of Veterans Affairs, Boston, Massachusetts, United States (Hanna Gerlovin, Daniel C. Posner, Yuk-Lam Ho, Katherine E. Kurgansky, Lauren Costa, J. Michael Gaziano, David R. Gagnon, Kelly Cho); VA Connecticut Healthcare System, US Department of Veterans Affairs, West Haven, Connecticut, United States (Christopher T. Rentsch, Janet P. Tate, Joseph T. King, Jr., Amy C. Justice); Faculty of Epidemiology and Population Health, London School of Hygiene and Tropical Medicine, London, United Kingdom (Christopher T. Rentsch); Department of Internal Medicine, Yale School of Medicine, New Haven, Connecticut, United States (Janet P. Tate, Amy C. Justice); Department of Neurosurgery, Yale School of Medicine, New Haven, Connecticut, United States (Joseph T. King, Jr.); Oak Ridge National Laboratory, Oak Ridge, Tennessee, USA (Ioana Danciu, Franciel A. Linares, Ian Goethert, Daniel A. Jacobson, Edmon Begoli, Georgia Tourassi); The Bredesen Center for Interdisciplinary Research and Graduate Education, University of Tennessee Knoxville, Knoxville, Tennessee, United States (Daniel A. Jacobson); Department of Psychology, University of Tennessee Knoxville, Knoxville, Tennessee, United States (Daniel A. Jacobson); Geriatric Research Education and Clinical Center, Tennessee Valley Healthcare System, US Department of Veterans Affairs, Nashville, Tennessee, United States (Matthew S. Freiberg); Department of Medicine, Vanderbilt University Medical Center, Nashville, Tennessee, United States (Matthew S. Freiberg); Office of Research and Development, US Department of Veterans Affairs, Washington, DC, United States (Sumitra Muralidhar,

Rachel B. Ramoni); Division of Aging, Brigham and Women's Hospital, Boston, Massachusetts, United States (J. Michael Gaziano, Kelly Cho); Department of Medicine, Harvard Medical School, Boston, Massachusetts, United States (J. Michael Gaziano, Kelly Cho); Yale School of Public Health, New Haven, Connecticut, United States (Amy C. Justice); and Department of Biostatistics, Boston University School of Public Health, Boston, Massachusetts, United States (David R. Gagnon).

The project was supported by the US Department of Veterans Affairs, Office of Research and Development, Million Veteran Program Core (grant MVP000) and the National Institute on Alcohol Abuse and Alcoholism (grants U01-AA026224, U24-AA020794, U01-AA020790, U10-AA013566). We used resources of the Knowledge Discovery Infrastructure at the Oak Ridge National Laboratory, which is supported by the Office of Science of the US Department of Energy (DOE) under contract DE-AC05-00OR22725, and the Department of Veterans Affairs Office of Information Technology Inter-Agency Agreement with the DOE (VA118-16-M-1062). Part of this work was also supported by the Laboratory Directed Research and Development (LDRD) Program of Oak Ridge National Laboratory, managed by UT-Battelle, LLC, for the DOE (LOIS: 10074). Additional support was provided by LDRD Program for Small Seed Funding of Oak Ridge National Laboratory, also managed by UT-Battelle, LLC, for the DOE (LOIS: 10137).

The authors acknowledge the support of the larger COVID-19 Insights Partnership, team members Ryan Tipton, Hope Cook, Joe C Lake, Everett Rush III, Jeremy Cohen, Ozgur Ozmen, Melissa Skanderson, Brian Ferolito, and Drs Walid Gellad, Fran Cunningham, Rachel Ward, JP Casas, and Stacey Whitbourne. Most importantly, the authors thank and acknowledge the veterans who chose to get their care at the VA.

The data used in this paper cannot be shared outside the VA, except as authorized under VA policy with respect to the Freedom of Information Act. Formal mechanisms of request for anonymized data, analytic codebooks, and programs must be approved by the Veterans Health Affairs Office of Research Oversight.

The views and opinions expressed in this manuscript are those of the authors and do not necessarily represent those of the Department of Veterans Affairs, Department of Energy, or the US Government. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Conflict of interest: none declared.

REFERENCES

1. Andreani J, Le Bideau M, Duflot I, et al. In vitro testing of combined hydroxychloroquine and azithromycin on SARS-CoV-2 shows synergistic effect. *Microb Pathog*. 2020;145:104228.
2. Kupferschmidt K, Cohen J. Race to find COVID-19 treatments accelerates. *Science*. 2020;367(6485):1412–1413.
3. Oscanoa TJ, Romero-Ortuno R, Carvajal A, et al. A pharmacological perspective of chloroquine in SARS-CoV-2 infection: an old drug for the fight against a new coronavirus? *Int J Antimicrob Agents*. 2020;56(3):106078.
4. Tarek M, Savarino A. Pharmacokinetic basis of the hydroxychloroquine response in COVID-19: implications for therapy and prevention. *Eur J Drug Metab Pharmacokinet*. 2020;45(6):715–723.
5. Bright R. Re: Request for Emergency Use Authorization For Use of Chloroquine Phosphate or Hydroxychloroquine Sulfate Supplied From the Strategic National Stockpile for Treatment of 2019 Coronavirus Disease. (https://www.fda.gov/media/136534/download). Published March 28, 2020. Accessed October 20, 2020.
6. Disbrow GL. [Letter revoking emergency use authorization for chloroquine phosphate and hydroxychloroquine sulfate]. June 15, 2020. https://www.fda.gov/media/138945/download. (Accessed October 20, 2020).
7. US Food Drug Administration. FDA drug safety communication: 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. April 24, 2020. (https://www.fda.gov/safety/medical-product-safety-information/hydroxychloroquine-or-chloroquine-covid-19-drug-safety-communication-fda-cautions-against-use). Accessed October 20, 2020.
8. Kiley JP. NIH halts clinical trial of hydroxychloroquine. June 20, 2020. (https://www.nih.gov/news-events/news-releases/nih-halts-clinical-trial-hydroxychloroquine). Accessed October 20, 2020.
9. Recovery Collaborative Group. Statement from the chief investigators of the Randomised Evaluation of COVid-19 thERapY (RECOVERY) Trial on hydroxychloroquine, June 5, 2020. (https://www.recoverytrial.net/files/hcq-recovery-statement-050620-final-002.pdf). Accessed October 20, 2020.
10. World Health Organization. WHO discontinues hydroxychloroquine and lopinavir/ritonavir treatment arms for COVID-19. July 4, 2020. (https://www.who.int/news/item/04-07-2020-who-discontinues-hydroxychloroquine-and-lopinavir-ritonavir-treatment-arms-for-covid-19). Accessed October 20, 2020.
11. Pan H, Peto R, Karim QA, et al. Repurposed antiviral drugs for COVID-19—interim WHO SOLIDARITY trial results [preprint]. *medRxiv*2020. (https://doi.org/10.1101/2020.10.15.20209817). Accessed November 1, 2020.
12. Recovery Collaborative Group, Horby P, Mafham M, et al. Effect of hydroxychloroquine in hospitalized patients with Covid-19. *N Engl J Med*. 2020;383(21):2030–2040.
13. Self WH, Semler MW, Leither LM, et al. Effect of hydroxychloroquine on clinical status at 14 days in hospitalized patients with COVID-19: a randomized clinical trial. *JAMA*. 2020;324(21):2165–2176.
14. Benson K, Hartz AJ. A comparison of observational studies and randomized, controlled trials. *N Engl J Med*. 2000;342(25):1878–1886.
15. Alexander PE, Debono VB, Mammen MJ, et al. COVID-19 coronavirus research has overall low methodological quality thus far: case in point for chloroquine/hydroxychloroquine. *J Clin Epidemiol*. 2020;123:120–126.
16. Fiolet T, Guihur A, Rebeaud ME, et al. Effect of hydroxychloroquine with or without azithromycin on the mortality of coronavirus disease 2019 (COVID-19) patients: a systematic review and meta-analysis. *Clin Microbiol Infect*. 2021;27(1):19–27.

17. Singh AK, Singh A, Singh R, et al. Hydroxychloroquine in patients with COVID-19: a systematic review and meta-analysis. *Diabetes Metab Syndr*. 2020;14(4):589–596.
18. Hernandez AV, Roman YM, Pasupuleti V, et al. Hydroxychloroquine or chloroquine for treatment or prophylaxis of COVID-19: a living systematic review. *Ann Intern Med*. 2020;173(4):287–296.
19. Patil VM, Singhal S. Masand N. A systematic review on use of aminoquinolines for the therapeutic management of COVID-19: efficacy, safety and clinical trials. *Life Sci*. 2020; 254:117775.
20. Magagnoli J, Narendran S, Pereira F, et al. Outcomes of hydroxychloroquine usage in United States veterans hospitalized with COVID-19. *Med (N Y)*. 2020;1(1): 114–127.e3.
21. Kim AHJ, Sparks JA, Liew JW, et al. A rush to judgment? Rapid reporting and dissemination of results and its consequences regarding the use of hydroxychloroquine for COVID-19. *Ann Intern Med*. 2020;172(12): 819–821.
22. Pottegård A, Kurz X, Moore N, et al. Considerations for pharmacoepidemiological analyses in the SARS-CoV-2 pandemic. *Pharmacoepidemiol Drug Saf*. 2020;29(8): 825–831.
23. Zhai MZ, Lye CT, Kesselheim AS. Need for transparency and reliable evidence in emergency use authorizations for coronavirus disease 2019 (COVID-19) therapies. *JAMA Intern Med*. 2020;180(9):1145–1146.
24. Austin PC. An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate Behav Res*. 2011;46(3):399–424.
25. Lee BK, Lessler J, Stuart EA. Improving propensity score weighting using machine learning. *Stat Med*. 2010;29(3): 337–346.
26. McCaffrey DF, Ridgeway G, Morral AR. Propensity score estimation with boosted regression for evaluating causal effects in observational studies. *Psychol Methods*. 2004;9(4): 403–425.
27. Setodji CM, McCaffrey DF, Burgette LF, et al. The right tool for the job: choosing between covariate-balancing and generalized boosted model propensity scores. *Epidemiology*. 2017;28(6):802–811.
28. Yang JY, Webster-Clark M, Lund JL, et al. Propensity score methods to control for confounding in observational cohort studies: a statistical primer and application to endoscopy research. *Gastrointest Endosc*. 2019;90(3):360–369.
29. Westreich DJ, Lessler J, Jonsson Funk M. Propensity score estimation: machine learning and classification methods as alternatives to logistic regression. *J Clin Epidemiol*. 2011; 63(8):826–833.
30. Schneeweiss S, Rassen JA, Brown JS, et al. Graphical depiction of longitudinal study designs in health care databases. *Ann Intern Med*. 2019;170(6):398–406.
31. Chapman AB, Peterson KS, Turano A, et al. A natural language processing system for national COVID-19 surveillance in the US Department of Veterans Affairs. In: *Proceedings of the 1st Workshop on NLP for COVID-19 at ACL 2020*. Stroudsburg, PA: Association for Computational Linguistics;, 2020. https://www.aclweb.org/anthology/2020.nlpcovid19-acl.10. Accessed October 20, 2020.
32. Sohn MW, Arnold N, Maynard C, et al. Accuracy and completeness of mortality data in the Department of Veterans Affairs. *Popul Health Metr*. 2006;4(1):2.
33. Orkaby AR, Nussbaum L, Ho YL, et al. The burden of frailty among U.S. veterans and its association with mortality, 2002–2012. *J Gerontol A Biol Sci Med Sci*. 2019;74(8): 1257–1264.
34. R Core Team. *R: A Language and Environment for Statistical Computing, release 3.6.1*. Vienna, Austria: R Foundation for Statistical Computing; 2019.
35. Doove LL, Van Buuren S, Dusseldorp E. Recursive partitioning for missing data imputation in the presence of interaction effects. *Comput Stat Data Anal*. 2014;72: 92–104.
36. van Buuren S, Groothuis-Oudshoorn K. Mice: multivariate imputation by chained equations. (R package version 3.11.0). https://cran.r-project.org/package=mice. Published August 5, 2020. Accessed October 20, 2020.
37. Rubin DB, Schenker N. Multiple imputation in health-care databases: an overview and some applications. *Stat Med*. 1991;10(4):585–598.
38. Hastie T, Tibshirani R, Friedman J. Boosting and Additive Trees. In: *In: The Elements of Statistical Learning. Springer Series in Statistics*. New York, NY: Springer; 2009:1–51.
39. Greenwell B, Boehmke B, Cunningham J, et al. Gbm: generalized boosted regression models. (R package version 2.1.8). https://cran.r-project.org/package=gbm. Published July 15, 2020. Accessed October 1, 2020.
40. Greifer N. WeightIt: weighting for covariate balance in observational studies. (R package version 0.10.2). https://cran.r-project.org/package=WeightIt. Published August 27, 2020. Accessed October 20, 2020.
41. Greifer N. Cobalt: covariate balance tables and plots. (R package version 4.2.3). https://cran.r-project.org/package=cobalt. Published August 31, 2020. Accessed October 20, 2020.
42. Natekin A, Knoll A. Gradient boosting machines, a tutorial. *Front Neurorobot*. 2013;7(Dec):21.
43. Therneau TM. Survival: a package for survival analysis in S. (R package version 3.2–3). https://cran.r-project.org/package=survival. Published June 13, 2020. Accessed October 20, 2020.
44. McCaffrey DF, Griffin BA, Almirall D, et al. A tutorial on propensity score estimation for multiple treatments using generalized boosted models. *Stat Med*. 2013;32(19): 3388–3414.
45. Nguyen TL, Collins GS, Spence J, et al. Comparison of the ability of double-robust estimators to correct bias in propensity score matching analysis. A Monte Carlo simulation study. *Pharmacoepidemiol Drug Saf.* 2017; 26(12):1513–1519.
46. US Department of Veterans Affairs. Department of Veterans Affairs COVID-19 National Summary. https://www.accesstocare.va.gov/Healthcare/COVID19NationalSummary. Accessed August 20, 2020.
47. Austin PC, Stuart EA. Moving towards best practice when using inverse probability of treatment weighting (IPTW) using the propensity score to estimate causal treatment effects in observational studies. *Stat Med*. 2015;34(28): 3661–3679.
48. Rosenberg ES, Dufort EM, Udo T, et al. Association of treatment with hydroxychloroquine or azithromycin with in-hospital mortality in patients with COVID-19 in New York state. *JAMA*. 2020;323(24):2493–2502.
49. Rivera DR, Peters S, Panagiotou OA, et al. Utilization of COVID-19 treatments and clinical outcomes among patients with cancer: a COVID-19 and Cancer Consortium (CCC19) cohort study. *Cancer Discov*. 2020;10(10):1514–1527.
50. Sbidian E, Josse J, Lemaitre G, et al. Hydroxychloroquine with or without azithromycin and in-hospital mortality or

discharge in patients hospitalized for COVID-19 infection: a cohort study of 4,642 in-patients in France [preprint]. *medRxiv*2020. https://doi.org/10.1101/2020.06.16.20132597. Accessed October 20, 2020.
51. Arshad S, Kilgore P, Chaudhry ZS, et al. Treatment with hydroxychloroquine, azithromycin, and combination in patients hospitalized with COVID-19. *Int J Infect Dis*. 2020; 97(0):396–403.
52. Lee TC, MacKenzie LJ, McDonald EG, et al. An observational cohort study of hydroxychloroquine and azithromycin for COVID-19: (can't get no) satisfaction. *Int J Infect Dis*. 2020;98:216–217.
53. Catalogue of Bias Collaboration, Spencer EA, Heneghan C. Chronological bias. In: *Catalogue of Bias 2017*. https://catalogofbias.org/biases/chronological-bias/. Accessed October 2, 2020.
54. Bull-Otterson L, Gray EB, Budnitz DS, et al. Hydroxychloroquine and chloroquine prescribing patterns by provider specialty following initial reports of potential benefit for COVID-19 treatment - United States, January-June 2020. *MMWR Morb Mortal Wkly Rep*. 2020;69(35):1210–1215.
55. Rentsch CT, Kidwai-Khan F, Tate JP, et al. Patterns of COVID-19 testing and mortality by race and ethnicity among United States veterans: a nationwide cohort study. *PLoS Med*. 2020;17(9):e1003379.
56. Geleris J, Sun Y, Platt J, et al. Observational study of hydroxychloroquine in hospitalized patients with Covid-19. *N Engl J Med*. 2020;382(25):2411–2418.
57. Greenland S, Lanes S, Jara M. Estimating effects from randomized trials with discontinuations: the need for intent-to-treat design and G-estimation. *Clin Trials*. 2008; 5(1):5–13.