



## Article

# Analyzing the Difference in the Length of Stay (LOS) in Moderate to Severe COVID-19 Patients Receiving Hydroxychloroquine or Favipiravir

Bandar Alosaimi <sup>1</sup>, Huda M. Alshanbari <sup>2</sup>, Muath Alturaiqy <sup>3</sup>, Halah Z. AlRawi <sup>1</sup>, Saad Alamri <sup>1</sup>, Asma Albujaidy <sup>4</sup>, Aljawharah Bin Sabaan <sup>5</sup>, Ahmed A. Alrashed <sup>6</sup>, Ahmad Alamer <sup>7,8</sup>, Fayez Alghofaili <sup>9</sup>, Khaled Al-Duraymih <sup>10</sup>, Abdulaziz J. Alshalani <sup>11</sup> and Wael Alturaiki <sup>9,\*</sup>

<sup>1</sup> Research Center, King Fahad Medical City, Riyadh Second Health Cluster, Riyadh 11525, Saudi Arabia

<sup>2</sup> Department of Mathematical Sciences, College of Science, Princess Nourah bint Abdulrahman University, Riyadh 11671, Saudi Arabia

<sup>3</sup> Department of Internal Medicine, Alzulfi General Hospital, Riyadh Second Health Cluster, Riyadh 11525, Saudi Arabia

<sup>4</sup> Department of Clinical Pharmacy Service, Prince Mohammed bin Abdulaziz Hospital, Riyadh Second Health Cluster, Riyadh 11525, Saudi Arabia

<sup>5</sup> Pharmacy College, Almaarifa University, Riyadh 11597, Saudi Arabia

<sup>6</sup> Department of Pharmaceutical Services, Main Hospital, King Fahad Medical City, Riyadh Second Health Cluster, Riyadh 11525, Saudi Arabia

<sup>7</sup> Department of Clinical Pharmacy, Prince Sattam Bin Abdulaziz University, Alkharj 11942, Saudi Arabia

<sup>8</sup> Department of Pharmacy Practice and Science, College of Pharmacy, University of Arizona, Tucson, AZ 85721, USA

<sup>9</sup> Department of Medical Laboratory Sciences, College of Applied Medical Sciences, Majmaah University, Majmaah 11952, Saudi Arabia

<sup>10</sup> Main Laboratory and Blood Bank, Alzulfi General Hospital, Riyadh Second Health Cluster, Riyadh 11525, Saudi Arabia

<sup>11</sup> Alzulfi General Hospital, Riyadh Second Health Cluster, Riyadh 11525, Saudi Arabia

\* Correspondence: w.alturaiki@mu.edu.sa



**Citation:** Alosaimi, B.; Alshanbari, H.M.; Alturaiqy, M.; AlRawi, H.Z.; Alamri, S.; Albujaidy, A.; Bin Sabaan, A.; Alrashed, A.A.; Alamer, A.; Alghofaili, F.; et al. Analyzing the Difference in the Length of Stay (LOS) in Moderate to Severe COVID-19 Patients Receiving Hydroxychloroquine or Favipiravir. *Pharmaceuticals* **2022**, *15*, 1456. <https://doi.org/10.3390/ph15121456>

Academic Editor: Marco Fiore

Received: 14 October 2022

Accepted: 18 November 2022

Published: 24 November 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Abstract:** *Background:* The coronavirus 2019 (COVID-19) disease, caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) virus led to a global pandemic. HCQ and FPV were used early in the pandemic as a treatment modality for COVID-19. Various studies evaluated the HCQ and FPV effectiveness, based on the mortality endpoint and showed conflicting results. We hypothesize that analyzing the difference in the LOS as a significant endpoint would be of a major interest, especially for healthcare providers, to prevent a lengthy hospitalization and disease progression. *Methods:* This is a retrospective observational study, conducted via a medical chart review of COVID-19 patients who were admitted between April 2020 and March 2021 with a moderate to severe illness. The LOS endpoint was tested using the paired Wilcoxon signed-rank (WSR) model. Prior to using the WSR model, the balance between the HCQ and FPV groups, the propensity score matching, the LOS distribution, and the normality assumptions were tested. Two sensitivity statistical analyses were conducted to confirm the results (stratified log-rank test and U Welch test after transforming the LOS by the squared root values). *Results:* A total of 200 patients were included for the analysis: 83 patients in the HCQ group and 117 patients in the FPV group. Thirty-seven patients were matched in each group. The LOS data was positively skewed and violated the normality (Shapiro–Wilk  $p < 0.001$ ) and had an unequal variance (Levene's test,  $p = 0.019$ ). The WSR test showed no statistical significance in the LOS endpoint, with a median of  $-0.75$  days (95% confidence interval:  $-4.0$  to  $2.5$ ,  $p = 0.629$ ), in favor of the HCQ group (four days), in comparison to seven days of the FPV group. The WSR findings were further confirmed with the stratified log rank test ( $p = 740$ ) and the U Welch test ( $p = 391$ ). *Conclusions:* The study concluded that the HCQ and FPV treatments have a comparable effectiveness in terms of the LOS in the moderate to severe COVID-19 patients. This study highlights the importance of analyzing the LOS as a relevant endpoint, in order to prevent the costs of a lengthy hospitalization and disease progression. The current study also emphasizes the

importance of applying the appropriate statistical testing when dealing with two-sample paired data and analyzing non-parametric data such as the LOS.

**Keywords:** favipiravir; hydroxychloroquine; length of stay; COVID-19; SARS-CoV-2; effectiveness; statistical analysis

## 1. Introduction

The emergence of SARS-CoV-2 has caused a pandemic, with nearly half a billion confirmed cases of COVID-19, including over 6 million deaths, reported to the WHO [1]. In Saudi Arabia, as of April 2022, there have been more than 750 thousand confirmed cases with at least 9 thousand deaths. Although the majority of SARS-CoV-2 infections produce mild symptoms or are asymptomatic, some infected individuals require hospital care [2].

Although the safety and effectiveness of new COVID-19 investigational therapies continue to be evaluated, the U.S. FDA issued an Emergency Use Authorization (EUA) for the drugs paxlovid, molnupiravir, baricitinib and tocilizumab, which were shown to reduce to the recovery time [3]. To control an outbreak for a particular infectious disease, the discovery of new drugs appears to be the most effective way to treat patients. However, the time needed to develop and assess a new drug is lengthy.

There are many available antivirals that have proven to be successful with already established safety profiles, such as HCQ and FPV [4]. Furthermore, many agents have been recorded as having anti-SARS-CoV-2 effects, such as lopinavir and ritonavir, which were originally used to treat the HIV virus; remdesivir, an Ebola antiviral drug [5]; and dexamethasone. All have shown both improved outcomes and a reduced mortality [6]. HCQ has long been used for the treatment of malaria, and has become the standard treatment for COVID-19. HCQ can reduce the infection of SARS-CoV-2 by glycolizing the viral spike proteins and the human lung receptors, preventing the viral entry [7]. FPV has also been used in severe influenza cases. The drug can inhibit the RNA polymerase activity by both blocking the replication process and adversely acting on the genetic copying [8].

The Saudi protocol to treat patients with a confirmed COVID-19 has included HCQ and FPV, since the beginning of the pandemic [9]. However, few studies have been reported, comparing the treatment modality, based on multiple endpoints, to evaluate the effectiveness of HCQ and FPV. In addition, there is insufficient comparative data on the effectiveness of treatment analyzing the difference in the LOS in hospitals as a significant endpoint. For these reasons, this study, aimed to evaluate the effectiveness of HCQ and FPV, among moderate to severe COVID-19 Saudi patients, by analyzing the difference in the LOS, using the paired Wilcoxon signed-rank (WSR) model, in order to prevent the costs of a lengthy hospitalization and disease progression.

## 2. Results

We identified 278 patients admitted to the hospitals, of which 127 patients received HCQ, and 130 patients received FPV. All medical records were screened for eligibility to include only adults ( $\geq 18$  years old) with a confirmed diagnosis of moderate to severe COVID-19 and who received treatment of either HCQ or FPV. A total of 200 patients were included for the analysis (83 patients in the HCQ and 117 patients in the FPV groups). Prior to matching, the mean (SD) age for the HCQ group was 49.48 (15.46), compared to 62.25 (15.9) in the FPV group ( $p < 0.001$ ). Male patients were in the majority in the HCQ (71.1%) and FPV groups (60.7%) ( $p = 0.171$ ). More Saudi nationals were in the FPV group (81.2%), compared to 36.1% who received the HCQ treatment. There was no difference in the body mass index ( $\text{kg}/\text{m}^2$ ) between the two groups, with a mean (SD) of 30.2 (7.1) for the HCQ group and 28.8 (3.1) for the FPV group ( $p = 0.065$ ). Patients who received HCQ had a more severe illness (16.1%) and received fewer steroids (7.2%), compared to the FPV group with 11.1% and 27.4% of patients, respectively. Diabetes and hypertension were common in

the two groups but not statistically significant. Cardiovascular disease was more frequent in the HCQ group (9.6%) versus 1.7% in the FPV group. There were more patients with asthma in the HCQ group: 19.3% versus 4.3% in the FPV group. Chronic kidney disease was more prevalent in the HCQ group (6.0%) than in the FPV group (0.0%), as shown in Table 1.

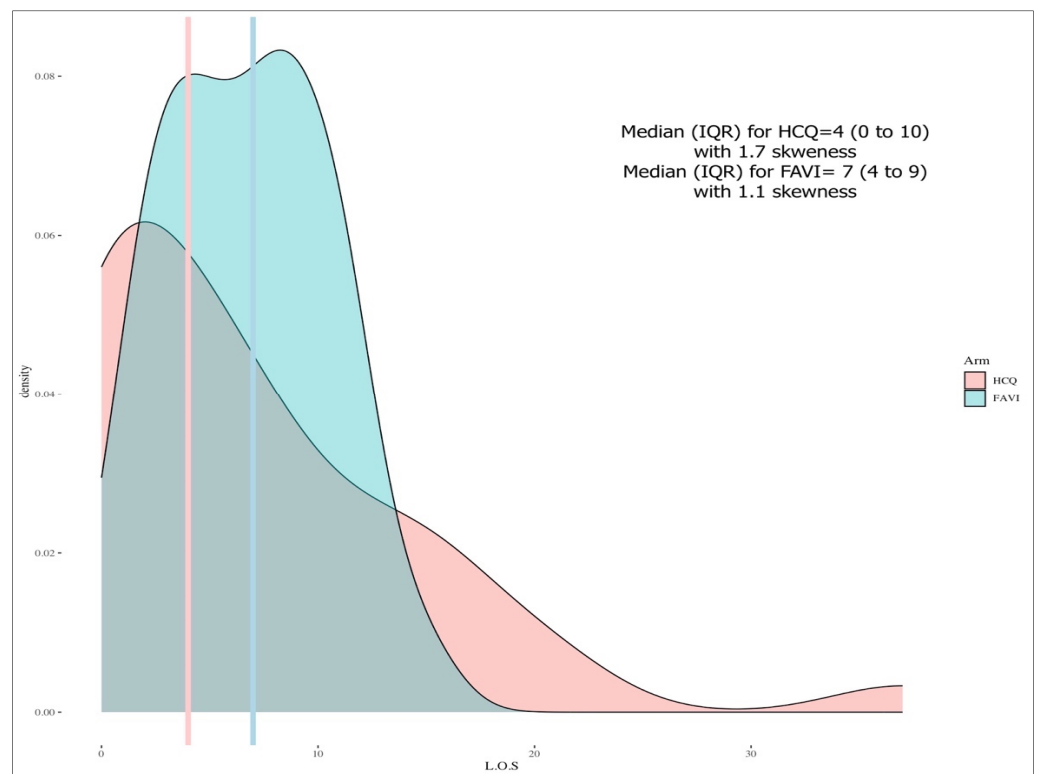
**Table 1.** Baseline characteristics (before and after the propensity score matching with the absolute standardized mean difference).

Variable	Total (n = 200)	HCQ (n = 83)	FPV (n = 117)	p-Value	Unadjusted Absolute SMD	HCQ (n = 37)	FPV (n = 37)	p-Value	Adjusted SMD
Age (years), mean (SD)	56.95 (16.9)	49.48 (15.46)	62.25 (15.9)	<0.001	0.8872	54.08 (17.49)	57.38 (16.22)	0.403	0.188
Male, n (%)	130 (65.0)	59 (71.1)	71 (60.7)	0.171		26 (70.3)	27 (73.0)	1.000	0.0417
Saudi nationality, n (%)	125 (62.5)	30 (36.1)	95 (81.2)	<0.001	0.4505	21 (56.8)	22 (59.5)	1.000	0.0487
Ethnicity, n(%) a									
Middle Eastern	136 (68.0)	37 (44.6)	99 (84.6)			25 (67.6)	25 (67.6)		
African	19 (9.5)	15 (18.1)	4 (3.4)			5 (13.5)	3 (8.1)		
South/Southeast Asia	43 (21.5)	29 (34.9)	14 (12.0)	<0.001	NA	7 (18.9)	9 (24.3)	NA	NA
Central Asia	1 (0.5)	1 (1.2)	0 (0.0)			0 (0.0)	0 (0.0)		
Unknown	1 (0.5)	1 (1.2)	0 (0.0)			0 (0.0)	0 (0.0)		
BMI (kg/m <sup>2</sup> )	29.40 (5.2)	30.2 (7.1)	28.83 (3.1)	0.065	0.3039	28.89 (4.20)	28.58 (2.91)	0.717	0.0650
Severe WHO classification, n (%)	27 (16.1)	14 (27.5)	13 (11.1)	0.015	0.1009	5 (13.5)	4 (10.8)	1.000	0.0208
Steroid, n (%)	38 (19.0)	6 (7.2)	32 (27.4)	0.001	0.2012	4 (10.8)	7 (18.9)	0.513	0.0587
Diabetes, n (%)	69 (34.5)	34 (41.0)	35 (29.9)	0.142	0.1105	8 (21.6)	9 (24.3)	1.000	0.0374
Hypertension, n (%)	82 (41.0)	33 (39.8)	49 (41.9)	0.877	0.0212	14 (37.8)	12 (32.4)	0.808	0.0529
Cardiovascular disease, n (%)	10 (5.0)	8 (9.6)	2 (1.7)	0.027	0.0793	1 (2.7)	2 (5.4)	1.000	0.0105
Asthma n (%)	21 (10.5)	16 (19.3)	5 (4.3)	<0.001	0.1500	5 (13.5)	4 (10.8)	1.000	0.0365
Chronic kidney disease, n (%)	5 (2.5)	5 (6.0)	0 (0.0)	0.026	0.0602	0 (0.0)	0 (0.0)	1.000	0.0000

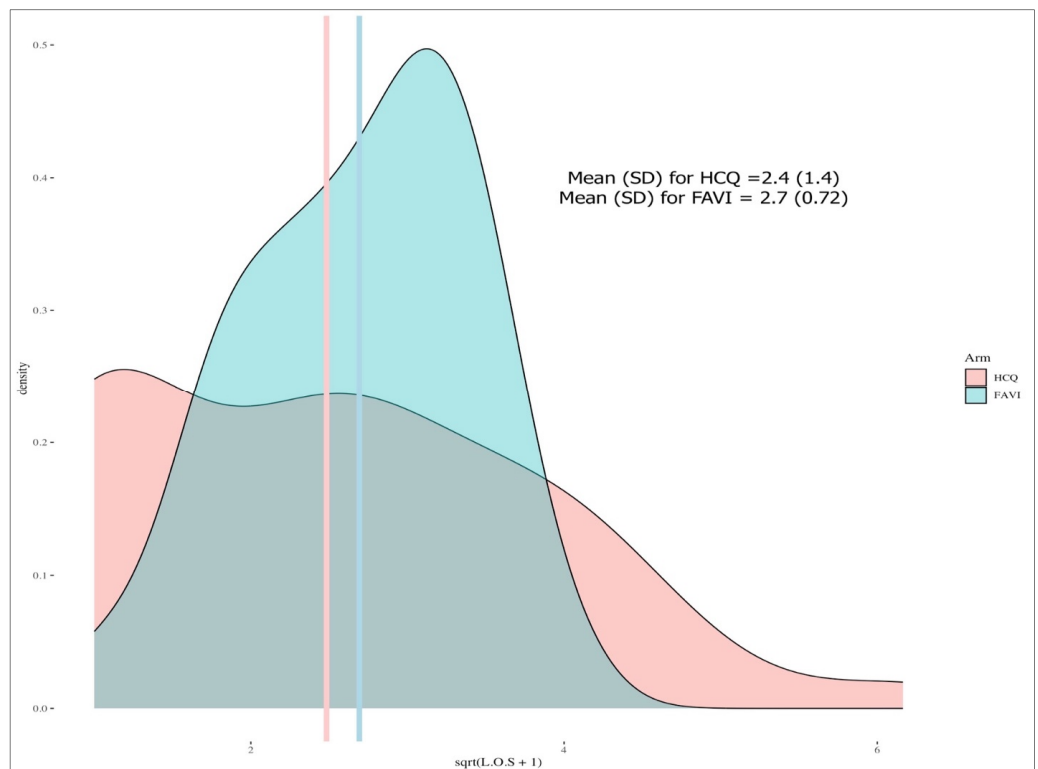
SD: Standard Deviation. SMD: standardized mean difference. BMI: body mass index. WHO: World Health Organization. Cardiovascular disease included the following: myocardial infarction, ischemic heart disease or stroke. (a): was not included in the propensity score model. Note: Multiple imputation using the multivariate imputation by the chained equations for the missing variables (only 0.5% in the BMI and severity (16%) were missing). SMD values are the preferred method to evaluate the covariate balance. Absolute SMD values < 0.2 indicated an adequate balance. Log transformation used age and the BMI for the propensity score modelling.

Once matched, the baseline characteristics were adequately balanced, with the SMD values < 0.2. The balance cohort is shown in Table 1. A total of 37 patients were matched in each group. The mirror plots for the propensity score distributions between the treated (FPV = 1) and control (HCQ = 0), indicated an adequate balance (Figure S1). The Love's plot provided a visual representation of the SMD values for the covariates of interest across the multiple inputted datasets (Figure S2).

The LOS distribution was positively skewed with a tail to the right (skewness of 1.7 in the HCQ group and 1.1 for the FPV group). The result of the Shapiro–Wilk test was a significant *p*-value < 0.001, as was the Levene's test, (*p*-value = 0.019), indicating an unequal variance. In the matched sample, the LOS median (IQR) for the HCQ was four days (0 to 10), and seven days (4 to 9) for the FPV, with no statistical difference, based on the Wilcoxon signed-rank test between the two groups (−0.75; 95% confidence interval (CI), −4.0 to 2.5, *p*-value = 0.629). A visual representation of these findings is shown in Figure 1A.



(A)

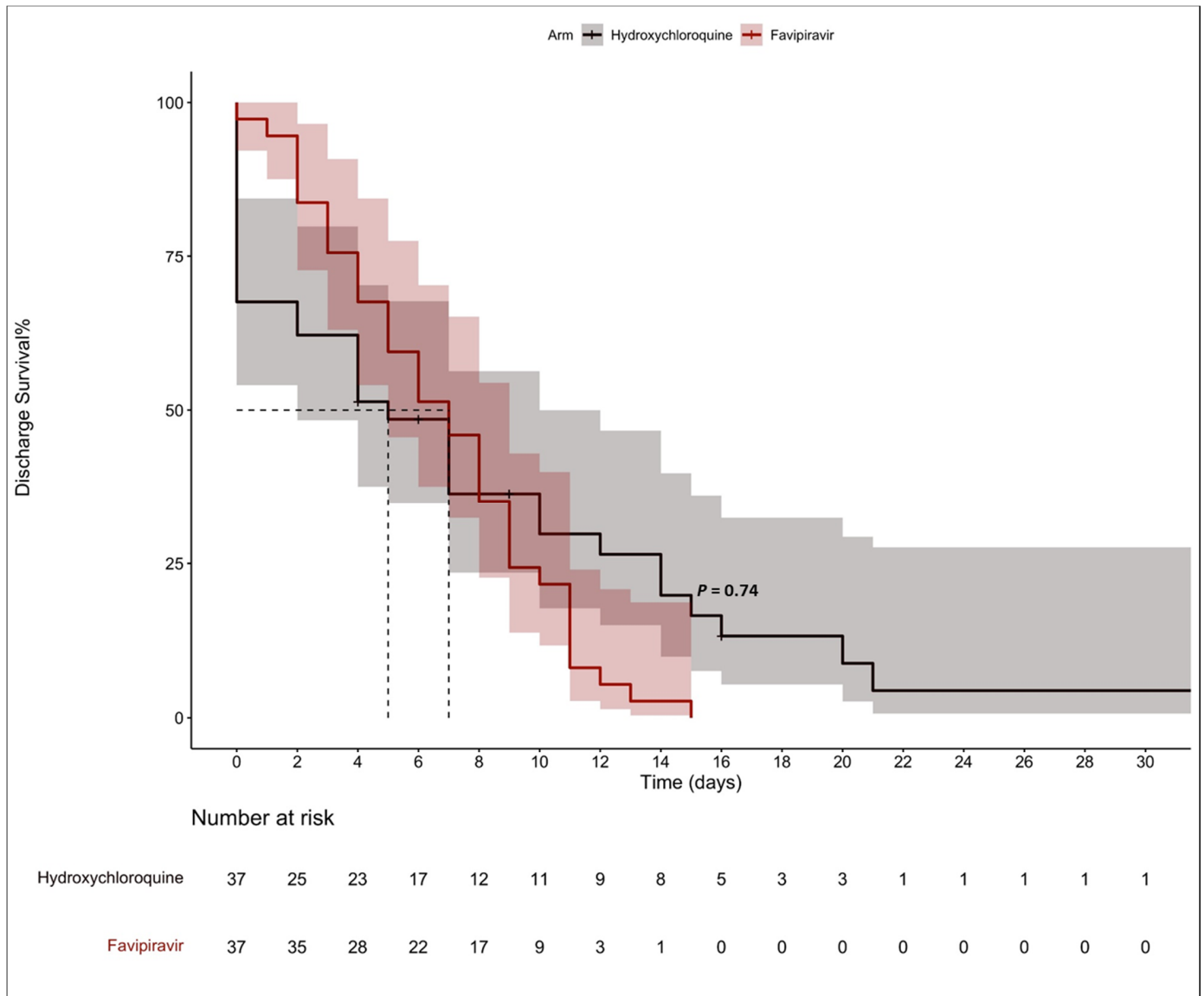


(B)

**Figure 1.** (A). Length of stay (LOS) density plot for the matched sample. Median presented for each group with the interquartile range (IQR). (B). square root length of stay (LOS) + 1-day density plot for the matched sample. Mean with the standard deviation (SD) presented for each group. HCQ = hydroxychloroquine. FAVI = Favipiravir. Vertical line for Figure 1A indicates median and Figure 1B are mean.

The Welch U test was also not significant, with a  $p$ -value = 0.391. The square root transformation made the LOS outcome close to the normal distribution; however, it did not improve the variance of the variable (Figure 1B).

The Kaplan–Meier model showed that the median time to discharge was five days (95% CI, 2 to 10) for the HCQ group and seven days for the FPV group (95% CI, 5 to 10), with no statistical significance between the two groups ( $p$ -value = 0.740), as shown in Figure 2. Two deaths occurred in the HCQ group and no deaths occurred in the FPV group,  $p$ -value = 0.493 (Table 2). None of the patients included in this study required mechanical ventilation during their hospital stay.



**Figure 2.** Kaplan–Meier for the median time to discharge in the matched sample. Dashed lines for the median. Decrease in survival indicates the discharge events.

**Table 2.** Clinical outcomes in the matched sample.

Matched Sample	HCQ (n = 37)	FPV (n = 37)	p-Value
Outcome			
LOS (days), median (IQR) <sup>a</sup>	4 (0 to 10)	7 (4 to 9)	p = 0.629
Transformed LOS (days), mean (SD) <sup>b</sup>	2.4(1.4)	2.7 (0.72)	p = 0.391
Median time to discharge (95% CI) <sup>c</sup>	5 (2 to 10)	7 (5 to 10)	p = 0.740
Death, n(%) <sup>d</sup>	2 (5.4)	0 (0.0)	p = 0.493

<sup>a</sup> Paired Wilcoxon signed rank test. Pseudo (median) difference was  $-0.75$  (95% confidence interval:  $-4$  to  $2.5$ ).

<sup>b</sup> Transformation by the square root. Welch U test was conducted. Mean difference was  $-0.21$  (95% confidence interval:  $-0.70$  to  $0.28$ ). <sup>c</sup> Kaplan–Meier estimates. *p* value was calculated using the stratified log rank test.

<sup>d</sup> Fisher exact test was performed.

### 3. Discussion

The study provided an opportunity to advance the knowledge of analyzing the effect of HCQ versus FPV on the LOS, as a surrogate for the marker effectiveness. We observed that both drugs showed a comparable median LOS in moderate to severe COVID-19 patients. There have been few head-to-head comparative quantitative studies that investigated the clinical outcomes of HCQ and FPV [4,10]. Previous studies have only focused on laboratory tests, such as the clearance of SARS-CoV-2 by the PCR, and radiological improvements which may not reflect the clinical recovery [4,10]. The little data in the literature are controversial, and there is no general agreement about the difference in the LOS and mortality outcomes between the HCQ and FPV groups.

The most likely explanations of the contradictory findings between both groups are poor study designs, heterogeneity of the study populations, and power of the study, small sample sizes, early vs. late initiation of the antiviral treatment, diversity of dosing, therapy regimen and duration of the treatment plan [10,11]. There are several studies comparing the LOS of an HCQ-based regimen to a control arm. In a randomized control trial, Abd-Elsalam et al., reported that the mean LOS did not differ between the HCQ and the supportive groups, with 11.04 days and 11.27 days, respectively (*p*-value = 0.52) [12]. Likewise, an open-label RCT study, compared the LOS of moderate COVID-19 cases in patients receiving HCQ, lopinavir/ritonavir, and lopinavir/ritonavir plus interferon, and found that they were identical, with a mean LOS of 11, 12, 11 days, respectively (*p*-value = 0.20) [13]. In contrast, a meta-analysis demonstrated that antiparasitic drugs, including HCQ, were associated with a shorter length of hospital stay, in comparison with the standard of care in a control group who were taking a placebo (*p*-value < 0.05) [14]. Conversely, Alamer et al., found that the FPV group had a shorter median time to discharge among the moderate COVID-19 cases, than the supportive-care group of nine days (95% CI, 9 to 10) versus 11 days (95% CI, 10 to 12) [15].

As many studies showed inconsistencies in the LOS outcomes, a debate about mortality with HCQ and FPV has also created interest within the scientific community. Studies reported varying findings, ranging from large decreases in mortality [16], no changes in mortality [17] and moderate to large increases in mortality [18]. One study raised enough red flags to stop the use of HCQ, with the preliminary analysis revealing that the risk unexpectedly outweighed the benefits [19]. The reported adverse events are an increased all-cause mortality, ICU admission, mechanical ventilation, QTc interval prolongation, and ventricular arrhythmias [19]. The UK Medicines and Healthcare Products Regulatory Agency requested that the data be scrutinized and further reviewed. Unfortunately, their analysis was prompted by the retractions of high-profile papers in journals, such as *The Lancet* [20].

Based on the current data available, the use of HCQ and FPV was limited because numerous anti-COVID-19 drugs were found to be potentially useful, with various levels of supporting evidence [5,6,21]. We noted in our study that the mean age in the unmatched

sample was 49 years in the HCQ group and 62 in the FPV group. This finding may be related to the prescribers' preferences or to institutional policies that considered HCQ to be unsafe for the elderly, given the multiple reports of safety-related concerns [20,22,23]. The age discrepancy favoring the young for HCQ is not unique to this study. The study by Alotaibi et al., conducted in Saudi Arabia, reported a similar mean age (mean = 49.6 with SD = 12.6), which was statistically significant ( $p$ -value < 0.001), compared with the FPV group (mean = 54.8 with SD = 14.9) [4]. The higher age range in the FPV may also be related to the relative safety of the medication, as more information has emerged since its approval [24].

As many COVID-19 treatments failed to show mortality benefits [25], it is critical to evaluate the LOS endpoint, as COVID-19 patients can put a strain on the healthcare system [26]. A caveat is that much of the published literature failed to report meeting the statistical test assumptions, giving rise to the question of the validity of the statistical tests conducted [27,28]. In the case of COVID-19, this may be related to rapid research reporting and dissemination, which can lead to the misinterpretation of the data [29].

In order to analyze our primary endpoint accurately, our study addressed the significant baseline characteristics imbalance we encountered, using propensity score matching. In our propensity score model, important variables of interest were included. The procedure reduced our sample size ( $n = 37$  in each group) but ensured the comparability between the two groups. Figure 1A shows the probability density function for the outcome of the LOS which was positively skewed with a tail to the right, with few outliers for the HCQ group.

Statistical test assumptions are often violated when analyzing the difference in the LOS between two samples. Failure to test for the LOS distribution can lead to selecting an inappropriate statistical test. For instance, a literature review from five major emergency medicine journals (data from 2004 to 2007) with a focus in the emergency department (ED) LOS endpoint, revealed that 80% of the included studies did not perform a test of normality on the LOS data and 43% of the studies failed to perform the appropriate non-parametric tests [30]. Applying a parametric test to the non-normally distributed data may result in a type II error-failure to detect the difference. Chazard et al., a total of 12 statistical tests commonly used to analyze the LOS data were evaluated. In the case of a balanced sample (equal sizes of  $n$ ), it is advisable to use either the Wilcoxon test, or the Kruskal–Wallis test (known as pure shift models) or Student's  $t$ -test with a log or rank transformation, when comparing the LOS non-parametric data. These tests give high power with relative efficiency, compared with Student's  $t$ -test (Student's  $t$ -test is very conservative, with a type I error < 0.5%) [31]. The log-rank test was not appropriate in the simulated study as a statistical test for the LOS, as it gave unacceptable high  $\alpha$  errors. In our study, the Wilcoxon test agreed with the stratified log-rank test that there was no difference between the two groups in the LOS endpoint. However, the simulated study did not address the problem of the unequal variance as a test assumption for the Wilcoxon test. The unequal variance could have affected the accuracy of the Wilcoxon test. Therefore, we decided to modify the LOS by the square root transformation. The square root transformation made the LOS outcome close to a normal distribution but did not improve the variance of the variable. A re-test for our hypothesis, using the Welch U test with the log LOS also produced non-significant results ( $p$ -value = 0.391).

Finally, several important limitations need to be considered. The study was retrospective in nature. We did not evaluate the safety profile of HCQ, despite the existing an international debate about the considerable cardiovascular adverse events reported. As the study data was collected early in the pandemic, it did not capture the effect of the vaccine status, improvements in the treatment protocols, nor emergence of COVID-19 variants. Moreover, we did not examine the effect of removing the outliers from the LOS outcomes, since this would have reduced the sample size and created unequal sizes, calling into question the validity of our statistical tests. The numbers of patients and controls were also relatively small, which may have affected the power of the study and the generalizability

of the results. The distribution of the LOS data could have been improved by increasing the sample size or selecting a statistical test that can achieve a greater power while controlling the type I errors. Finally, despite the use of the propensity score matching to account for the confounders, the residual confounding cannot be entirely ruled out.

## 4. Methodology

### 4.1. Study Design and Setting

The research was a retrospective observational study conducted via a medical chart review of two Riyadh Second Health Cluster hospitals: Prince Mohammed Bin Abdulaziz Hospital (500 beds) and Al-Zulfi General Hospital (150 beds). The Institutional Review Board at the tertiary specialty referral hospital in the cluster, known as King Fahad Medical City in Riyadh, approved the study (IRB 22-175).

### 4.2. Selection Criteria

A list of COVID-19 patients admitted to the two hospitals between April 2020 and March 2021 and who had received either FPV or HCQ treatments, was obtained. To minimize bias, a random-selection process was implemented while screening for eligibility. Patients with a confirmed diagnosis of COVID-19 by a PCR had an equal opportunity to be selected if they were at least 18 years old, received either FPV or HCQ treatment and had been admitted with a moderate to severe infection, as classified by the World Health Organization (WHO). Patients with mild or critical infections, classified by the WHO, who had an incomplete medical record, or were admitted outside of the study duration, were excluded from the study.

### 4.3. Definitions

We classified patients according to the WHO definition into moderate or severe [32]. Patients with non-severe pneumonia, who were breathing normally without supplemental oxygen, were classified as moderate patients. A severe COVID-19 disease was defined as the development of fever, in addition to one or more of the following symptoms: respiratory rate  $\geq 30$ /min, dyspnea, respiratory distress,  $SpO_2 \leq 93\%$  on room air,  $PaO_2/FiO_2$  ratio  $< 300$  or lung infiltrate  $> 50\%$  of the lung field within one to two days. The length of stay (LOS) was defined as the date of admission minus the date of discharge and was reported in days.

### 4.4. Study Outcomes

The primary endpoint was the difference in the LOS between the two treatments. The second endpoint was the mortality difference between the two treatments.

### 4.5. Baseline Information Collection

The demographic data that included age, sex, body mass index ( $kg/m^2$ ), ethnicity, comorbidities, steroid use, length of stay, respiratory assistance and patients' outcomes, were extracted. Data quality during the data collection was periodically checked to assess the consistency and to resolve any discrepancies.

### 4.6. Sample Size Calculation and Statistical Analysis

We calculated the sample size, based on the assumption that the LOS distribution would violate the normality distribution. The sample size was calculated using the MKpower package [33]. The package is written in R language with various functions to calculate the sample size for the Wilcoxon rank sum and signed rank tests. In our parameter inputs, we assumed a mean difference of  $-2$  and a standard deviation (SD) of 2 between the two treatment groups for a  $0.05 \alpha$  significance level (type 1 error rate). Based on our calculations, a sample size of 20 patients in each group was needed to achieve a power (beta)  $> 80\%$ .



Continuous variables with normal distributions were presented as mean  $\pm$  (standard deviation [SD]) and the statistical significance between the two groups was tested using Student's t-test. The Chi-squared test and the Fisher exact test were used to compare the categorical data. There was a few missing data points in the dataset (such as the body mass index), due to the nature of the study (observational and dependent on a chart review). In order to perform an adequate analysis for the propensity score matching, a complete dataset is required to calculate the propensity scores for each case. For the main analysis, the missing data were first inputted using the multivariate imputation by chained equations (MICE), based on the assumption that the data were missing at random (five imputations with 10 iterations). To induce a balance between the two treatment groups, the treatment assignment was modelled, using a logistic regression model that included the measured pre-treatment covariates of interest. The log transformations were applied to the continuous variables for modeling. Within the datasets, the nearest neighbor method with a caliper of 0.2, were the specifications for the propensity scores matching procedure. Less than a 0.2 absolute standardized mean difference (SMD) indicated an adequate balance on the measured covariate. Love's plot for the covariate balance and mirror plots for the propensity score distributions were generated (Figures S1 and S2).

The null hypothesis (H0) for the LOS outcome stated that there was no difference between the medians or means of the LOS of the two treatment groups. The analysis was carried out using only matched data. The LOS was presented as the median with interquartile ranges (IQRs). The normality was checked for the LOS outcome using the Shapiro–Wilk test and the Q-Q plots. An equal variance assumption was tested using Levene's test. The paired Wilcoxon signed-rank test [34] was chosen to test the median differences in the LOS between the two treatment groups to account for the matched data [13]. Two sensitivity statistical analyses were conducted to confirm the results. First, the Welch U test was used, assuming an unequal variance with the square root transformation of the LOS, to test the difference between the two means. All days were shifted by one day for this test. The second test was using the survival analysis. For the second method, a Kaplan–Meier model was used to estimate the median time to discharge. Patients who were discharged were considered as events. To test for the significance in the survival curves, a stratified log-rank test was used for the matched pairs [35].

#### 4.7. Software Used

R Core Team (2021) software (R Foundation for Statistical Computing, Version 4.1.2, Vienna, Austria) was used. The following R packages were used to generate results: ggplot2 [36], matchthem [37], cobalt [38], MKpower [33], mice [39] and survival [40].

## 5. Conclusions

Our study concluded that HCQ and FPV have a comparable effectiveness, in terms of the LOS in moderate to severe COVID-19 patients. Although not statistically significant, the median interquartile range of the LOS for patients treated with HCQ was nearly half of the LOS time for patients treated with FPV. The study highlights the importance of healthcare providers analyzing the LOS as a relevant endpoint to reduce the costs of a lengthy hospitalization and disease progression. The current study also emphasizes the importance of applying appropriate statistical testing when dealing with non-parametric data, such as the LOS.

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/ph15121456/s1>, Figure S1: Mirror plot for the propensity scores distribution balance between the two groups (both prior and after matching) after multiple imputation. 0 = hydroxychloroquine and 1 = Favipravir; Figure S2: Love's plot for the covariate balance between the two groups (prior and after matching) after multiple imputation.

**Author Contributions:** Conceptualization, B.A., A.A. (Ahmad Alamer), M.A. and W.A.; formal analysis, A.A. (Ahmad Alamer), H.M.A. and A.A.A.; investigation, H.Z.A. and S.A.; visualization, F.A., K.A.-D. and A.J.A.; sample collection, A.A. (Asma Albujaidyand) and A.B.S.; writing—original draft, B.A., A.A. (Ahmad Alamer), H.M.A. and A.A.A.; writing—review and editing, B.A., A.A. (Ahmad Alamer) and W.A. All authors have read and agreed to the published version of the manuscript.

**Funding:** Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2022R299), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia.

**Institutional Review Board Statement:** The protocol was approved by the Institutional Review Board at King Fahad Medical City (IRB 22-175). Hence, the written informed consent was waived due to the retrospective nature of the study and it was not required, since only unidentifiable data were extracted from the medical records.

**Informed Consent Statement:** Informed consent was waived due to the retrospective nature of the study.

**Data Availability Statement:** Data is contained within the article and Supplementary Material.

**Acknowledgments:** The authors also thank the Research Center at King Fahad Medical City for their technical support. Extended thanks to the administration of Alzulfi General Hospital and Prince Mohammad bin Abdulaziz hospital for their help in providing the data.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. WHO. *Coronavirus Disease (COVID-19) Situation Report*; World Health Organization: Geneva, Switzerland, 2022.
2. Wu, Z.; McGoogan, J.M. Characteristics of and important lessons from the coronavirus disease 2019 (COVID-19) outbreak in China: Summary of a report of 72,314 cases from the Chinese Center for Disease Control and Prevention. *JAMA* **2020**, *323*, 1239–1242. [[CrossRef](#)] [[PubMed](#)]
3. COVID-19 Treatment Guidelines Panel. Coronavirus Disease 2019 (COVID-19) Treatment Guidelines. National Institutes of Health. Available online: <https://www.covid19treatmentguidelines.nih.gov/> (accessed on 4 April 2022).
4. Alotaibi, M.; Ali, A.; Bakhshwin, D.; Alatawi, Y.; Alotaibi, S.; Alhifany, A.; Alharthi, B.; Alharthi, N.; Alyazidi, A.; Alharthi, Y.; et al. Effectiveness and Safety of Favipiravir Compared to Hydroxychloroquine for Management of COVID-19: A Retrospective Study. *Int. J. Gen. Med.* **2021**, *14*, 5597–5606. [[CrossRef](#)] [[PubMed](#)]
5. Costanzo, M.; De Giglio, M.A.R.; Roviello, G.N. SARS-CoV-2: Recent Reports on Antiviral Therapies Based on Lopinavir/Ritonavir, Darunavir/Umifenovir, Hydroxychloroquine, Remdesivir, Favipiravir and other Drugs for the Treatment of the New Coronavirus. *Curr. Med. Chem.* **2020**, *27*, 4536–4541. [[CrossRef](#)]
6. Group, T.R.C. Dexamethasone in hospitalized patients with COVID-19—Preliminary report. *N. Engl. J. Med.* **2020**, *384*, 693–704.
7. Wei, P.F. Diagnosis and Treatment Protocol for Novel Coronavirus Pneumonia (Trial Version 7). *Chin. Med. J.* **2020**, *133*, 1087–1095. [[CrossRef](#)]
8. Udawadia, Z.F.; Singh, P.; Barkate, H.; Patil, S.; Rangwala, S.; Pendse, A.; Kadam, J.; Wu, W.; Caracta, C.F.; Tandon, M. Efficacy and safety of favipiravir, an oral RNA-dependent RNA polymerase inhibitor, in mild-to-moderate COVID-19: A randomized, comparative, open-label, multicenter, phase 3 clinical trial. *Int. J. Infect. Dis.* **2020**, *103*, 62–71. [[CrossRef](#)]
9. Saudi Arabia Ministry of Health (MOH). Saudi MoH Protocol for Patients Suspected of/Confirmed with COVID-19. Available online: <https://www.moh.gov.sa/Ministry/MediaCenter/Publications/Documents/MOH-therapeutic-protocol-for-COVID-19.pdf> (accessed on 2 September 2020).
10. Saudi Arabia Ministry of Health (MOH). Coronavirus Disease COVID-19 Guidelines, v1.3. Available online: <https://www.moh.gov.sa/Ministry/MediaCenter/Publications/Documents/Coronavirus-Disease-2019-Guidelines-v1.2.pdf> (accessed on 2 September 2020).
11. Kohl, M. Package MKpower: Wilcoxon Rank Sum and Signed Rank Tests. 2020. Available online: <https://cran.r-project.org/web/packages/MKpower/vignettes/MKpower.html#wilcoxon-rank-sum-and-signed-rank-tests> (accessed on 19 July 2022).
12. Scheff, S. *Fundamental Statistical Principles for the Neurobiologist, Chapter 8—Nonparametric Statistics*; Academic Press: Cambridge, MA, USA, 2016. [[CrossRef](#)]
13. Austin, P.C. A Tutorial and Case Study in Propensity Score Analysis: An Application to Estimating the Effect of In-Hospital Smoking Cessation Counseling on Mortality. *Multivar. Behav. Res.* **2011**, *46*, 119–151. [[CrossRef](#)]
14. Wickham, H. ggplot2: Create Elegant Data Visualisations Using the Grammar of Graphics. 2020. Available online: <https://cran.r-project.org/web/packages/ggplot2/index.html> (accessed on 19 July 2022).
15. Pishgar, F. MatchThem: Matching and Weighting Multiply Imputed Datasets. 2020. Available online: <https://cran.r-project.org/web/packages/MatchThem/index.html> (accessed on 19 July 2022).

16. Greife, N. Cobalt: Covariate Balance Tables and Plots. 2020. Available online: <https://cran.r-project.org/web/packages/cobalt/index.html> (accessed on 19 July 2022).
17. Buuren, S.V. Mice: Multivariate Imputation by Chained Equations. 2020. Available online: <https://cran.r-project.org/web/packages/mice/index.html> (accessed on 19 July 2022).
18. Therneau, T. Survival: Survival Analysis. 2020. Available online: <https://cran.r-project.org/web/packages/survival/index.html> (accessed on 19 July 2022).
19. Eroglu, A.; Kartal, S.; Kongur, E. Comparison of the intensive care unit admission rate and mortalities of COVID-19 patients who received Hydroxychloroquine and Favipiravir in the ward. *J. Anesth. Crit. Care Open Access* **2022**, *14*, 40–43. [[CrossRef](#)]
20. Turan, D.B.; Menteş, M.; Özel, Y.; Şerefhanoglu, K.; Aydoğan, B.; İbil, N.; Güneşdoğdu, F.; Orucova, H.M.; Saltürk, C.; Çelik, H. The effect of favipiravir versus hydroxychloroquine on clinical and laboratory findings in COVID-19 in healthcare workers. *Braz. J. Infect. Dis.* **2022**, *26*, 102328. [[CrossRef](#)]
21. Abd-El Salam, S.; Esmail, E.S.; Khalaf, M.; Abdo, E.F.; Medhat, M.A.; Abd El Ghafar, M.S.; Ahmed, O.A.; Soliman, S.; Serangawy, G.N.; Alborai, M. Hydroxychloroquine in the treatment of COVID-19: A multicenter randomized controlled study. *Am. J. Trop. Med. Hyg.* **2020**, *103*, 1635–1639. [[CrossRef](#)] [[PubMed](#)]
22. Ader, F.; Peiffer-Smadja, N.; Poissy, J.; Bouscambert-Duchamp, M.; Belhadi, D.; Diallo, A.; Delmas, C.; Saillard, J.; Dechanet, A.; Mercier, N.; et al. An open-label randomized controlled trial of the effect of lopinavir/ritonavir, lopinavir/ritonavir plus IFN- $\beta$ -1a and hydroxychloroquine in hospitalized patients with COVID-19. *Clin. Microbiol. Infect.* **2021**, *27*, 1826–1837. [[CrossRef](#)] [[PubMed](#)]
23. Akbari, A.; Razmi, M.; Sedaghat, A.; Dana, S.M.M.A.; Amiri, M.; Halvani, A.M.; Yazdani, S.; Sahab-Negah, S. Comparative effectiveness of pharmacological interventions on mortality and the average length of hospital stay of patients with COVID-19: A systematic review and meta-analysis of randomized controlled trials. *Expert Rev. Anti-Infect. Ther.* **2021**, *20*, 585–609. [[CrossRef](#)] [[PubMed](#)]
24. Alamer, A.; Alrashed, A.A.; Alfaihi, M.; Alosaimi, B.; AlHassar, F.; Almutairi, M.M.; Howaidi, J.; Almutairi, W.; Mohzari, Y.; Sulaiman, T.; et al. Effectiveness and safety of favipiravir compared to supportive care in moderately to critically ill COVID-19 patients: A retrospective study with propensity score matching sensitivity analysis. *Curr. Med. Res. Opin.* **2021**, *37*, 1085–1097. [[CrossRef](#)]
25. Membrillo, F.J.; Ramírez-Olivencia, G.; Estébanez, M.; de Dios, B.; Herrero, M.D.; Mata, T.; Borobia, A.M.; Gutiérrez, C.; Simón, M.; Ochoa, A.; et al. Early hydroxychloroquine is associated with an increase of survival in COVID-19 patients: An observational study. *Preprints* **2020**, 2020050057. [[CrossRef](#)]
26. Mahévas, M.; Tran, V.T.; Roumier, M.; Chabrol, A.; Paule, R.; Guillaud, C.; Gallien, S.; Lepeule, R.; Szwebel, T.A.; Lescure, X.; et al. No evidence of clinical efficacy of hydroxychloroquine in patients hospitalised for COVID-19 infection and requiring oxygen: Results of a study using routinely collected data to emulate a target trial. *MedRxiv* **2020**. [[CrossRef](#)]
27. Geleris, J.; Sun, Y.; Platt, J.; Zucker, J.; Baldwin, M.; Hripcsak, G.; Labella, A.; Manson, D.K.; Kubin, C.; Barr, R.G.; et al. Observational study of hydroxychloroquine in hospitalized patients with COVID-19. *N. Engl. J. Med.* **2020**, *382*, 2411–2418. [[CrossRef](#)]
28. Borba, M.G.S.; Val, F.F.A.; Sampaio, V.S.; Alexandre, M.A.A.; Melo, G.C.; Brito, M.; Mourão, M.P.G.; Brito-Sousa, J.D.; Baía-da-Silva, D.; Guerra, M.V.F.; et al. Chloroquine diphosphate in two different dosages as adjunctive therapy of hospitalized patients with severe respiratory syndrome in the context of coronavirus (SARS-CoV-2) infection: Preliminary safety results of a randomized, double-blinded, phase IIb clinical trial (CloroCOVID-19 Study). *MedRxiv* **2020**. [[CrossRef](#)]
29. Torjesen, I. COVID-19: Hydroxychloroquine does not benefit hospitalised patients, UK trial finds. *BMJ* **2020**, *369*, m2263. [[CrossRef](#)]
30. Extance, A. COVID-19: What is the evidence for the antiviral Paxlovid? *BMJ* **2022**, *37*, o10377. [[CrossRef](#)]
31. Gao, J.; Tian, Z.; Yang, X. Breakthrough: Chloroquine phosphate has shown apparent efficacy in treatment of COVID-19 associated pneumonia in clinical studies. *Biosci. Trends* **2020**, *14*, 72–73. [[CrossRef](#)]
32. Gautret, P.; Lagier, J.C.; Parola, P.; Hoang, V.T.; Meddeb, L.; Mailhe, M.; Doudier, B.; Courjon, J.; Giordanengo, V.; Vieira, V.E.; et al. Hydroxychloroquine and azithromycin as a treatment of COVID-19: Results of an open-label non-randomized clinical trial. *Int. J. Antimicrob. Agents* **2020**, *56*, 105949. [[CrossRef](#)] [[PubMed](#)]
33. Pilkington, V.; Pepperrell, T.; Hill, A. A review of the safety of favipiravir—A potential treatment in the COVID-19 pandemic? *J. Virus Erad.* **2020**, *6*, 45–51. [[CrossRef](#)] [[PubMed](#)]
34. Yan, V.C.; Muller, F.L. Why Remdesivir Failed: Preclinical Assumptions Overestimate the Clinical Efficacy of Remdesivir for COVID-19 and Ebola. *Antimicrob. Agents Chemother.* **2021**, *65*, e0111721. [[CrossRef](#)] [[PubMed](#)]
35. Guo, A.; Lu, J.; Tan, H.; Kuang, Z.; Luo, Y.; Yang, T.; Xu, J.; Yu, J.; Wen, C.; Shen, A. Risk factors on admission associated with hospital length of stay in patients with COVID-19: A retrospective cohort study. *Sci. Rep.* **2021**, *11*, 7310. [[CrossRef](#)]
36. Patino, C.M.; Ferreira, J.C. Meeting the assumptions of statistical tests: An important and often forgotten step to reporting valid results. *J. Bras. Pneumol.* **2018**, *44*, 353. [[CrossRef](#)] [[PubMed](#)]
37. Hoekstra, R.; Kiers, H.; Johnson, A. Are Assumptions of Well-Known Statistical Techniques Checked, and Why (Not)? *Front. Psychol.* **2012**, *3*, 137. [[CrossRef](#)]

38. Kim, A.H.; Sparks, J.A.; Liew, J.W.; Putman, M.S.; Berenbaum, F.; Duarte-García, A.; Graef, E.R.; Korsten, P.; Sattui, S.E.; Sirolich, E.; 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*, 819–821. [[CrossRef](#)]
39. Qualls, M.; Pallin, D.J.; Schuur, J.D. Parametric Versus Nonparametric Statistical Tests: The Length of Stay Example. *Acad. Emerg. Med.* **2010**, *17*, 1113–1121. [[CrossRef](#)]
40. Chazard, E.; Ficheur, G.; Beuscart, J.-B.; Preda, C. How to Compare the Length of Stay of Two Samples of Inpatients? A Simulation Study to Compare Type I and Type II Errors of 12 Statistical Tests. *Value Health* **2017**, *20*, 992–998. [[CrossRef](#)]