PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Examination of patient characteristics and hydroxychloroquine use based on U.S. Food and Drug Administration's recommendation: a cross-sectional analysis in New York
AUTHORS	Kim, Eun Ji; Coppa, Kevin; Hirsch, Jamie; Abrahams, Sara; Johnson, Jennifer; Lesser, Martin; Davidson, Karina W.; Conigliaro, Joseph

VERSION 1 – REVIEW

REVIEWER	Jean-David Zeitoun Centre d'Epidémiologie Clinique, Hôtel Dieu Hospital, APHP, Paris, France Shareholder of a company whose Sanofi is a customer
REVIEW RETURNED	01-Sep-2020

GENERAL COMMENTS	Thank you very much for allowing me to review this manuscript of high interest. I congratulate the authors for their efforts and findings. Overall, I think that the manuscript deserves publication. However, I have some significant remarks to suggest. Most of them relate to the clarity of the work and its purposes (which were not obvious to me at first sight) First, the abstract needs a revision. It is not even mentioned in the Objective that this is related to the COVID-19 pandemic. I know that this has become the only subject of conversation but still. I think in particular about the long term life of the article. In the results (of the abstract), the total number of patients after the FDA warning seems to be missing. (ditto for the article summary) Last, I am not sure that the first sentence of the conclusion replies to the objectives of the work. The second does so. However, perhaps a little bit of interpretation would help to capture the message to be taken home Intro Second sentence, an addition such as "in many countries" or "in most countries" would be relevant. Third sentence: I know it is moving very fast yet more recent report found a much lower mortality both in regular wards and ICUs. Perhaps the authors could add that the presented figures are likely to be early high estimates and that variations over time have been observed (and are likely to be further reported) Fourth sentence: it will need to be updated just before the publication ideally I'm sorry but as of now, I think that the whole paragraph regarding hydroxychloroquine needs to be changed. Since then, many rigorous papers have shown no positive effect of HCQ against Covid-19 (https://www.clinicalmicrobiologyandinfection.com/article/S1198-

	743X(20)30505-X/fulltext) I think that the whole orientation of this paragraph should be rebuilt. This does not mean in any case that the rational for studying changing patterns as the authors did is outdated. Yet something like "at some point of time, there has been a hypothesis suggesting that HCQ could be effective…" would be more adequate with respect to the rapidly evolving state of knowledge As already said for the Abstract, the research objectives of the authors should be more finely explained (last para before Methods). What is unclear to me at this stage is whether the authors want to study compliance with (evolving) FDA recommendations or comparative outcomes of patients receiving or not HCQ I would suggest to the authors first to state as an objective the description of the evolution of chloroquine use over the beginning of the pandemic according to FDA's recommendations Discussion My comments are here again general and pertain to the purposes of the work and to their presentation I leave this decision to the editors that will be also advised by other reviewers yet I may have framed the paper that way. First objective, to describe the pattern of HCQ use over time according to FDA's positions. Second, to compare the outcomes of both cohorts (with and without HCQ) according to an observational design and with best efforts to adjust on confounding by indication (severity or others). This would lead to observe an absence of positive effect of HCQ on patients' outcome if my understanding is correct, which is consistent with the quasi-totality of rigorous reports. Third, to examine whether there is an interaction with FDA's positions (the current main focus) Otherwise, I found the paper well written and the presentation honest with inherent limitations of the observational design in particular

REVIEWER	Susan Xu Houston Methodist Research Institute, USA
REVIEW RETURNED	29-Oct-2020

GENERAL COMMENTS	This is a well-written manuscript. The statistical methods are appropriate. I have a few comments. 1. Which variables were included in the model for the propensity score matching? 2. Which variables were included in the multivariable analysis and propensity-score matched analyses? 3. After the propensity score matching, there were still statistically significant differences between HCQ and no HCQ for some variables. Did you consider them in the multivariable analysis? 4. Did you consider the clustering (propensity score matched pairs) in the modeling? 5. Did you check the interaction between the treatment and the FDA periods before the subgroup analysis? Did you check the proportional assumption for the modeling?

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1

Comments to the Author
Thank you very much for allowing me to review this manuscript of high interest. I congratulate the authors for their efforts and findings. Overall, I think that the manuscript deserves publication. However, I have some significant remarks to suggest. Most of them relate to the clarity of the work and its purposes (which were not obvious to me at first sight)
First, the abstract needs a revision. It is not even mentioned in the Objective that this is related to the COVID-19 pandemic. I know that this has become the only subject of conversation but still. I think in particular about the long term life of the article.

We agree with the reviewer and mentioned COVID-19 pandemic in the objective.

In the results (of the abstract), the total number of patients after the FDA warning seems to be missing. (ditto for the article summary).

Thank you for finding this error. We included the total number of patients in the FDA warning period.

Last, I am not sure that the first sentence of the conclusion replies to the objectives of the work. The second does so. However, perhaps a little bit of interpretation would help to capture the message to be taken home.

We combined two sentences, with an emphasis on the second sentence to relay a clear message about the study.

Intro
Second sentence, an addition such as "in many countries" or "in most countries" would be relevant.

We fixed this.

Third sentence: I know it is moving very fast yet more recent report found a much lower mortality both in regular wards and ICUs. Perhaps the authors could add that the presented figures are likely to be early high estimates and that variations over time have been observed (and are likely to be further reported).

As suggested, we edited our manuscript to reflect more recent study findings, including lower inpatient mortality in the recent period compared to beginning of the pandemic.

Fourth sentence: it will need to be updated just before the publication ideally.

We agree and will update the statistics before the publication.

I'm sorry but as of now, I think that the whole paragraph regarding hydroxychloroquine needs to be changed. Since then, many rigorous papers have shown no positive effect of HCQ against Covid-19 (https://urldefense.com/v3/__https://www.clinicalmicrobiologyandinfection.com/article/S1198-743X(20)30505-X/fulltext__;!!BWcElQ!hfXbB6XdBJtULlwpqadSDYtF2MXe3Hh6jsNTXKb-lHLTlkqlFxDR_O0poTVKfvNK7-VkypE$)
I think that the whole orientation of this paragraph should be rebuilt. This does not mean in any case that the rational for studying changing patterns ashe authors did is outdated. Yet something like "at some point of time, there has been a hypothesis suggesting that HCQ could be effective…" would be more adequate with respect to the rapidly evolving state of knowledge
As already said for the Abstract, the research objectives of the authors should be more finely explained (last para before Methods). What is unclear to me at this stage is whether the authors want to study compliance with (evolving) FDA recommendations or comparative outcomes of patients receiving or not HCQ
I would suggest to the authors first to state as an objective the description of the evolution of chloroquine use over the beginning of the pandemic according to FDA's recommendations.

We agree with the reviewer that the second paragraph of the Introduction section needs a major revision given the recent publications on the topic. We have included more recent literature on this topic, including papers on systemic review and randomized clinical trials. We also agree that the study is still unique in examining the association between hydroxychloroquine use and clinical outcomes by FDA recommendations, which has a significant interaction with changes in patient case mix.

Discussion

My comments are here again general and pertain to the purposes of the work and to their presentation

I leave this decision to the editors that will be also advised by other reviewers yet I may have framed the paper that way. First objective, to describe the pattern of HCQ use over time according to FDA's positions. Second, to compare the outcomes of both cohorts (with and without HCQ) according to an observational design and with best efforts to adjust on confounding by indication (severity or others). This would lead to observe an absence of positive effect of HCQ on patients' outcome if my understanding is correct, which is consistent with the quasi-totality of rigorous reports. Third, to examine whether there is an interaction with FDA's positions (the current main focus)

Otherwise, I found the paper well written and the presentation honest with inherent limitations of the observational design in particular.

We thank the reviewer for raising an important point. We agree that there are two objectives to the study: 1) describe the pattern of hydroxychloroquine use over time according to FDA's position and 2) compare the outcomes between patients treated with and without hydroxychloroquine. Also, although not the main objective of the study, we agree that there is a significant interaction because patients' disease severity and FDA's position. We have edited the manuscript in multiple places (abstract, introduction, and discussion) to reflect that there are two objectives to the study.

Reviewer: 2

Comments to the Author

This is a well-written manuscript. The statistical methods are appropriate. I have a few comments.

1. Which variables were included in the model for the propensity score matching?

The variables included in the propensity score matching model were age, gender, race, obesity, and insurance. Also included are binary variables for the presence of the following comorbidities by *International Statistical Classification of Disease and Related Health Problems, Tenth Revision (ICD-10)* coding: cancer, coronary artery disease, hypertension, asthma, chronic obstructive pulmonary disease, diabetes, chronic liver disease, chronic kidney disease, and end stage renal disease. Lastly, we calculated the Charlson Comorbidity Index, which is an index that predicts the 10-year survival of patients with multiple comorbidities, as a measure of total comorbidity burden. The only covariate with missing data was BMI, and we categorized the BMI group as not obese (BMI less than 30kg/m^2), obese (BMI greater than or equal to 30kg/m^2), and missing BMI.

2. Which variables were included in the multivariable analysis and propensity-score matched analyses?

We thank the reviewer for raising this important question. After the propensity score match, we ran a univariate analysis with hydroxychloroquine treatment in the analyses. We did not think it would be statistically sound to include the same variables in the statistical models that were used in the propensity score matching model. We corrected this in the manuscript.

3. After the propensity score matching, there were still statistically significant differences between HCQ and no HCQ for some variables. Did you consider them in the multivariable analysis?

We did not consider including them in the analyses because there is only one variable in one time period that had a p-value less than 0.05 (signifying a statistically significant difference). Also, the standard mean differences between HCQ and No HCQ matched groups (found in Table 2) were largely below the 0.1 threshold.

4. Did you consider the clustering (propensity score matched pairs) in the modeling?

To address clustering in the model, we considered the propensity score matched pairs within each FDA period and applied the nearest-neighbor method to create matched samples.

5. Did you check the interaction between the treatment and the FDA periods before the subgroup analysis? Did you check the proportional assumption for the modeling?

We thank the reviewer for raising this issue. We addressed any possible interactions between the treatment and the FDA periods by having individual matching models for each FDA period. The

proportional hazards assumption was met in the Cox regression model and we included this in the Results section.

VERSION 2 – REVIEW

REVIEWER	Jean Davod ZEITOUN Centre d'Epidémiologie Clinique, Hôtel Dieu Hospital, Paris None related to the authors Sharelholder of a company whose Sanofi is a customer (Sanofi being the manufacturer of HCQ in France at least)
REVIEW RETURNED	10-Dec-2020

GENERAL COMMENTS	I have reviewed the revised version and the response letter. The authors adequately replied to my comments I was happy to see that the other reviewer had proposed complementary modifications I congratulate them for their work that deserves publication to my opinion

REVIEWER	Susan Xu Houston Methodist Research Institute
REVIEW RETURNED	18-Nov-2020

GENERAL COMMENTS	Thank the authors for addressing my previous comments. However, I still have a few comments. Table 2 presents patient characteristics after propensity score matching (PSM). Firstly, the p-value presented in this table should be removed because non-significance after PSM may simply due to the reduced sample size rather than improved balance. Secondly, after PSM, quite a few covariates still had SMD>0.1 (not balanced). Did the authors check the PS model for mis-specification (such as non-linearity)? If the model specified is fine, then all the unbalanced covariates should be included in the multivariable model. Finally, the clustering (PSM pairs) should be considered in the Cox proportional hazards model.

VERSION 2 – AUTHOR RESPONSE

Reviewer: 2
Thank the authors for addressing my previous comments. However, I still have a few comments. Table 2 presents patient characteristics after propensity score matching (PSM). Firstly, the p-value presented in this table should be removed because non-significance after PSM may simply due to the reduced sample size rather than improved balance.
Thank you for the comment, we have removed the p-values in table 2.

Secondly, after PSM, quite a few covariates still had SMD>0.1 (not balanced). Did the authors check the PS model for mis-specification (such as non-linearity)? If the model specified is fine, then all the unbalanced covariates should be included in the multivariable model.
Almost all of the covariates with a SMD > 0.1 occurred in the Pre-FDA time period. In the pre-FDA time period, there was no guideline in terms of who received HCQ. Therefore, it was possible that there might have been more variability in the characteristic of the patients who received HCQ in

the pre-FDA approval period, resulting in more difficulty in appropriate matching. We constructed a graph that plotted logic(p) versus age group, where p is the probability of the HCQ (receiving HCQ), and we did not appreciate non-linearity for the age variable.

Finally, the clustering (PSM pairs) should be considered in the Cox proportional hazards model.
We found mixed literature on maintaining matched pair identities from propensity score matching in the Cox proportional hazards model. Some papers discouraged the use of stratification on the propensity score to estimate hazard ratios because it can result in biased estimation of the conditional hazard ratio.
References:

- Austin PC. The use of propensity score methods with survival or time-to-event outcomes: reporting measures of effect similar to those used in randomized experiments. Statistics in medicine. 2014 Mar 30;33(7):1242-58.
- Shinozaki T, Mansournia MA, Matsuyama Y. On hazard ratio estimators by proportional hazards models in matched-pair cohort studies. Emerging themes in epidemiology. 2017 Dec 1;14(1):6.

Reviewer: 1
I have reviewed the revised version and the response letter. The authors adequately replied to my comments
I was happy to see that the other reviewer had proposed complementary modifications
I congratulate them for their work that deserves publication to my opinion
Thank you.

VERSION 3 – REVIEW

REVIEWER	Susan Xu Houston Methodist Research Institute
REVIEW RETURNED	30-Dec-2020

GENERAL COMMENTS	Thanks the authors for addressing my previous comments. I don't have any further questions.