Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Journal of Infection and Public Health 15 (2022) 65–67

Contents lists available at ScienceDirect

Journal of Infection and Public Health

journal homepage: http://www.elsevier.com/locate/jiph

Response to the letter to the editor

Dear Editor,

We have read the "Letter to the Editor" responded to our study "Comparing ICU admission rates of mild/moderate COVID-19 patients treated with hydroxychloroquine, favipiravir, and hydroxychloroquine plus favipiravir" with careful consideration. Since the author/authors mention several points, we will clarify in detail why and how our research findings are valid based on literature to prevent further misunderstandings.

Comment 1: "Power calculation was not presented. It would have helped to evaluate the limitations and implications of the study."

Retrospective power analysis is controversial as is evidenced by several papers in the literature such as the papers of Zhang et al. [1], Reed and Blaustein [2], and Hayes and Steidl [3]. As mentioned in Zhang et al., retrospective power analysis is misleading since post-hoc power estimates vary in range and can differ from the true power. Hence, we do not present post-hoc power calculation.

On the other hand, we calculated the required sample size for the study as suggested in Powell and Sweeting [4] as the sample size should be based on the anticipated outcomes. Furthermore, as suggested by Röhrig et al. [5] required sample size can be calculated before the study using information gathered from the literature. Therefore, we used the confidence interval approach to calculate the required sample size before the study where ICU admission rate was our primary outcome. The admission rate had been reported as 9.4% to 45.9% in different studies at the planning stage of our study [6,7]. Moreover, the ICU admission rate was not uniform among different patient groups [8–11]. Since there are different ICU admission rates reported in the literature, we estimated the required sample size using Population ICU admission rate as P = 0.50 where it produces the largest sample size for estimation of the confidence interval using 95% alpha level and 5% precision for proportion where it was 385 patients overall. However, we aimed to evaluate all patient records which met inclusion criteria.

Comment 2: "The inclusion rate was too low and difficult to relate to the exclusion criteria. A flowchart depicting eligible subjects and the reason for excluding them should have been provided"

As mentioned in the article, only mild and moderate patients were included, as the severe patient group received different treatments for cytokine storm. To be able to comment on the efficacy of the drugs used, patients who died within 72 h after the admission and pregnant patients who would not be able to use the evaluated drugs were not included.

Comment 3: "Factors related to poor prognosis were more common in favipiravir treatment groups than hydroxychloroquine only group, including older age, male sex, comorbidities (diabetes mellitus, hypertension, malignancy, etc.), inflammatory markers, radiological signs of more severe pulmonary disease, and higher SOFA scores. Propensity score matching should have been used for the adjustment."

There are several different methods for adjusting the effects of confounding and/or baseline variables in the literature that are suitable when there are more than two treatment groups. Inverse Probability Treatment Weighting (IPTW) is used to balance/adjust the covariate/confounding effects to estimate the Average Treatment Effect. IPTW is especially useful where there are several confounding/covariate effects to adjust, and the researcher/researchers can't afford to lose any patients who could not be matched due to a small study population. There is a wide range of resources on the method itself since the work of Rosenbaum [12–15]. and the method is commonly used in studies related with COVID-19 [16,17] in the literature. On the other hand, we should emphasize that the multiple-group propensity score studies shouldn't be evaluated by only considering the studies on two-group propensity score methods. The authors of the letter base their criticism on such studies. It is obvious that there are three treatment groups in our study and there are a small number of subjects who required intensive care. Hence, the matching approach is no longer practical to use for covariate adjustment since we can't afford to lose any observations where the overall ICU admission rate was approximately %3. As mentioned in Leite [18], there are three methods that have the same three assumptions to estimate the unbiased effects of multiple treatments. When the assumptions are satisfied, the choice of the method depends on the study and when the propensity score model was correctly specified, the results of the matching and weighting approaches are similar [19]. Furthermore, IPTW estimates are generalizable to the entire population from which the observed sample was taken [15].

As mentioned in Yoshida et al. [20], simultaneous matching where there are multiple treatment groups is computationally challenging and often causes excluding too many patients. Furthermore, in a good covariate overlap setting, covariate balance was achieved better in IPTW than three-way matching according to Yoshida et al. [20]. However, we should note that no study evaluates the performance of IPTW in estimating Odds Ratio when there are three treatment groups. Additionally, we evaluate the validity of our results as suggested by Leite [18] by trimming the 99th percentile of the weights. Since it could be an indicator of poor common support or misspecified propensity score model [18], we evaluate if there is any substantial change in the estimations after the trimming as a part of sensitivity analysis. As a result, we showed that our results are valid since there was no substantial difference between the results after the trimming. As mentioned in Leite [18], there is no clear guideline to evaluate what magnitude of the weights are too large to consider for stabilization of the weights. We concluded that there are no extreme weights in our propensity score model as we follow the steps suggested by Leite [18]. As can be seen in findings on IPTW application of our original research, the covariate balance was evaluated using standardized difference where no covariate's value exceeded 0.10 as suggested by Leite [18].

https://doi.org/10.1016/j.jiph.2021.12.007

Comment 4: "In propensity models, standardized differences with a threshold of 10% were chosen to define imbalance. This threshold signified a wide caliper problem, which could cause substantial bias for such imbalanced data. Tighter caliper and closer matches would be preferred to modifying the propensity score, as mentioned above [2]*."*

The covariate balance can be evaluated using several diagnostic approaches when there are two treatment groups. One of the methods is evaluating standardized differences as mentioned by the author/authors of the letter in their reference to the study of Lunt [21]. However, the aforementioned study does not include the scenario where there are more than two treatment groups and therefore, it is unfortunate to generalize the study findings to include multiple group scenarios. In our study, there are three treatment groups. As mentioned in the paper of Austin [13], "although there is no universally agreed-upon criterion as to what threshold of the standardized difference can be used to indicate an important imbalance, a standard difference that is less than 0.1 has been taken to indicate a negligible difference in the mean or prevalence of a covariate between treatment groups". In addition, there are several studies using 0.1 as the maximum standardized difference to indicate poor covariate balance. However, those studies were constructed with two treatment arms, therefore, cannot be generalized for multiple groups setting. On the other hand, there are several researches on multiple groups propensity score methods where the conventional 0.10 is used as the threshold [18,22]. To the best of our knowledge, a study that provides simulation results on the optimal threshold for IPTW in multiple treatment groups doesn't exist.

Comment 5. "Analysis with trimming could have introduced selection bias."

Trimming is a method where if the method is applied correctly the first level of trimming (here, 1%) improves precision by eliminating subjects with high weights who inflate the variance of estimates as concluded by Conover et al. [23]. There are different trimming strategies where researchers should carefully consider which one is suitable for their study [24] when there are two treatment groups. On the other hand, we must inform the readers that the extension of these methods for multiple groups setting is an active research area [25].

In contrast to the author/authors, trimming could reduce bias [25] or produce similar results in certain settings [26]. Even though it is a controversial topic, we use the trimming method only to determine whether some patients effect our results *i.e.* robustness of our model.

Comment 6. "Synthetic Minority Oversampling Technique (SMOTE) is an approach for classification of data in case of an imbalance [3]*. However, the imbalance between treatment groups in that study had an impact on the outcome, but not on classification."*

As the author/authors pointed out, we use SMOTE to balance the ICU admission groups. The aim was to achieve classification (i.e. ICU admitted and not admitted) despite the imbalance in the data at a ratio of 1:40 as mentioned in our study findings: "The need for ICU follow-up was observed in 3 (0.5%) patients in HCQ group, 7 (7%) in favipiravir group, and 13 (10.8%) in HCQ plus favipiravir group" among 824 eligible patients. SMOTE is one of the several methods suggested to overcome such high imbalances and was employed to validate our model results in combination with the Bootstrap sampling technique.

Comment 7. "Sub-group analysis is a common variation of sensitivity analysis, which is relevant for this study comparing the treatment outcomes but was not performed in the analysis [4]*."*

Unfortunately, due to the high imbalance in the data, we can't conduct a sub-group analysis. However, we constructed a sensitivity analysis which aimed "assessing how sensitive the model is to fluctuations in the parameters and data on which it is built" [27] by examining whether some patients (whose weights are high in the IPTW) could be highly effective in our findings. Therefore, we used trimming as a tool to assesses the robustness of our model by comparing our preliminary findings with trimmed results.

Comment 8. "Data about the duration of hospitalization, treatment decision, and drug-related adverse events were not given. The study used patient data between March 15 and June 1, 2020. During this period Turkish Ministry of Health released four documents on the treatment algorithm and criteria for hospitalization of COVID-19 patients [5]*. Favipiravir was mainly reserved for more severe disease, which was evident in the higher frequency of more severe COVID-19 patients receiving favipiravir treatment. Thus, time matching should have been added to adjust for these differences due to changes in the treatment algorithm."*

Treatment algorithm changes in the study period are implicitly accounted for when patient's situation is considered by adjusting the effects of clinical and laboratory findings using IPTW which is also the suggested method for adjusting time-varying covariates [28]. Furthermore, a pre-analysis evaluation of the admission date revealed that it wasn't significant.

Briefly, due to the requirements of the nature of highly imbalanced data, it is necessary to adjust the effects of confounding factors and the baseline characteristics with IPTW. We employed IPTW to balance the effects of covariates and possible confounding factors, evaluated the common support (*i.e.* the covariate balance) using suggested approaches in Leite [18], and utilized SMOTE in combination with bootstrap approach for validation of the model. From a statistical standpoint, the results are valid. It must also be noted that more research is needed to evaluate the effectiveness of these treatment regimens.

Acknowledgement

The authors would like to thank Hatice Yagmur Zengin for her support in statistical analysis and evaluation both in the article and the response to the letter.

References

[1] Zhang Y, Hedo R, Rivera A, Rull R, Richardson S, Tu XM. Post hoc power analysis: is it an informative and meaningful analysis? Gen Psychiatry 2019;32(4).
[2] Reed JM, Blaustein AR. Assessment of "nondeclining" amphibian populations using power analysis. Conserv Biol 1995;9(5):1299–300.
[3] Hayes JP, Steidl RJ. Statistical power analysis and amphibian population trends. Conserv Biol 1997;11(1):273–5.
[4] Powell JT, Sweeting MJ. Retrospective studies. Eur J Vasc Endovasc Surg 2015;50(5):675.
[5] Röhrig B, du Prel JB, Wachtlin D, Kwiecien R, Blettner M. Sample size calculation in clinical trials: part 13 of a series on evaluation of scientific publications. Deutsches Ärzteblatt International 2010;107(31–32):552.
[6] Immovilli P, Morelli N, Antonucci E, Radaelli G, Barbera M, Guidetti D. COVID-19 mortality and ICU admission: the Italian experience. Crit Care 2020;24(May):1.
[7] Cunningham AC, Goh HP, Koh D. Treatment of COVID-19: old tricks for new challenges; 2020.
[8] Eventhough we do not consider those references by the time of the Project proposal we would like to give detailes on the subject fort he readers.
[9] Pijls BG, Jolani S, Atherley A, Derckx RT, Dijkstra JI, Franssen GH, et al. Demographic risk factors for COVID-19 infection, severity, ICU admission and death: a meta-analysis of 59 studies. BMJ Open 2021;11(1):e044640.
[10] Subudhi S, Verma A, Patel AB, Hardin CC, Khandekar MJ, Leite H, et al. Comparing machine learning algorithms for predicting ICU admission and mortality in COVID-19. NPJ Digit Med 2021;4(1):1–7.
[11] Lammers AJJ, Brohet RM, Theunissen REP, Koster C, Rood R, Verhagen DWM, et al. Early hydroxychloroquine but not chloroquine use reduces ICU admission in COVID-19 patients. Int J Infect Dis 2020;101:283–9.
[12] Rosenbaum PR. Model-based direct adjustment. J Am Stat 1987;82:387–94.
[13] 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.
[14] The role of the propensity score in estimating dose-response functions. In: Imbens Guido W, Guo S, Fraser MW, editors. Propensity score analysis: statistical methods and applications, vol. 11. SAGE publications; 2014.
[15] Brookhart MA, Wyss R, Layton JB, Stürmer T. Propensity score methods for confounding control in nonexperimental research. Circ Cardiovasc Qual Outcomes 2013;6(5):604–11.

[16] Shin HJ, Chow R, Noh H, Leite J, Leite J, Choi YG. The association between immunosuppressants use and COVID-19 adverse outcome: national COVID-19 cohort in South Korea. medRxiv 2021, preprint.
[17] Somers EC, Eschenauer GA, Troost JP, Golob JL, Gandhi TN, Wang L, et al. Tocilizumab for treatment of mechanically ventilated patients with COVID-19. Clin Infect Dis 2021;73(2):e445–54.
[18] Leite Walter. Practical propensity score methods using R. Sage Publications; 2016.
[19] Austin PC, Stuart EA. The performance of inverse probability of treatment weighting and full matching on the propensity score in the presence of model misspecification when estimating the effect of treatment on survival outcomes. Stat Methods Med Res 2017;26(4):1654–70.
[20] Yoshida K, Hernández-Díaz S, Solomon DH, Jackson JW, Gagne JJ, Glynn RJ, et al. Matching weights to simultaneously compare three treatment groups: comparison to three-way matching. Epidemiology (Cambridge, Mass.) 2017;28(3):387.
[21] Lunt M. Selecting an appropriate caliper can be essential for achieving good balance with propensity score matching. Am J Epidemiol 2014;179(2):226–35, http://dx.doi.org/10.1093/aje/kwt212.
[22] Yoshida K, Hernández-Díaz S, Solomon DH, Jackson JW, Gagne JJ, Glynn RJ, et al. Matching weights to simultaneously compare three treatment groups: comparison to three-way matching. Epidemiology 2017;28(3):387.
[23] Conover MM, Rothman KJ, Stürmer T, Ellis AR, Poole C, Jonsson Funk M. Propensity score trimming mitigates bias due to covariate measurement error in inverse probability of treatment weighted analyses: a plasmode simulation. Stat Med 2021;40(9):2101–12.
[24] Stürmer T, Webster-Clark M, Lund JL, Wyss R, Ellis AR, Lunt M, et al. Propensity score weighting and trimming strategies for reducing variance and Bias of treatment effect estimates: a simulation study. Am J Epidemiol 2021;190(8):1659–70.
[25] Yoshida K, Solomon DH, Haneuse S, Kim SC, Patorno E, Tedeschi SK, et al. Multinomial extension of propensity score trimming methods: a simulation study. Am J Epidemiol 2019;188(3):609–16.
[26] Luna-Bazaldua D, O'Dwyer L, July Multiple-group propensity score inverse weight trimming and its impact on covariate balance and bias in treatment effect estimation. In: The annual meeting of the psychometric society. Cham: Springer; 2019. p. 147–59.
[27] Salciccioli JD, Crutain Y, Komorowski M, Marshall DC. Sensitivity analysis and model validation. In: MIT Critical Data, editor. Secondary analysis of electronic health records [Internet]. Cham (CH): Springer; 2016., http://dx.doi.org/10.1007/978-3-319-43742-2_17 [chapter 17].
[28] 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):3661–79.

Rahmet Guner
Imran Hasanoglu*
Bircan Kayaaslan
Department of Infectious Disease and Clinical Microbiology, Ankara Yildirim Beyazit University, Ankara City Hospital, Ankara, Turkey

Adalet Aypak
Department of Infectious Disease and Clinical Microbiology, Ankara City Hospital, Ankara, Turkey

Esragul Akinci
Hurrem Bodur
Department of Infectious Disease and Clinical Microbiology, University of Health Sciences, Ankara City Hospital, Ankara, Turkey

Fatma Eser[a]
Ayse Kaya Kalem[a]
[a] *Department of Infectious Disease and Clinical Microbiology, Ankara Yildirim Beyazit University, Ankara City Hospital, Ankara, Turkey*

Orhan Kucuksahin
Department of Rheumatology, Ankara Yildirim Beyazit University, Ankara City Hospital, Ankara, Turkey

Ihsan Ates
Department of Internal Medicine, Ankara City Hospital, Ankara, Turkey

Aliye Bastug

Yasemin Tezer Tekce
Zeynep Bilgic
Fahriye Melis Gursoy
Hatice Nisa Akca
Department of Infectious Disease and Clinical Microbiology, University of Health Sciences, Ankara City Hospital, Ankara, Turkey

Seval Izdes
Department of Anaesthesiology and Reanimation Ankara, Yildirim Beyazit University, Ankara City Hospital, Ankara, Turkey

Deniz Erdem
Department of Anaesthesiology and Reanimation Ankara, Ankara City Hospital, Ankara, Turkey

Emra Asfuroglu
Department of Internal Medicine, Ankara City Hospital, Ankara, Turkey

Habibe Hezer
Department of Pulmonary Diseases, Ankara City Hospital, Ankara, Turkey

Hatice Kilic
Department of Pulmonary Diseases, Ankara Yildirim Beyazit University, Ankara City Hospital, Ankara, Turkey

Musa Civak
Department of Internal Medicine Ankara City Hospital, Ankara, Turkey

Sibel Aydogan
Department of Virology, Ankara City Hospital, Ankara, Turkey

Turan Buzgan
Department of Infectious Disease and Clinical Microbiology, Ankara Yildirim Beyazit University, Ankara City Hospital, Ankara, Turkey

* Corresponding author.
E-mail addresses: rahmetguner@yahoo.com (R. Guner), imran.solak@gmail.com (I. Hasanoglu), drbican@gmail.com (B. Kayaaslan), aadalet@yahoo.com (A. Aypak), esragulakinci@gmail.com (E. Akinci), hurrembodur@gmail.com (H. Bodur), fatmacivelekeser@hotmail.com (F. Eser), dr.aysekaya09@hotmail.com (A. Kaya Kalem), orhankcs@gmail.com (O. Kucuksahin), dr.ihsanates@hotmail.com (I. Ates), dr.aliye@yahoo.com (A. Bastug), ayasmintezer@gmail.com (Y. Tezer Tekce), zeynepunsal3860@gmail.com (Z. Bilgic), mlsgursoy@gmail.com (F.M. Gursoy), htcnskc@gmail.com (H.N. Akca), sevalizdes@yahoo.com (S. Izdes), dh2erdem@gmail.com (D. Erdem), emra.kalkan@hotmail.com (E. Asfuroglu), hoflaz@yahoo.com (H. Hezer), drhaticeb@yahoo.com (H. Kilic), drhgk@hotmail.com (M. Civak), drsaydogan72@gmail.com (S. Aydogan), turanbuzgan@yahoo.com (T. Buzgan).

25 November 2021