Short Communication

In Reply to the Discovery Trial Report Evaluating the Potential Benefit of Hydroxychloroquine, Lopinavir and Ritonavir with and Without Interferon B-1a in Hospitalized Covid-19 Patients

Valère Lounnas[1*], Xavier Azalbert[2], Christian Perronne[3]

[1]EMBL Heidelberg alumni, Meyerhofstraße 1, 69117 Heidelberg, Germany

[2]Toulouse School of Economics, 1988 Econometrics, France

[3]Service des maladies infectieuses et tropicales, University of Versailles Saint Quentin – Paris Saclay, University Hospital Raymond Poincaré, Garches, France

***Corresponding author:** Valère Lounnas, EMBL Heidelberg alumni, Meyerhofstraße 1, 69117 Heidelberg, Germany

Received: 28 May 2021; **Accepted:** 21 June 2021; **Published:** 10 August 2021

Citation: Valère Lounnas, Xavier Azalbert, Christian Perronne. In reply to the Discovery trial report evaluating the potential benefit of hydroxychloroquine, lopinavir and ritonavir with and without interferon β-1a in hospitalized Covid-19 patients. Archives of Microbiology & Immunology 5 (2021): 310-315.

Keywords: Covid-19, Discovery, randomized trial, multicenter, phase 3, hospitalized patients, hydroxychloroquine, HCQ, lopinavir, ritonavir, interferon β-1a, proportional-odds model, treatment toxicity, statistical power, treatment-timing

1. Discovery, an international multicenter Covid-19 trial aborted too early

Several points of concern arise regarding the methodology and the presented results of Discovery, the large French multicenter, randomized, controlled trial which evaluated the efficacy and safety of lopinavir + ritonavir (L/r), lopinavir/ritonavir + Interferon β-1a (L/r + IFN), and hydroxychloroquine (HCQ) in adults hospitalized for COVID-19. Results were made public on January 8, 2021 via a manuscript deposited on MedRxiv that was finally published in the CMI journal on May 25 one year after the trial was stopped [1,2]. The section "Declaration of interest" of this article is really impressive and raises questions. Sixteen investigators of the trial declare receiving multiple personal fees and/or grants from Gilead and others pharmaceutical companies including MSD, Bio Merieux, Merck, Thera

Technologies, Abbvie, BMS, J&J, Pfizer, and ViiV Healthcare, Baxter, Jansen. One prominent clinician who signs as co-last author and seats at the High Council for Public Health (HCSP), the scientific advisory board of the French government, declares he had already stopped receiving money in March 2020 when he was nominated to the board: "Y.Y. has nothing to disclose. He has been a board member receiving consultancy fees from ABBVIE, BMS, Gilead, MSD, J&J, Pfizer, and ViiV Healthcare, however all these activities have been stopped in the 03 past years".

Discovery was aborted due to a combination of reasons. Its international organization failed due to a lack of collaborative effort among EU nations and the high cost of participation (5000 euros/patient) [3]. The HCQ treatment arm was prematurely stopped on May 25th, 2020, by the trial's independent data safety and monitoring board (DSMB), at the request of the French health authority (ANSM), following the safety warning resulting from the article of Mehra et al. published in the Lancet (article that was retracted short after). The L/r arm was stopped by the DSMB for the lack of demonstrating sufficient efficacy (futility) and, short after, the L/r + IFN arm was stopped as well for excess toxicity. Sample size was thus reduced by 75% for all arms (145 patients/arm instead of the 600 initially planed) abrogating any possible statistically significant conclusion.

2. Questionable meaning of using an ordinal scale score as primary output in a Covid-19 trial

We were surprised by the primary outcome of Discovery: a WHO 7-point ordinal scale score. This primary outcome may not be adequate with measuring real survival benefit of treatment in the setting of Covid-19, especially for hospitalized patients requiring oxygen with the possibility of progression towards mechanical ventilation and death. The authors have substituted the classically used survival primary measure for a 7 item ordinal scale (1. not hospitalized, no limitation on activities; 2. not hospitalized, limitation on activities; 3. hospitalized, not requiring supplemental oxygen ; 4. hospitalized, requiring supplemental oxygen; 5. hospitalized, on non-invasive ventilation or high flow oxygen devices; 6. hospitalized, on invasive mechanical ventilation or extracorporeal membrane oxygenation (ECMO) ; 7. death) the statistical treatment of which implies the use of a proportional-odds model (a linear logistic model) [4]. It may be used in clinical research, for instance to evaluate the global benefit of cancer treatments with respect to covariates, where survival must be balanced with quality of life to palliate with treatment toxicity and discomfort. This is particularly appropriate when dealing with diseases developing over several months or years. It is also useful to determine parameters that crucially impact on the quality of survival, measured several months after intervention, like the measurement of neurological functional disability subsequent to brain damage following a stroke [5]. Death may be included as well in the ordinal scale besides other essential components in the evaluation of a complex situation. However, when only life is at stake over a time span of 2 or 3 weeks, like in Covid-19, the primary outcome should be uniquely survival because the risk is too high of misinterpreting results with a model which adds an unnecessary level of complexity.

The proportional-odds model differs radically from a Cox model adjusted across subgroups of patients with well defined specific disease characteristics for which proportional hazards ratios can be evaluated. Lack of statistical power due to too-small sample size or deviation from the hazards proportionality (chaotic or crossing curves) are defects that can be easily appraised visually. On the contrary, using the proportional-odds model implies in principle

the assessment of its likelihood. It requires a mathematical treatment, that commercial software packages do not generally provide although they fit the model [4]. This shows the non triviality of the model and the risk of being mislaid when applying it blindly in the setting of Covid-19.

3. Analysis of raw survival data of Discovery suggests L/r and HCQ potential efficacy

As a matter of fact, analyzing the results of Discovery, we note that the odds ratios for survival at 15 days with respect to standard of care (SOC), calculated from raw death data and sample sizes with the Fisher's exact test from the R software [6], are $OR_{L/r}$ = 0.72 [0.22 – 2.15] (p = 0.62) for L/r, $OR_{L/r + IFN}$ = 1.33 [0.52 – 3.49] (p = 0.67) for L/r + IFN, and OR_{HCQ} = 0.72 [0.22 – 2.15] (p = 0.62) for HCQ. Inspection of the baseline patients characteristics does not reveal major imbalances between the 4 groups SOC, L/r, L/r + IFN and HCQ, except for kidney and liver diseases present more predominantly in the HCQ (14 pts, 9.6%) and SOC (13 pts, 8.8%) arm compared with vs L/r (5 pts, 3.4%) and L/r + IFN (5 pts, 3.4%). These unbalances are due to the fact that lopinavir and ritonavir are known, although they are rare events, for inflammation of the liver and kidney failure causing fragile patients in that respect to be allocated to SOC or HCQ. Despite this, the unadjusted survival ratios can serve fairly as comparison because classical comorbidities are evenly distributed across treatment groups. They suggest a possible benefit for HCQ and L/r treatment, although statistical power is dramatically lacking, forbidding any conclusion on treatment effect on survival.

In contrast, the OR ratios, produced by the proportional-odds model: OR=0.83 [0.55 to 1.26] (p=0.39) for L/r, (Table 2 of ref. 1 and 2), OR=0.69 [0.45 to 1.04] (p=0.08) for L/r + IFN and OR=0.93 [0.62 to 1.41] (p=0.75) for HCQ are presented in a way that may cast confusion among readers who are not familiar with the proportional-odds model interpretation. In the results section, the authors state: “adjusted OR (aOR) were not in favor of investigational treatments (i.e., < 1.)”. In this case, adjusted ORs from the proportional-odds model must be understood as the odds of having 1 point deterioration in the outcome ordinal scale compared with SOC, irrespective of the specific score values considered. This is an overall estimate measure that mixes survival with overall patient condition. Here, it is consistent with the fact that the L/r + IFN arm was stopped for excess toxicity with dramatic impact on survival (13 deaths at 15 days twice as much as for HCQ) and the overall patient clinical condition.

In fact, the proportional-odds modeling of the 7-item scale has brought to the fore, with near statistical significance, a very detrimental effect for L/r + IFN with 3 deaths imputed to the experimental treatment, but less detrimental for L/r (1 treatment-related death) and certainly not for HCQ (0 treatment-related death). However, due to the amalgamation of all scores in one single signal, it conceals the potential direct survival benefit of HCQ suggested by the classical odds ratios. The authors do not discuss the meaning and limitations of the ORs produced by the model they used, possibly installing some confusion in their interpretation. Neglecting the lack of statistical power, they take as granted that the proportional-odds model provides a good estimate of the reality of the trial. From this, they conclude to an absence of potential survival benefit of HCQ. To substantiate their assertion, they refer to the trial of Cavalcanti et al. [14] where the same 7-item ordinal scale was used as primary outcome at 15 days. That trial was performed in the setting of hospitalized patients, with mild to moderate diseases, only 40% of which required

oxygen support. HCQ and HCQ plus azithromycin (AZI) were the evaluated treatments compared with SOC. Results with the 7-item ordinal scale revealed also a distortion between classically calculated hazard ratios (HR) and the odds ratio derived from the proportional-odds model. Thus, for HCQ + AZI a HR = 0.64 [0.18 to 2.21] was obtained in contrast with the 7-item ordinal scale derived OR = 0.99 [0.57 to 1.73] (no effect) while, for HCQ the result was consistent between HR = 1.47 [0.48 to 4.53] and the 7-item ordinal scale OR = 1.21 [0.69 to 2.11]. (with OR > 1. values meaning here a detrimental effect of the treatment) Additionally and more crippling was that the Calvalcanti's trial, like Discovery, dramatically lacked power, with too few patient inclusions: 70 pts/arm.

4. Secondary outcome analysis

Classical survival ORs were calculated at 28 days as a secondary outcome of the trial whereas it should have been the primary outcome. We find a very good agreement between the reported ORs (Table 2 of ref 2) and those that can be inferred from the unadjusted raw data with $OR_{L/r}$ = 1.19 [0.49 – 2.92] (p = 0.69) for L/r, $OR_{L/r + IFN}$ = 1.44 [0.62 – 3.44] (p = 0.43) for L/r + IFN, and OR_{HCQ} = 0.94 [0.36 – 2.40] (p = 1.). These values shows the deleterious treatment effect of the L/r and L/r + IFN treatment and, contrary to the values recorded at 15 days, suggest an absence of treatment effect for HCQ.

The curves reporting the evolution of the viral load in the lower respiratory track (LRT) within 15 days after inclusion show a continuous decrease without difference with respect to SOC for any of the tested treatments. This reveal that patients were included at quite an advanced stage where the viral load starts to get cleared from the body. It is in contrast with the article published by Gautret et al. on the viral load decrease (measured by PCR in nasopharyngeal swabs) in patients with predominantly upper respiratory track symptoms (28/36 pts) [7].

5. Absence of HCQ toxicity compared with SOC and the lack of statistical power indicate HCQ may remain potentially a valid treatment

Despite all their trial limitations, the authors of Discovery firmly state in their discussion of the preprint manuscript that: “these investigational treatments failed to improve the clinical course of COVID-19, nor to enhance SARS-CoV-2 clearance and were associated with more Serious Adverse Events (SAEs)” [1] which is in contradiction with the interpretation of the adjusted OR for the 7-item ordinal scale suggesting HCQ was not associated with a deterioration of the patient overall condition. The supposed high toxicity of HCQ has been the moto of many clinicians during the year 2020. However, this misleading assertion was corrected in the published version of the manuscript. This inaccuracy of the MedRxiv manuscript was particularly annoying since it was presented as the official report of the trial. Indeed, the numbers of Adverse Events (AEs) and SAEs per arm reported in Table 3 of their manuscript show a statistically significant increase of AE and SAEs for both L/r and L/r + IFN ($p < 0.05$) compared with SOC, but not for HCQ (p= 0.35). For any grade 3 and 4 SAEs in the HCQ arm there is absolutely no difference with SOC (p=0.71). This added to the unadjusted survival OR_{HCQ} = 0.70 [8] for HCQ suggests that statistical significance of a potential benefit of HCQ at the dosage of the trial (400 mg twice on day 1 and then 400 mg once daily for 9 days) may have been reached provided accrual had not been interrupted due to political reasons.

6. Optimal antiviral efficacy requires early patient management

Thirdly, we note generally that in Discovery, as in many other trials, patients were treated at a quite advanced stage, at the end of the viral primary phase, possibly just before or presumably at the beginning of the acute inflammatory phase (cytokines storm). Regarding HCQ, the authors of Discovery changed their discussion, between the MedRxiv version and the published version in CMI, acknowledging that "the target plasma concentration was reached in only 25% of participants at day 3, and optimal intra-pulmonary exposure might have been only achieved at day 10 [10][15]". This must be put in perspective with the fact that "The median time from symptoms onset to randomization was 9 days (IQR, 7-12)" indicating that placement of patients to HCQ treatment occurred too late. Furthermore, no indication is provided on the median time from hospitalization to randomization in Discovery.

Hospitalized patients had an oxygen saturation of less than 94%, with symptoms of fluid penetrating small airways (rales and crackles), which indicates pulmonary damage. However, for optimal antiviral efficacy, it is necessary to treat as early as possible, i.e. before pulmonary involvement requiring hospitalization, at the time when the virus begins to multiply [7]. Detailed analysis of the trials conducted between March and June 2020 shows that poorly established disease severity, inhomogeneous treatment protocols and patient's selection, as well as variable delay between hospitalization and treatment initiation, in addition to ill-defined treatment dosage and length, has forbidden any clear conclusion on antiviral treatments real efficacy [9,10]. At a quite late stage, where antiviral treatment effect may not to be optimal, the patient survival depends crucially on other synergistic factors such as comedication and timing of cares [11,12,13]. At an even more advanced stage, the potential positive effect of HCQ may become essentially anti-inflammatory since it down-regulates massive cytokines production [12]. We note that when solely severe disease cases are taken into account at 28 days OR_{HCQ} = 0.74 [0.17 – 2.92] (p = 0.76) ($OR_{L/r}$ = 1.50, $OR_{L/r + IFN}$ = 1.84).

As for the level of virus in the nasopharyngeal secretions, at the end of the viral multiplication phase, the virus level has reached a maximum, which then gradually decreases until the patient may no longer have detectable virus. As the virus is hardly multiplying at the beginning of oxygen desaturated, no real difference was observed between any of the experimental treatment and SOC. It cannot be ruled out that the natural virus decaying process dominates here. But this does not mean that no effect on the viral load would have been observed at an earlier stage of the disease as observed by Gautret et al. [7].

7. Conclusion

In short, poor inclusion criteria as well as intra- and inter-trial variability in treatment timing and patient management may explain why the effects of a wide spectrum anti-viral drugs such as HCQ are difficult to demonstrate. In Discovery, we note that at D15, there is a difference in mortality between the hydroxychloroquine arm and the SOC arm, and also at D28 for patients with severe disease, although not statistically significant due to a dramatic lack of power. The authors of Discovery claim that their trial had some limitations as it was not a placebo-controlled, double-blinded trial but they avoid to mention that inclusion of patients in the HCQ arm was stopped too hastily after the publication of the article of Mehra et al. in the Lancet, unduly abrogating the possibility to measure

the potential effect of HCQ. Additionally, "the trial did not target patients at the early phase of the disease, nor included anti-inflammatory agents. Being considered all this, a significant difference in survival could have emerged for HCQ if the trial had been conducted according to the statistical plan.

8. Conflict of interest: none

9. Acknowledgment: association bonsens.org covers the publication fees.

References

1. Ader F, Peiffer-Smadja N, Poissy J, Bouscambert-Duchamp M, Belhadi D, Diallo A, et al. Antiviral drugs in hospitalized patients with COVID-19 - the DisCoVeRy trial medRxiv (2021) 20248149.
2. Ader F, Peiffer-Smadja N, Poissy J, Bouscambert-Duchamp M, Belhadi D, Diallo A, et al. An open-label randomized, controlled trial of the effect of lopinavir/ritonavir, lopinavir/ritonavir plus IFN-β-1a and hydroxychloroquine in hospitalized patients with COVID-19 CMI, Published (2021).
3. Coronavirus : l'essai clinique Discovery englué faute de coopération européenne », Le Monde.fr, 7 mai (2020).
4. Peter McCullag "The proportional-odds model : Theoretical Background" Whiley (2014).
5. Goyal M, Demchuk AM, Menon BK, Eesa M, Rempel JL, Thornton J, et al. Randomized assessment of rapid endovascular treatment of ischemic stroke. N Engl J Med 372 (2015): 1019-30.
6. The R Project for Statistical Computing.
7. Gautret P, Lagier JC, Parola P, Hoang VT, Meddeb L, Mailhe M, et al. Hydroxychloroquine and azithromycin as a treatment of COVID-19: results of an open-label non-randomized clinical trial. Int J Antimicrob Agents. 56 (2020): 105949.
8. Steven Tenny; Mary R. Hoffman. Odds Ratio In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing (2021).
9. Lounnas V, Lacout A, Azalbert X, Perronne C. Revisiting a Meta-analysis Show that Hydroxychloroquine with Azithromycin may be Efficient in Covid-19 patients. Archives of Microbiology & Immunology 5 (2021): 154-175.
10. Lounnas V, Lacout A, Azalbert X, Perronne C. Shunt due to Hydroxychloroquine Sub-lethal Dosage Resulted in Excess Transfer to Mechanical Ventilation and Death in Hospitalized Patients with Covid-19. Archives of Microbiology & Immunology 5 (2021): 176-181.
11. Bousquet G, Falgarone G, Deutsch D, et al. ADL-dependency, D-Dimers, LDH and absence of anticoagulation are independently associated with one-month mortality in older in patients with Covid-19. Aging (Albany NY) 12 (2020): 11306-11313.
12. Lagier JC, Million M, Gautret P, et al. Outcomes of 3,737 COVID-19 patients treated with hydroxychloroquine/azithromycin and other regimens in Marseille, France: A retrospective analysis. Travel Medicine and Infectious Disease 36 (2020): 101791.
13. Lacout A, Lounnas V, Perronne C. Timing and dosage may be the key in the realisation of hydroxychloroquine + azithromycin treatment benefit in Covid-19 elderly patients. Int J Antimicrob Agents 57 (2021):106314.
14. Cavalcanti AB, Zampieri FG, Rosa RG et al. Hydroxychloroquine with or without Azithromycin in Mild-to-Moderate Covid-19. N Engl J Med 383 (2020): 2041-2052.