PEER REVIEW HISTORY

BMJ Medicine publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

TITLE (PROVISIONAL)	COVID-19 convalescent plasma to treat hospitalised COVID-19 patients with or without underlying immunodeficiency: a randomized trial
AUTHORS	Lacombe, Karine; Hueso, Thomas; Porcher, Raphael; Mekinian, Arsene; Chiarabini, Thibault; Georgin-Lavialle, Sophie; Ader, Florence; Saison, Julien; Martin-Blondet, Guillaume; De Castro, Nathalie A; Bonnet, Fabrice; Cazanave, Charles; Francois, Anne; Morel, Pascal; Hermine, Olivier; Pourcher, Valerie; Michel, Marc; Lescure, Xavier; Soussi, Nora; Brun, Phillipe; Pommeret, Fanny; Sellier, Pierre; Rousset, Stella; Piroth, Lionel; Michot, Jean-Marie; baron, gabriel; de Lamballerie, Xavier; Mariette, Xavier; Tharaux, Pierre-Louis; Resche-Rigon, Matthieu; Ravaud, Philippe; Simon, Tabassome; Tiberghien, Pierre

VERSION 1 - REVIEW

REVIEWER 1	Casedevall, Arturo. Competing Interest: I received funds to support in part a study of convalescent plasma that was published in the NEJM (PMID: 35353960).
REVIEW RETURNED	15-Nov-2022

GENERAL COMMENTS	This paper reports the outcome of Bayesian randomized clinical trial evaluating the efficacy of high volume and high titer COVID-19 convalescent plasma (CCP) in a population that included a high percentage of immunocompromised patients. The results show reduced mortality in the CCP treated group, specially in the immunocompromised patient subset. This is an important paper given that the latest omicron variants have now defeated all the monoclonal antibodies that were critical therapeutics for immunocompromised patients. Hence, the results that CCP reduces mortality in immunocompromised is welcomed news. The study is innovative in that that the patients received the relatively large dose of 4 units of CCP in two days, which would have provided a larger quantity of specific antibody as well as the biological diversity of polyclonal responses in several units but possibly also contributed to the greater pulmonary problems observed in the treatment group. The major limitation of the study is its small to moderate size. This reviewer has no major criticisms of the study and a few suggestions for improvement. 1. How long after randomization was CCP administered? This information is important to include since all evidence suggests that efficacy drops with time. 2. The finding that there were 4 cases of acute pulmonary edema in the CCP group and none in the usual case suggest that the administration of the large volume of plasma could have tipped these patients. Can the authors clarify whether these four cases received the full intended doses – e.g. 4 units? 3. As a follow up to point 2, the early respiratory worsening

	observed in the CCP group is different than reported in the almost three dozen RCTs for CCP in the literature making this study an outlier in this regard. This study differs from these other RCTs in using a protocol of 4 units over 2 days. This dose may have been too much over such short time. This reviewer thinks these issues should be addressed more clearly in the discussion and perhaps a take home lesson is that if 4 units are to be given maybe these should be spaced out over more days. Minor points 4. The wording in the first line of the discussion that CCP 'failed to show a better efficacy' is not accurate and should be reworded to state that no difference was found. In fact, the sentence is internally inconsistent as it initially states that it failed to show better efficacy but ends by pointing out that those treated had lower fewer deaths, which de facto suggests greater efficacy. 5. Page 15 line 18 states that the RECOVERY trial 'have found no evidence of survival benefit with CCP' is not quite correct. Several subgroups showed reduced mortality albeit closely missing significance at 0.05. RECOVERY was a very problematic study with 9% of patients in the CCP not receiving plasma. 6. Page 9, line 28. Omicron not micron 7. The manuscript states in several places that overall efficacy of CCP has not been established in hospitalized patients. However, that conclusion is reasonable only if one considers only RCTs data. Given how problematic the RCT design has been for CCP (late treatment in most, low titers in some, etc) a case can be made that real world data provides better evidence and the large HCA study in the United States reported a large drop in mortality in those treated with CCP (PMID: 34464352). Similarly, the large USA registry study reported a reduction in mortality. Hence, the authors may want to modify these statements or state that they reflect RCT evidence only.

REVIEWER 2	Besson, Caroline. Competing Interest: None
REVIEW RETURNED	27-Nov-2022

GENERAL COMMENTS	This paper is very clear and well-written. It confirms the lack of efficacy of COVID-19 convalescent plasma to treat hospitalised COVID-19 patients. However, it also highlights the efficiency of this treatment in immunocompromised patients.

REVIEWER 3	Malagnino, Vincenzo; University of Rome Tor Vergata, Medicin of Systems. Competing Interest: None
REVIEW RETURNED	29-Nov-2022

GENERAL COMMENTS	The work of Lacombe et all reports on the experience of a French multicentre cohort on the use of convalescent plasma in SARS-COV 2 positive patients. Although the data available on the use of this type of treatment have not produced results in terms of mortality, even in this work, the group clearly defines the possible role of this treatment within the decision-making algorithm of therapy for COVID19. The trend of the pandemic, with the advent of variants and subvariants, some with immune-escape mutations, makes the use of convalescent plasma interesting, also given the data on the alternating efficacy of oral antivirals and monoclonal antibodies. The setting of immunocompromised patients, in particular

	haematological patients, is currently the most vulnerable both in terms of clinical outcome and interruption of diagnostic-therapeutic procedures. In this sense, the definition of additional therapeutic lines that can perhaps support the drugs available. The work is extremely well defined, the methodology is accurate and linear and the contents are clear, despite the complex and poorly defined strategy in clinical practice. I consider the paper suitable for publication in the journal

REVIEWER 4	Estcourt, Lise. Competing Interest: I declare that I know and work as part of an EU collaboration with one of the authors
REVIEW RETURNED	01-Dec-2022

GENERAL COMMENTS	This is a manuscript describing a randomised controlled trial of convalescent plasma adminisistration in patients who required hospitalisation for COVID-19 compared to usual care, It was conducted between April 2020 to April 2021. The authors need to be congratulated for completing a well conducted trial during the pandemic which included a large proportion of participants who had immunocompromise. The manuscript would benefit from some clarifications. 1. It would be useful to know whether the majority of participants would have wild type or alpha variant. Looking at the recruitment graph the majority of participants were recruited from November 2020 onwards. Did the study test participants, and if not which were the main circulating variants in the country at the time of the recruitment as a guide to the likely underlying variant. 2. The authors have described the mean Euroimmun per unit as 6.1 with an SD of 2.9 for the convalescent plasma. The authors have not described a minimum titre threshold. Can the authors please clarify the minimum threshold used if there was a cut-off. Also, the distribution of titres is likely to be a skewed distribution, can the authors instead report the median, IQR, and minimum titre used. 3. Are the authors able to provide more detailed baseline characteristics describing the immunocompromised subgroup within the appendix as this was the group that was shown to benefit from convalescent plasma. 4. Were any of the adverse events that occurred in the convalescent plasma arm reported to the French Health Products Safety Agency (AFSSAPS)? If so please clarify 5. Can the authors explain why in a similar group of participants (RECOVERY trial) there did not appear to be an increase in respiratory issues with the CP group compared to control? Could this be due to higher neutralising activity within all plasma units in the RECOVERY trial - minimum titre was Euroimmun of 6 in the RECOVERY trial. In the Canadian trial (Begin et al) plasma with a lower titre led to harm. Or could it be related to time since infection that donors were able to donate plasma. This was 15 days in France but at least 28 days in UK? Minor issues 1. P6 line 28 - micron rather than omicron 2. P7 line 11 - unnecessary word early within the sentence

REVIEWER 5	Calmy, Alexandra; Hôpitaux Universitaires Genève, infectious diseases. Competing Interest: Research Funds: MSD (Why women aren't accessing clinical trials) Unrestricted Educational Funds (to the Institution): MSD, AbbVie, Gilead Sciences, ViV Healthcare
REVIEW RETURNED	05-Dec-2022

GENERAL COMMENTS	The investigators conducted a open-label randomised clinical trial assessing the efficacy of 4 units of convalescent plasma (CCP) over 2 days (versus usual care alone) on the proportion of patients with a WHO-Clinical Progression score at Day 4 and on Survival at Day 14. The clinical trial has been designed in order to also answer the question of efficacy among the most fragile populations of immuno-suppressed patients. We agree with the investigators that the antibody-based therapy has raised interest in the new era of immune-resistant SARS COV-2 variants. 1. Method: - Inclusion period from April 2020 to April 2021 - was the one year inclusion period related to feasability issues with CCP? - Statistical analysis: can you provide some comments on the your choice of the Bayesian framework to analyse the primary endpoint? This method generally allows for an update based on the most recent data (for example with a prior distribution informed by Recovery?). In addition, in the abstract the term "CI" is used but unclear whether this refers to Confidence Interval or Credibility Interval. - You do not mention if you could assess low vs high dose of CCP (antibody potency in transfused CCP). 2. Results: - 50 up to 60 patients randomized in the intervention arm did receive the full intervention of 4 CCP. Please comment on the feasbility issue (physician's fear of adverse event? Difficulties to access CCP?) - Safety: It would be interesting to have more information of the causality of the SAE (for example 4 acute pulmonary oedema were reported in the CCP arm, and none in the UC. (I can't see imputability in table S10?). 3. Discussion - Such an intervention is theoretically interesting, however I would advise to soften the conclusion based on the possible risks (pulmonary oedema). - Discussion on more concentrated CCP - with less frequent transfusion may be worth mentioning (less fluids, one donor). Finally I find the publication of these finding important after a careful review regarding the above-mentioned concerns.

REVIEWER 6	Riley, Richard; University of Birmingham, Institute of Applied Health Research. Competing Interest: None
REVIEW RETURNED	20-Jan-2023

GENERAL COMMENTS	A well-conducted trial that has clearly been a major undertaking by the authors. Also good to see a Bayesian approach to analysis.

	A major limitation is the small sample size which is reflected in the generally wide credible intervals, and this makes strong inferences problematic. However, it is still important to publish trials in this situation, regardless, as they can be important toward meta-analyses in the future. Why is the continuous (ordinal) outcome of WHO-CPS dichotomised into a binary outcome using WHO-CPS ≥6? Similarly, why was duration of symptoms modelled as a binary variable using cut-point of 5, rather than as a continuous variable (with potential non-linear trend)? It would seem to me that the analysis would be both more powerful and more interpretable using the original scales. The analysis appears to not always account for clustering of patients with centers – can the authors clarify? The tables provide p-values (I think), but these are Bayesian analyses, so please clarify. A major finding is that there is a stronger effect in the subgroup of immunocompromised individuals, with HR 0.37 [95%CI 0.14-0.97]). However, there are many subgroup analyses done, at multiple days, and events and sample sizes are sparse, therefore I am not confident in this finding. Moreover, the findings are only shown in supp material RTable S8, and there it is clear that there is much uncertainty in the interaction (difference in subgroup results) for those with and without immunodeficiency. The p-value for the interaction is 0.11. The authors should use day 14 instead of d14, and likewise for other days. In the main paper, we should have a table showing results for every day considered, including the full and subgroup results. At the moment I am scrambling to collate all these from the main paper and the supp material. Is there an issue of censoring, where some individuals are lost to follow-up? How was this handled? Figure 2a provides two HRs, one for day 14 and one for day 28. However, HRs can relate to the whole time-period, so how are these derived? Is one day 0 to 14, and the other day 14 to 28? Given that multiple time-points are of interest, should the analyses have been undertaken using a longitudinal model that allows for repeated measures per individual? Figures 3b and 3c need to express the uncertainty in the results, either by adding a CI around the lines, or adding the HR estimate with a CI underneath. The authors must re-write to avoid language about ‘trend toward …’ – as this is not scientific. Simply state that all the findings are uncertain due to the small sample size, but could be useful in meta-analyses moving forwards. I hope these comments are constructive and helpful to the authors moving forward.

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1
Comments to the Author
This paper reports the outcome of Bayesian randomized clinical trial evaluating the efficacy of high volume and high titer COVID-19 convalescent plasma (CCP) in a population that included a high percentage of immunocompromised patients. The results show reduced mortality in the CCP treated group, specially in the immunocompromised patient subset. This is an important paper given that the latest omicron variants have now defeated all the monoclonal antibodies that were critical therapeutics for immunocompromised patients. Hence, the results that CCP reduces mortality in immunocompromised is welcomed news. The study is innovative in that that the patients received the relatively large dose of 4 units of CCP in two days, which would have provided a larger quantity of specific antibody as well as the biological diversity of polyclonal responses in several units but possibly also contributed to the greater pulmonary problems observed in the treatment group. The major limitation of the study is its small to moderate size. This reviewer has no major criticisms of the study and a few suggestions for improvement.

1. How long after randomization was CCP administered? This information is important to include since all evidence suggests that efficacy drops with time.

CCP was administered within one day of randomization in 46 patients, the next day for 12 patients and two days after randomization for 1 patient. One patient was transferred to ICU after randomization and before receiving CCP. Given the fact that a large majority received CCP within 24 hours after randomization, it is unlikely that the results may have been influenced by a drop in efficacy.

2. The finding that there were 4 cases of acute pulmonary edema in the CCP group and none in the usual case suggest that the administration of the large volume of plasma could have tipped these patients. Can the authors clarify whether these four cases received the full intended doses – e.g. 4 units?

Among the 4 cases of acute pulmonary edema in the CCP group, 1 received 4 units and 3 received only 2 because of worsening health conditions. We have added this information P 11 of the manuscript. We think that Covid-19 is likely the main cause for the aggravation. Given the small number of events, it is difficult to draw any conclusion regarding a possible statistical difference between both arms with regard to those side effects.

3. As a follow up to point 2, the early respiratory worsening observed in the CCP group is different than reported in the almost three dozen RCTs for CCP in the literature making this study an outlier in this regard. This study differs from these other RCTs in using a protocol of 4 units over 2 days. This dose may have been too much over such short time. This reviewer thinks these issues should be addressed more clearly in the discussion and perhaps a take home lesson is that if 4 units are to be given maybe these should be spaced out over more days.

We agree with the reviewer and have made the following modifications in the text:
“The transfusion of 4 units of plasma may have contributed to circulatory overload in some patients. Further spacing of CCP administration (i.e. 1 unit / day over 4 days) could reduce such a risk”. P 14 of the manuscript.

Minor points

4. The wording in the first line of the discussion that CCP ‘failed to show a better efficacy’ is not accurate and should be reworded to state that no difference was found. In fact, the sentence is internally inconsistent as it initially states that it failed to show better efficacy but ends by pointing out that those treated had lower fewer deaths, which de facto suggests greater efficacy.

We agree with the reviewer and have modified the sentence according to his suggestion. P 13 of the manuscript

5. Page 15 line 18 states that the RECOVERY trial ‘have found no evidence of survival benefit with CCP’ is not quite correct. Several subgroups showed reduced mortality albeit closely missing significance at 0.05. RECOVERY was a very problematic study with 9% of patients in the CCP not receiving plasma.

We agree with that remark and have slightly changed our wording « Of note, pre-specified sub groups in the RECOVERY trial such as less than 7 days of symptoms or no oxygen showed reduced mortality with CCP albeit missing significance at 0.05." P 13 of the manuscript

6. Page 9, line 28. omicron not Omicron

Thank you for highlighting this mistake, now changed throughout in the manuscript, except in references where it was used that way by authors.

7. The manuscript states in several places that overall efficacy of CCP has not been established in hospitalized patients. However, that conclusion is reasonable only if one considers only RCTs data. Given how problematic the RCT design has been for CCP (late treatment in most, low titers in some, etc) a case can be made that real world data provides better evidence and the large HCA study in the United States reported a large drop in mortality in those treated with CCP (PMID: 34464352). Similarly, the large USA registry study reported a reduction in mortality. Hence, the authors may want to modify these statements or state that they reflect RCT evidence only.

We thank the reviewer for highlighting the studies by Egloff et al in JCI as well the USA registry study (as reported by Joyner et al). Both of these studies are now mentioned P 13 of the manuscript.
We have reworded the text to highlight those studies: "Furthermore, large retrospective studies in the US reported evidence of reduced mortality associated with CCP treatment in patients hospitalized with COVID-19 (Joyner et al, NEJM, 2021, Egloff et al, JCI, 2021).

Reviewer: 2
Comments to the Author:
This paper is very clear and well-written. It confirms the lack of efficacy of COVID-19 convalescent plasma to treat hospitalised COVID-19 patients. However, it also highlights the efficiency of this treatment in immunocompromised patients.
We thank the reviewer for her comment.
Reviewer: 3
The work of Lacombe et all reports on the experience of a French multicentre cohort on the use of convalescent plasma in SARS-COV 2 positive patients. Although the data available on the use of this type of treatment have not produced results in terms of mortality, even in this work, the group clearly defines the possible role of this treatment within the decision-making algorithm of therapy for COVID19. The trend of the pandemic, with the advent of variants and subvariants, some with immune-escape mutations, makes the use of convalescent plasma interesting, also given the data on the alternating efficacy of oral antivirals and monoclonal antibodies. The setting of immunocompromised patients, in particular haematological patients, is currently the most vulnerable both in terms of clinical outcome and interruption of diagnostic-therapeutic procedures. In this sense, the definition of additional therapeutic lines that can perhaps support the drugs available. The work is extremely well defined, the methodology is accurate and linear and the contents are clear, despite the complex and poorly defined strategy in clinical practice. I consider the paper suitable for publication in the journal
We thank the reviewer for his comment.

Reviewer: 4
This is a manuscript describing a randomised controlled trial of convalescent plasma adminisistration in patients who required hospitalisation for COVID-19 compared to usual care, It was conducted between April 2020 to April 2021.
The authors need to be congratulated for completing a well conducted trial during the pandemic which included a large proportion of participants who had immunocompromise.
We thank the reviewer for her comment.
The manuscript would benefit from some clarifications.
1. It would be useful to know whether the majority of participants would have wild type or alpha variant. Looking at the recruitment graph the majority of participants were recruited from November 2020 onwards. Did the study test participants, and if not which were the main circulating variants in the country at the time of the recruitment as a guide to the likely underlying variant.

Some participants were tested for SARS-COV2 genotype and some others were not, depending on the center's procedures, as it was not a requirement for the trial. However, at the time of recruitment, circulating variants were first the ancestral strain from Wuhan then alpha strain.

2. **The authors have described the mean Euroimmun per unit as 6.1 with an SD of 2.9 for the convalescent plasma. The authors have not described a minimum titre threshold. Can the authors please clarify the minimum threshold used if there was a cut-off. Also, the distribution of titres is likely to be a skewed distribution, can the authors instead report the median, IQR, and minimum titre used.**

In fact PCC were qualified on the basis of a seroneutralisation titer >= 1/40. Among the qualified PCC the mean ELISA (Euroommun IgG) ratio was 6.1 with an SD of 2.9, an IQR of 5.4, min. of 0.4 and maximum of 13.0. The text P 6 has been amended for clarification and provision of the ELISA IQR, min and max. ratio.

3. Are the authors able to provide more detailed baseline characteristics describing the immunocompromised subgroup within the appendix as this was the group that was shown to benefit from convalescent plasma.

We thank the reviewer for this suggestion. A table describing baseline characteristics of the immunocompromised participants has been added to the appendix.

4. Were any of the adverse events that occurred in the convalescent plasma arm reported to the French Health Products Safety Agency (AFSSAPS)? If so please clarify.

All adverse events associated with CCP transfusion (probable to certain imputabiliy) were reported to ANSM (new name for AFFSAPS) as per current regulation regarding blood components in France (P 6 of the manuscript). Additionally, all serious adverse events recorded in study eCRF (table S10) were reported to ANSM as well as per current regulation regarding clinical trials in France.

5. Can the authors explain why in a similar group of participants (RECOVERY trial) there did not appear to be an increase in respiratory issues with the CP group compared to control? Could this be due to higher neutralising activity within all plasma units in the RECOVERY trial - minimum titre was Euroimmun of 6 in the RECOVERY trial. In the Canadian trial (Begin et al) plasma with a lower titre led to harm. Or could it be related to time since infection that donors were able to donate plasma. This was 15 days in France but at least 28 days in UK?

We do not have an explanation for the apparent discrepancies regarding possible early respiratory issues after PCC administration. As highlighted by Dr Casadevall, the volume of plasma transfused was larger (800-880 ml) than most other studies. As mentioned as well in the discussion, early pulmonary worsening in a limited number of patients has been reported in another study involving PCC (ref 27; Avendano-Sola et al, JCI). Furthermore our study had an early primary endpoint (day 4) which may have resulted in a very close patient monitoring early on. Lastly, and as mentioned in the discussion P 14, distinguishing spontaneous COVID-19 worsening from circulatory overload or CCP-induced inflammatory antibody dependent enhancement (ADE) type reaction may be challenging. The possibility that low titer CCP may be playing a role is intriguing. However, if we are not mistaken, the harm reported in Begin et al does not relate to early, possibly transient, pulmonary worsening. Furthermore, in our study the 4 plasmas transfused in each patient was provided by different donors, thus probability that a patient may have received more than 1 plasma with a low titer (out of 4) was low. Lastly, while there was a trend towards early pulmonary worsening (disappearing by day 12) in patients treated by CCP vs controls, this potential association remained not statistically significant and therefore warrants further studies.

Minor issues

P6 line 28 - micron rather than omicron 2. P7 line 11 - unnecessary word early within the sentence

This has been changed.

Reviewer: 5

Comments to the Author The investigators conducted a open-label randomised clinical trial assessing the efficacy of 4 units of convalescent plasma (CCP) over 2 days (versus usual care alone) on the proportion of patients with a WHO-Clinical Progression score at Day 4 and on Survival at Day 14. The clinical trial has been designed in order to also answer the question of efficacy among the most fragile populations of immuno-suppressed patients.

We agree with the investigators that the antibody-based therapy has raised interest in the new era of immune-resistant SARS COV-2 variants.

1. Method:

- Inclusion period from April 2020 to April 2021 - was the one year inclusion period related to feasability issues with CCP?

The context in France at that period was that of a competition between trials: a very high number of trials in Covid19 had been initiated within a year and it has been difficult to include the required number of patients in a shorter time. Besides, the recruitement has been impacted by the successive waves with periods where few patients have been hospitalized.

- Statistical analysis: can you provide some comments on the your choice of the Bayesian framework to analyse the primary endpoint? This method generally allows for an update based on the most recent data (for example with a prior distribution informed by Recovery?). In addition, in the abstract the term "CI" is used but unclear whether this refers to Confidence Interval or Credibility Interval.

The choice of a Bayesian analysis for primary outcomes was initially made for the entire CORIMUNO19 platform, for two specific aims that actually match what the reviewer mentions. First, it was use to allow flexible and informative monitoring of each trial, by providing the DSMB with different posterior probabilities (probability of any benefit and probability of moderate or greater benefit) for early stopping for efficacy or futility, or recommending trail continuation or extension. We were convinced (and still are) that Bayesian statistics are a good way to convey such information for decision making. Second, it was also planned that the results of previous CORIMUNO19 trials on similar drugs could be used to inform the prior distributions for a specific trial. This was not the case for CORIPLASM, since we had no prior information on convalescent plasma efficacy within CORIMUNO19, but also outside CORIMUNO19 (inclusions started on April 16, 2020, while no randomized evidence was public; in particular inclusions in the RECOVERY convalescent trial started on May 28, 2020 and those in PLACID on April 22, 2020).

Regarding the acronym “CI” in the abstract, we apologise for any confusion. We have made it clear and used “CrI” for credible intervals and “CI” for confidence intervals (for frequentist analyses).

- You do not mention if you could assess low vs high dose of CCP (antibody potency in transfused CCP).

We report P12 (data Table S9) that post-hoc analysis of antibody potency in transfused CCP in relation with to outcome did not reveal a significant dose-effect. As mentioned in the discussion P 13, the requirement in our study that the 4 CCP transfused by provided by different donors resulted in limited variation in mean Ab content from patient to patient (to the difference of the CAPSID trial for example). This difference may have contributed to reducing the ability to identify an Ab dose effect in our study.

2. Results:

- 50 up to 60 patients randomized in the intervention arm did receive the full intervention of 4 CCP. Please comment on the feasbility issue (physician's fear of adverse event? Difficulties to access CCP?)

There were 3 reasons for not receiving 4 units: randomized but worsening of health condition leading to ICU before transfusion (1), 2 units as the beginning of the trial as per protocol (in order to test for safety (3), worsening of health condition leading to ICU transfer after transfusion of 2 units. As discussed in the discussion section, we cannot differenciate a potential deleterious effect of CPP from a worsening due to Covid19. However the clinicians’ perception was that the patient was already deteriorating where CCP was transfused.

- Safety: It would be interesting to have more information of the causality of the SAE (for example 4 acute pulmonary oedema were reported in the CCP arm, and none in the UC. (I can't see imputability in table S10?).

In table S10, the number in brackets for all figures under the heading **Type of SAE** reports the number of events imputable to CCP according to the investigator. We have amended table S10 for increased clarity.

3. Discussion

- Such an intervention is theoretically interesting, however I would advise to soften the conclusion based on the possible risks (pulmonary oedema).

We have reworded the conclusion and introduced a more balanced discussion around the risk and benefist of such a therapeutic, given the results in terms of efficacy especially in iummunocompromised patients and gobal safety with the risk of pulmonary oedema.

- Discussion on more concentrated CCP - with less frequent transfusion may be worth mentioning (less fluids, one donor).

We are not sure to understand the reviewer's comment : does she refer to the possibility to reconstitute units from concentrated plasma ? this is technologically challenging. Plus we have decided to mix donors in order to have a better mean concentration of antibodies, given the challenge of getting enough donors during the pandemic (dealing with lockdowns and confinments).

Finally I find the publication of these finding important after a careful review regarding the above-mentioned concerns.

We thank the reviewer for her insightful comments.

Reviewer: 6

A well-conducted trial that has clearly been a major undertaking by the authors. Also good to see a Bayesian approach to analysis.

A major limitation is the small sample size which is reflected in the generally wide credible intervals, and this makes strong inferences problematic. However, it is still important to publish trials in this situation, regardless, as they can be important toward meta-analyses in the future.

Why is the continuous (ordinal) outcome of WHO-CPS dichotomised into a binary outcome using WHO-CPS ≥6? Similarly, why was duration of symptoms modelled as a binary variable using cut-point of 5, rather than as a continuous variable (with potential non-linear trend)? It would seem to me that the analysis would be both more powerful and more interpretable using the original scales.

We thank the reviewer for raising that point. The reason for dichotomising the ordinal WHO-CPS scale for primary outcomes was two-fold. The first reason was that when planning the CORIMUNO19 trials at the end of March 2020, the WHO scales were not well-known. Multiple versions existed, with larger or smaller numbers of levels, and physicians had no experience in those, especially as a clinical trial outcome. There were debates as to whether we should use of scale with more levels (thus providing more insights into the clinical status of patients) or with less levels (with more confidence for low—out of the hospital—and high scores, the latter depending on PaO2/FiO2 ratios that may be variable depending e.g. on switching to a supine to prone position of patients). The choice of the most precise scale was made, but investigators also preferred that clear cut changes in clinical status and use of resources would be used for primary outcomes. For instance, a WHO-CPS score of 6 or more denotes

the need for at least a high-flow oxygen device, which implied a significant worsening in patient's status and using more resources. If a treatment was able to prevent worsening with a score of 6 or more, this would have both clinical and practical implications which would make the treatment more valuable, compared to a treatment leading to lower scores on average, but mostly driven by scores below 4 or between 7 and 8, for instance. However, the scores were also analysed on their original scale, both a prespecified timepoints and as longitudinal data. We agree that those analyses may be more powerful (and there was some work by Frank Harrell advocating such analyses in that particular context, but this came out approximately at the time the CORIMUNO19 platform was designed and we only became aware of it after the start of some trials of the platform). We still preferred to keep the design as it was to favour interpretability by investigators over statistical properties. Last, in their work comparing several COVID-19 outcome measures, Dodd et al. showed that time-to-event analyses methods have advantages in the COVID-19 setting, and were also well-suited for interim analyses, despite reducing an ordinal scale to a binary state (https://doi.org/10.48550/arXiv.2006.10533).
We apologise if there was a misunderstanding concerning the duration of symptom. The duration of symptoms below or larger than 5 was predefined for subgroup analyses. We agree that studying an interaction with possibly non-linear terms could have been a better choice, but we did not plan such analysis. Please note that longer symptom durations, as reported by the patients, may be less reliable.

The analysis appears to not always account for clustering of patients with centers – can the authors clarify?

We apologise if this was unclear. Actually, all analyses were adjusted for centre (with random effects) and age, the absolute risk difference in early primary outcome, which was simply analysed using a beta-binomial model. This was because at time of trial planning simulations of the operational characteristics of the design were undertaken using this easy to implement model. We have clarified in the revised manuscript that the Bayesian proportional odds model was also adjusted for age and centre.

The tables provide p-values (I think), but these are Bayesian analyses, so please clarify.

Sorry for the lack of clarity. The only p-values reported are for interaction tests of subgroup analyses, and for analyses of safety. As described in the Statistical Analysis Plan (but we acknowledge that this information is buried in the 35 pages of protocol plus 17 pages of SAP), those analyses were planned to be frequentist, as were some analyses of secondary outcomes.

A major finding is that there is a stronger effect in the subgroup of immunocompromised individuals, with HR 0.37 [95%CI 0.14-0.97]). However, there are many subgroup analyses done, at multiple days, and events and sample sizes are sparse, therefore I am not confident in this finding. Moreover, the findings are only shown in supp material RTable S8, and there it is clear that there is much uncertainty in the interaction (difference in subgroup results) for those with and without immunodeficiency. The p-value for the interaction is 0.11.

We totally agree that this result is of low certainty for the reasons you mentioned (many subgroups analysis, non significant interaction). As suggested by reviewer 4, we wanted to describe the characteristics of immunocompromised patients. Unfortunately, we realized at that time that we had made a computer glitch in our definition of immunocompromised patients (3 patients had been wrongly classified as immunocompromised and 1 patient should have been). Also, we started the analysis again after correcting this error. The estimate of the new subgroup analysis in immunocompromised patients is very close to the previous result but the confidence interval now

contains the value 1. We have corrected results in the paper and removed the focus on this subgroup in the abstract. Again, we apologize for this mistake.

The authors should use day 14 instead of d14, and likewise for other days.

We have changed the wording throughout the text.

In the main paper, we should have a table showing results for every day considered, including the full and subgroup results. At the moment I am scrambling to collate all these from the main paper and the supp material.

We think that this table may be very heavy, we have on purpose chosen to include the main results pertaining to primary outcomes in the paper and all other results in the supplementary material. However, if the editor thinks that including all the results in the main paper would read better, we can prepare this table.

Is there an issue of censoring, where some individuals are lost to follow-up? How was this handled?

No patient was lost to follow-up before day 14, and five (two and three in each group, respectively) were lost of follow-up at day 28. For time-to-event outcomes (overall survival, time to discharge and time to oxygen supply independency), those were handled using methods for censored data. Other outcomes were counted up to day 14.

Figure 2a provides two HRs, one for day 14 and one for day 28. However, HRs can relate to the whole time-period, so how are these derived? Is one day 0 to 14, and the other day 14 to 28?

The reviewer raises a good point, and we apologise for the lack of clarity. The HR for day 14 was for day 0 to 14, and the one for day 28 for day 0 to 28.

Given that multiple time-points are of interest, should the analyses have been undertaken using a longitudinal model that allows for repeated measures per individual?

This is what was also done for WHO-CPS over time, with a Bayesian longitudinal proportional odds models, as suggested by Frank Harrell in his (https://hbiostat.org/proj/covid19/statdesign.html#longitudinal-ordinal-outcome).

Figures 3b and 3c need to express the uncertainty in the results, either by adding a CI around the lines, or adding the HR estimate with a CI underneath.

We agree with the reviewer that displaying uncertainty is important, and we added HR estimates and CIs, as well as pointwise confidence bars, as suggested.

The authors must re-write to avoid language about 'trend toward …' – as this is not scientific. Simply state that all the findings are uncertain due to the small sample size, but could be useful in meta-analyses moving forwards.

We have deleted all mention to a "trend" throughout the paper and reworded the text to make it clear that differences were not significant.

VERSION 2 – REVIEW

REVIEWER 6	Riley, Richard; University of Birmingham, Institute of Applied Health Research. Competing Interest: None
REVIEW RETURNED	13-Jun-2023

GENERAL COMMENTS	I thank the authors for their clear and good response. It is evident that they have worked hard, not just to conduct this trial during a very challenging period, but also to deal with the many comments received. It is refreshing to see a Bayesian approach with posterior probabilities, examining impact of priors, and so forth. It is also good to see their honesty and transparency about the updated/corrected analysis of the interaction of the treatment effect with the subgroup of immunocompromised individuals, which has widened the confidence interval. I think this will be an excellent addition to the journal, but some important comments remain to be considered 1) The authors continue to focus on a dichotomous WHO scale outcome and dichotomise duration of symptoms. I personally don't see the rationale for this based on sticking to the protocol if it means the analysis methods are not as powerful as they could be. I also

	disagree that the analysis is more clinically interpretable, as noted in their response, as the cut-points can still be used after the more powerful analysis – plus then the actual cut-point is not enforced. However, it is good to see that the WHO ordinal scale was analysed using a Bayesian proportional odds models in other analysis, on the full scale comparing the distribution of ordinal scores at day 4, 7 and 14, adjusted for age and centre. Further, to deal with repeated measures, a longitudinal version of the model with a time effect and a random subject effect to analyse all scores up to day is now included. So these more full analyses are available for the reader – which I can accept as a compromise. Though duration of symptoms is still dichotomised. 2) What I do disagree with is the proportional odds model being labelled a 'secondary outcome' in the results – this is not fair, as it relates to a better analysis of the primary outcome, and it should be mentioned and discussed in the primary outcome section. Table 2 should also list this in the primary outcome section. I also do not see the proportional odds model results discussed at all in the results section of the paper. 3) The authors responses are not all translated into the revision. Please can the authors address this? For example, I asked them about how the HRs for 'day 14' and 'day 28' – which they clarify correspond to 0-14 days and 14-28 days periods, but this is not explained in the paper that I can find. To clarify, these HRs relate to different periods, not specific time-points per se. Although I question whether the periods are 0 to 14, and 0 to 28? 4) I still think a table showing the results for all days would be informative – the non primary days could be labelled as such (indeed table 2 already has some of the multiple time points included). 5) I had not spotted this before but "Time to discharge and time to oxygen supply independency were analysed in a competing risks framework using Fine-Gray models adjusted for age and centre (as a random effect), death being the competing event" – this produces sub-distribution hazard ratios rather than cause-specific hazard ratios – is the former not harder to interpret? Surely we want cause-specific hazard ratios, don't we? 6) "We computed that the trial would have a frequentist power of 97.2% to detect a decrease in event rate from 0.50 to 0.20, and 73.9% to detect a decrease in event rates from 0.50 to 0.30" – the word 'rate' is confusing in the context of time to event data, I think the authors mean event risk (or event proportion) here? 7) I think the main statistical analysis section (in the methods) should also explain how interactions were examined, as this is not currently mentioned 8) The sensitivity analyses are not mentioned in the results section (like impact of choice of priors, which is good addition) 9) "the rate of a WHO-CPS ≥6 at day 4 was not 21 statistically different in the CCP arm compared to the UC arm' – rogue 21 in the sentence? 10) "At day 28, there was a better survival of patients with underlying immunodeficiency who received CCP compared to those with UC, although there was no evidence of an interaction between immunodeficiency status and treatment (p=0.34)" – this is a non-significant finding for a secondary outcome, so I do not think the 'better survival' argument should be made. 11) Abstract and elsewhere define whether higher CPS values are good or bad. 12) "In CORIMUNO-CORIPLASM trial, no difference was found in terms of mortality between CCP and usual care for hospitalised

	COVID-19 patients not requiring assisted ventilation, with an improved survival in the CCP arm observed as early as day 14 and confirmed at day 28, without reaching a level of significance." This first sentence in the discussion is hard to read – is it saying there is a benefit or is it saying there is no benefit? Suggest rewrite more clearly – I would also emphasise that the trial is small with wide CIs and thus to conclude for sure about relationships is not sensible from this study alone. 13) For the same reason, the authors need to be very careful in the 'what this study adds' which currently seems to be strong. They say 'In this multicentre randomised clinical trial, high titre CCP in a population hospitalised with a mild to moderate form of COVID-19 within 9 days of symptoms onset did not prevent disease progression or death assessed at day 4 and day 14 after treatment.' – but isn't the CI very wide? 14) 'However, in the subgroup of patients with immunosuppression, there was a lower odds of death 28 days after CCP transfusion, albeit not reaching statistical significance.' – day 4 is the main outcome, yet here the authors push a non-significant interaction at 28 days. I really think this is inappropriate. They even then say 'The result of study, along with the recent data obtained from other trials and cohort studies may support the use of CCP in patients with underlying immunosuppression for whom therapeutic options are currently scarce' 15) 'CCP-associated early respiratory worsening as well as CCP-associated reduced day 14 and day 28 mortality were observed' – how are these findings compatible? It is associated with respiratory worsening, and yet associated with reduced mortality? Sorry to be raising points and issues for clarity again, as I appreciate the authors hard work, but I am working hard to try to get the right interpretations here. I hope the authors find it constructive.

VERSION 2 – AUTHOR RESPONSE

Reviewer 6

I thank the authors for their clear and good response. It is evident that they have worked hard, not just to conduct this trial during a very challenging period, but also to deal with the many comments received. It is refreshing to see a Bayesian approach with posterior probabilities, examining impact of priors, and so forth. It is also good to see their honesty and transparency about the updated/corrected analysis of the interaction of the treatment effect with the subgroup of immunocompromised individuals, which has widened the confidence interval. I think this will be an excellent addition to the journal, but some important comments remain to be considered.

We thank the reviewer for his encouraging comments, and for his thorough and thoughtful review of our revised manuscript. We have tried to address all specific points, as explained below.

1) **The authors continue to focus on a dichotomous WHO scale outcome and dichotomise duration of symptoms. I personally don't see the rationale for this based on sticking to the protocol if it means the analysis methods are not as powerful as they could be. I also disagree that the analysis is more clinically interpretable, as noted in their response, as the cut-points can still be used after the more powerful analysis – plus then the actual cut-point is not enforced. However, it is good to see that the WHO ordinal scale was analysed using a Bayesian proportional odds models in other analysis, on the full scale comparing the distribution of ordinal scores at day 4, 7 and 14, adjusted for age and centre. Further, to deal with repeated measures, a longitudinal version of the model with a time effect and a random subject effect to analyse all scores up to day is now included. So these more full**

analyses are available for the reader – which I can accept as a compromise. Though duration of symptoms is still dichotomized.

We thank the reviewer for summarising those points. We understand the personal view of the reviewer on the need to stick to the protocol or not, if a more powerful analysis can be conducted, but we have tried also to avoid criticisms related to post-hoc changes that may be considered ad-hoc changes. Given the results of the non-dichotomised outcomes are fully presented, we considered that the readers would be provided with the overall evidence. Of note the analysis of the WHO scale as a continuous outcome, including longitudinal proportional odds model were part of the original plans. Concerning the dichotomized duration of symptoms, we may have lacked clarity. The only dichotomized outcome was the WHO scale. We only dichotomized duration of symptoms at 5 days for performing subgroup analyses (page 7, 1st paragraph); we apologise if this was unclear.

2) **What I do disagree with is the proportional odds model being labelled a 'secondary outcome' in the results – this is not fair, as it relates to a better analysis of the primary outcome, and it should be mentioned and discussed in the primary outcome section. Table 2 should also list this in the primary outcome section. I also do not see the proportional odds model results discussed at all in the results section of the paper.**

We perhaps sticked too close to the protocol, which labelled the dichotomised score at day 4 as primary outcome. We understand the reviewer's point that the analysis of the day 4 score with a proportional odds model is actually a better analysis of the same outcome. But on the other hand, we do not want to be accused of post-hoc changes in primary outcomes compared to the protocol and the record on clinicaltrials.gov, which would be interpreted as a marker of poor trial reporting. As a matter of compromise, we have kept the reporting of day 4 continuous scores under "Secondary outcomes" in the table, but expanded the description of those results in the results, under "primary outcomes". We hope this will be considered as acceptable by the reviewer. Since day 7 and day 14 scores, as well as the longitudinal scores, were secondary outcomes, we would prefer keeping those labelled as such. We apologise for having describe the proportional odds models too briefly and have and have better discussed those results in the revised manuscript. In particular, as mentioned above, we now describe the analysis of continuous day 4 scores under "primary outcomes".

"Analysis of the day 4 WHO-CPS score was analysed as an ordinal outcome in a proportional odds model, yielded a median posterior adjusted odds ratio of 1·42 (95% CrI 0·70–2·91), therefore showing higher scores in the CCP arm, although the difference was not significant."

3) **The authors responses are not all translated into the revision. Please can the authors address this? For example, I asked them about how the HRs for 'day 14' and 'day 28' – which they clarify correspond to 0-14 days and 14-28 days periods, but this is not explained in the paper that I can find. To clarify, these HRs relate to different periods, not specific time-points per se. Although I question whether the periods are 0 to 14, and 0 to 28?**

We apologise for failing to report all changes in the manuscript. Those changes have been now made explicit in the revised manuscript. Of note, the HR for "day 28" mortality was the HR for the period day 1 to day 28, as we explained in our previous answers to the reviewer. Sorry if this was unclear. The reviewer is therefore right in questioning whether the periods were day 1–14 and 1–28.

The revised text now reads:

"Secondary endpoints included WHO-CPS at 4, 7 and 14 days after randomization, overall survival at 14 and 28 days after randomisation (i.e. for the periods from day 1 to day 14 and from day 1 to day 28, respectively)."

4) **I still think a table showing the results for all days would be informative – the non primary days could be labelled as such (indeed table 2 already has some of the multiple time points included).**

We had decided to comprehensively present the results for all available days in a figure (figure 2c) rather than a table. But we also understand that the figure does not provide odds ratios, for instance. We have therefore added the results at day 2, day 3, day 5, and day 6 to the supplementary table S6, with a note clearly identifying which analyses were predefined, and which were added at a reviewer's request. The table is reproduced below with changes highlighted in yellow.

Table S6. WHO-CPS scores during follow-up.
Odds ratios (ORs) were obtained from Bayesian proportional odds models adjusted for age and centre. For longitudinal data, time was used as a main effect in the model, and the model was also adjusted on the baseline WHO-CPS score. Missing values for patients discharged were imputed at a score 3, and a window of plus/minus 2 days was used for day 14 scores. For longitudinal analyses, only missing values on the day after discharge were imputed at a score 3; subsequent missing values were not imputed. An OR < 1 indicates efficacy of convalescent plasma compared to usual care. CrI: credible interval. The timepoints predefined in the protocol were day 4, day 7 and day 14. To provide more information, results have been added at the other available timepoints following the request of one reviewer.

	Convalescent plasma (N=60)		**Usual care (N=60)**		
	N	**Median (IQR)**	**N**	**Median (IQR)**	**Adjusted OR (95% CrI)**
Day 2	60	5 (5 to 5)	60	5 (5 to 5)	1.86 (0.71 to 5.22)
Day 3	60	5 (5 to 5)	60	5 (4 to 5)	2.99 (1.40 to 6.80)
Day 4*	60	5 (5 to 5)	60	5 (4 to 5)	1.42 (0.70 to 2.91)
Day 5	59	5 (4 to 5)	60	5 (4 to 5)	1.32 (0.67 to 2.63)
Day 6	59	5 (4 to 5)	59	5 (4 to 5)	1.08 (0.54 to 2.14)
Day 7*	58	5 (4 to 5)	59	5 (4 to 5)	1.20 (0.61 to 2.37)
Day 14*	59	3 (2 to 4)	59	3 (2 to 5)	0.59 (0.30 to 1.13)
Longitudinal analysis*	60	—	60	—	1.04 (0.37 to 2.86)

IQR: inter-quartile range; OR: odds ratio; CrI: credible interval. * Predefined analysis in the protocol.

5) I had not spotted this before but "Time to discharge and time to oxygen supply independency were analysed in a competing risks framework using Fine-Gray models adjusted for age and centre (as a random effect), death being the competing event" – this produces sub-distribution hazard ratios rather than cause-specific hazard ratios – is the former not harder to interpret? Surely we want cause-specific hazard ratios, don't we?

The reviewer is right in underlying that we estimated subdistribution hazard ratios for those outcomes, and we thank him for raising the interesting question of subdistribution vs. cause-specific hazard analysis in clinical trials. While we would generally tend to agree with the reviewer, we have considered that in the present case a Fine-Gray analysis would be preferable. Indeed we considered that the proportion of discharge alive and oxygen independency over time (cumulative incidence) would be a more clinically relevant than the rate at which those events would occur. Given the one-one relationship between the subdistribution hazard ratio and the cumulative incidence (while the relationship between

the cause-specific hazard and the cumulative incidence also involves the cause-specific hazard of competing risks), our choice was a subdistribution analysis. We also based our decision on previous work of one co-author of the manuscript (Resche-Rigon, Azoulay, Chevret. Crit Care. 2006 Feb;10(1):R5).

6) **"We computed that the trial would have a frequentist power of 97.2% to detect a decrease in event rate from 0.50 to 0.20, and 73.9% to detect a decrease in event rates from 0.50 to 0.30" – the word 'rate' is confusing in the context of time to event data, I think the authors mean event risk (or event proportion) here?**

We apologise for the wording. The reviewer is right, we meant event proportion and used "rate" as in "response rate" commonly used in cancer studies, which is improper. This has been corrected in the revised manuscript.

The revised text now reads:
"We computed that the trial would have a frequentist power of 97.2% to detect a decrease in event proportion from 0.50 to 0.20, and 73.9% to detect a decrease in event proportions from 0.50 to 0.30."

7) **I think the main statistical analysis section (in the methods) should also explain how interactions were examined, as this is not currently mentioned.**

The use of interaction tests is now mentioned in the revised manuscript.

The revised text now reads:
"Interaction tests between the treatment group and subgroups were used to tests for treatment effect heterogeneity between subgroups."

8) **The sensitivity analyses are not mentioned in the results section (like impact of choice of priors, which is good addition)**

Thank you for your appraisal of those sensitivity analyses. They are now mentioned in the results section.

The revised text now reads:
"Results remained highly consistent across the range of prior distributions used in sensitivity analyses (figure S3)."

9) **"the rate of a WHO-CPS ≥6 at day 4 was not 21 statistically different in the CCP arm compared to the UC arm' – rogue 21 in the sentence?**

We apologise for this type, which has been corrected in the revised manuscript.

The revised text now reads:
"the rate of a WHO-CPS ≥6 at day 4 was not statistically different in the CCP arm compared to the UC arm"

10) **"At day 28, there was a better survival of patients with underlying immunodeficiency who received CCP compared to those with UC, although there was no evidence of an interaction between immunodeficiency status and treatment (p=0.34)" – this is a non-significant finding for a secondary outcome, so I do not think the 'better survival' argument should be made.**

In accordance with the reviewer, we have reworded the results at day 28 (page 13 1st paragraph)

11) **Abstract and elsewhere define whether higher CPS values are good or bad.**

We have now clarified in the revised manuscript that higher WHO-CPS score values indicated a worse outcome.

Abstract:
“Primary outcome was proportion of patients with a WHO-Clinical Progression Score (CPS) ≥6 on the 10-point scale on day (d) 4 (higher values indicating a worse outcome) and survival without ventilation or additional immunomodulatory treatment by day 14.”

Manuscript
“As in all CORIMUNO19 nested trials, there was an early primary endpoint defined as a WHO Clinical Progression Scale (WHO-CPS) score ≥ 6 (Supp. material, Appendix I) at day 4 of randomisation, higher values of the WHO-CPS indicating a worse outcome.”

12) **“In CORIMUNO-CORIPLASM trial, no difference was found in terms of mortality between CCP and usual care for hospitalised COVID-19 patients not requiring assisted ventilation, with an improved survival in the CCP arm observed as early as day 14 and confirmed at day 28, without reaching a level of significance.” This first sentence in the discussion is hard to read – is it saying there is a benefit or is it saying there is no benefit? Suggest rewrite more clearly – I would also emphasise that the trial is small with wide CIs and thus to conclude for sure about relationships is not sensible from this study alone.**

in accordance withe the reviewer’s suggestion, we have rephrased this paragraph.

Now it reads : “ In CORIMUNO-CORIPLASM trial, no difference was found in terms of early outcome between CCP and usual care for hospitalised COVID-19 patients not requiring assisted ventilation. An improved survival in the CCP arm was observed as early as day 14 and again at day 28, however without reaching significance. “ page 13, 1st paragraph of discussion.

13) **For the same reason, the authors need to be very careful in the ‘what this study adds’ which currently seems to be strong. They say ‘In this multicentre randomised clinical trial, high titre CCP in a population hospitalised with a mild to moderate form of COVID-19 within 9 days of symptoms onset did not prevent disease progression or death assessed at day 4 and day 14 after treatment.’ – but isn’t the CI very wide?**

We agree that the wide CI hinder making firm conclusions. The sentence has been amended to reflect this limitation.
Now it reads: “In these patients with underlying immunosuppression, CCP treatment was associated with improved survival. However, a large confidence interval as well as limitations owing to a subgroup analysis prevent any firm conclusion regarding this association. “ page 14, 2nd paragraph

14) **'However, in the subgroup of patients with immunosuppression, there was a lower odds of death 28 days after CCP transfusion, albeit not reaching statistical significance.’ – day 4 is the main outcome, yet here the authors push a non-significant interaction at 28 days. I really think this is inappropriate. They even then say ‘The result of study, along with the recent data obtained from other trials and cohort studies may support the use of CCP in patients with underlying immunosuppression for whom therapeutic options are currently scarce’**

We agree that those results were overstated. Of note, there were two primary outcomes: one at day 4 (mostly for trial monitoring and interim analyses), and one at day 14. But we recognise that overall survival was a secondary outcome, and he interaction non-significant. This part of the manuscript has been toned down to better reflect the state of analyses.

15) **‘CCP-associated early respiratory worsening as well as CCP-associated reduced day 14 and day 28 mortality were observed’ – how are these findings compatible? It is associated with respiratory worsening, and yet associated with reduced mortality?**

We thank the reviewer for raising this point. As shown on the figure 2c (but also 2a and 2b), it seems that participants in the CCP group initially worsened, compared to those in the usual care group, up to about 5 days. Then, this seemed to reverse. Although the differences in survival were not significant,

and remained limited in absolute value, at about 8% between day 13 and 28, mortality was in favour of CCP. Such early worsening have been reported with the use of CCP, maybe due to a transient pulmonary oedema that did not lead to death but rather later to recovery, explaining the discrepancy between early and mid-term outcomes.

Sorry to be raising points and issues for clarity again, as I appreciate the authors hard work, but I am working hard to try to get the right interpretations here. I hope the authors find it constructive.

We thank again the reviewer for all those constructive comments, and hope we have satisfactorily clarified or corrected all remaining issues.

VERSION 3 - REVIEW

REVIEWER 1	Riley, Richard; University of Birmingham, Institute of Applied Health Research. Competing Interest: None
REVIEW RETURNED	29-Jul-2023

GENERAL COMMENTS	I thank the authors for their clear response. My only remaining comment is for the paper are: 1) to clear state that sub-distribution hazard ratios are presented, and to add an explanation of what these mean (and why they differ to cause-specific hazard ratios), as I think the reader may otherwise confuse them with 'normal' hazard ratios. 2) To re-word the conclusion that 'CCP did not improve early outcomes in patients with moderate COVID-19.' Because the CIs are all wide, and so actually we need further research don't we? We cannot be sure of anything based on this trial alone, I think. 3) To re-word the comment that 'In the subgroup of patients with immunosuppression, there was evidence suggesting a lower odds of death 14 and 28 days after CCP transfusion, albeit without reaching significance.' – this continues to be overstating something for which there is no clear evidence for.

VERSION 3 – AUTHOR RESPONSE

Reviewer 6

I thank the authors for their clear response. My only remaining comment is for the paper are:

1) to clear state that sub-distribution hazard ratios are presented, and to add an explanation of what these mean (and why they differ to cause-specific hazard ratios), as I think the reader may otherwise confuse them with 'normal' hazard ratios.

We thank the reviewer and have followed the advice to add details on subdistribution hazards modeling, and modified the manuscript accordingly.

In the methods section page 7 of the manuscript, it now read :

« For time to discharge and time to oxygen supply independency, we estimated adjusted sub distribution hazard ratios (SHRs) using Fine-Gray models, death being the competing event. Estimating SHRs was preferred over cause-specific HRs because they have a one-to-one relationship with the cumulative incidence, i.e. the proportion, of events, and we considered they would therefore be more relevant that the ratio of rates at which those events occur in time.

2) To re-word the conclusion that 'CCP did not improve early outcomes in patients with moderate COVID-19.' Because the CIs are all wide, and so actually we need further research don't we? We cannot be sure of anything based on this trial alone, I think.

According to the editor's suggestion, we have deleted this sentence.

3) To re-word the comment that 'In the subgroup of patients with immunosuppression, there was evidence suggesting a lower odds of death 14 and 28 days after CCP transfusion, albeit without reaching significance.' – this continues to be overstating something for which there is no clear evidence for.

This sentence has been deleted.