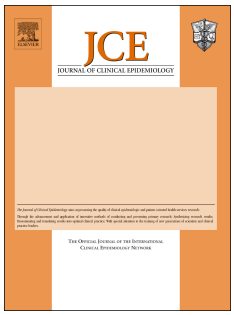
Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Journal Pre-proof



The relationship between pragmatism, timing, and study size on impact of treatment trials: a qualitative, hypothesis generating study of systemic corticosteroids for COVID-19

Aileen Liang, Katrina Domenica Cirone, Xiaoxiao (Daisy) Deng, Merrick Zwarenstein, MBBCh, PhD

PII: S0895-4356(22)00239-6

DOI: https://doi.org/10.1016/j.jclinepi.2022.09.018

Reference: JCE 10921

To appear in: *Journal of Clinical Epidemiology*

Received Date: 5 January 2022

Revised Date: 15 September 2022

Accepted Date: 30 September 2022

Please cite this article as: Liang A, Cirone KD, Deng X(D), Zwarenstein M, The relationship between pragmatism, timing, and study size on impact of treatment trials: a qualitative, hypothesis generating study of systemic corticosteroids for COVID-19, *Journal of Clinical Epidemiology* (2022), doi: https://doi.org/10.1016/j.jclinepi.2022.09.018.

The relationship between pragmatism, timing, and study size on impact of treatment trials: a qualitative, hypothesis generating study of systemic corticosteroids for COVID-19

Aileen Liang[1*]; Katrina Domenica Cirone[1*]; Xiaoxiao (Daisy) Deng[1*]; Merrick Zwarenstein [2,3 **], MBBCh, PhD

[1]*Medical students,' Schulich School of Medicine and Dentistry, The University of Western Ontario, London, ON, Canada.*

[2]*Department of Family Medicine, Centre for Studies in Family Medicine, Schulich School of Medicine & Dentistry, The University of Western Ontario, London, ON, Canada.*

[3]*Department of Epidemiology and Biostatistics, Schulich School of Medicine and Dentistry, The University of Western Ontario, London, ON, Canada.*

**These authors contributed equally to this work*

***Corresponding author:* Merrick Zwarenstein, Department of Family Medicine, Centre for Studies in Family Medicine, Schulich School of Medicine & Dentistry, Western University, 1151 Richmond Street, London, Ontario N6A 3K7, Canada. Electronic address: merrick.zwarenstein@ices.on.ca. Phone: 1-647-284-5052

Abstract

Objective — To explore qualitatively the relationship between selected trial design choices and proxies for scientific and clinical uptake in a cohort of published randomized controlled trials (RCTs) of corticosteroids for COVID-19, to identify design characteristics that may result in trials with potential to eliminate equipoise, achieve uptake and help reduce research waste.

Study Design & Setting — A systematic literature search and qualitative, narrative review of published RCTs (up to April 13, 2021) evaluating the effectiveness of systemic corticosteroids in treatment of COVID-19. We extracted information on sample size, number of centres, single or multi-country conduct, dates of initiation and of publication, risk of bias and pragmatism scores, and also on impact measured by citation in scientific literature and in clinical guidelines. We qualitatively compared design features of the highest impact versus other trials.

Results— RECOVERY was by the most impactful of the seven eligible RCTs as it was 10 times more frequently cited in peer-reviewed literature and influenced all the selected COVID-19 treatment guidelines. All trials started recruiting from similar dates. RECOVERY was a single country, multi-centre platform trial at low risk of bias, features which also fail to distinguish it from the other trials. RECOVERY was distinguished by more strongly pragmatic design features, more centres, and more rapid recruitment resulting in a larger sample size, and early publication.

Conclusion — Higher pragmatism scores may contribute to recruiting more centres and more rapid recruitment of patients at each centre, leading to larger size, earlier publication, and greater scientific and guideline uptake. By eliminating equipoise RECOVERY rendered other simultaneous trials redundant. Further work is needed to confirm these findings in a larger quantitative study and to identify the individual contribution of each characteristic of pragmatism to conduct and impact of trials, and their interaction in different national contexts. Until then

research waste might be reduced by designing trials with as many of the characteristics of RECOVERY as is feasible.

Keywords — Corticosteroids, COVID-19, PRECIS-2, Randomized controlled trial

Abstract Word Count: 317

What is new?

- RECOVERY is a large multicenter, single country, platform randomized trial of several treatments repurposed for COVID-19 care. The sub-trial of corticosteroids had a low risk of bias and highly pragmatic design features that facilitated wide implementation and rapid recruitment. This trial was cited ten times more often than the next most cited trial and relied on in all the prominent guidelines we reviewed, changing clinical practice globally. It eliminated equipoise, rendering redundant the other simultaneous trials of steroids that lacked one or more of these features. This research waste should be reduced.
- Large, pragmatic, unbiassed, single country platform trials of repurposed drugs and interventions, covering different potential pandemic conditions and multiple treatments, for a range of sociodemographic situations and healthcare capacities could be a valuable investment in readiness for future pandemics, resulting in trials with greater scientific and clinical impact and less research waste.
- Generic protocols for trials aimed at each kind of pandemic threat could be prepared, in advance, by a multi-national consortium of research agencies and public health bodies. Each participating country team could adapt the generic protocol and maintain preapprovals from their ethics and health system committees in their own country. This could ensure rapid publication of a set of trials, tailored for local applicability, designed

prospectively for metanalysis, and could speed study closure, saving lives and eliminating the research waste arising from a flurry of uncoordinated trials.

The relationship between pragmatism, timing, and study size on impact of treatment trials: a qualitative, hypothesis generating study of systemic corticosteroids for COVID-19

1. Introduction

The novel SARS-CoV-2 (COVID-19) virus, detected in November of 2019, has quickly caused a global pandemic with a huge public health burden[1]. Because of the long timeline for producing new tailored pharmaceuticals, many researchers focused on repurposing existing medications for treatment of COVID-19 infection [2,3]. From November 2019 to December 2021, over 2,500 interventional trials were registered on ClinicalTrials.gov evaluating interventions for COVID-19[4]. Fatality from COVID-19 infection appeared to be partly an immune system overreaction and thus trials of anti-inflammatory treatments, such as corticosteroids, were a priority[4,5]. Corticosteroids have previously been used to mitigate severe organ injuries in other viral pneumonias, but the initial recommendations for using corticosteroids for COVID-19 were uncertain as small-scale, non-randomized studies

93 demonstrated contradictory results [6–8]. This uncertainty drove some to use more rigorous
94 methods to assess the effectiveness of steroids forCOVID-19[2,3]. Due to their high internal
95 validity randomized-controlled trials (RCTs) are considered the gold standard among designs for
96 evaluation of interventions in healthcare[9]. The basis of an RCT is random assignment of
97 participants into experimental and control groups with allocation concealment which, within the
98 bounds of chance, helps to balance the characteristics of the groups between arms at baseline.
99 Unless biases arise during the trial, the outcomes are attributable to differences in
100 intervention[10].

101 The urgent circumstances of the COVID-19 pandemic provides a unique opportunity for
102 our examination of design determinants of trial impact: a large number of RCTs on the same
103 drug treatment for a single disease, thus likely to share a common effect size, all trials conducted
104 at the same time, in different health systems but all under similar pressure, and with urgent
105 demand for clinical guidance, allowing the impacts to be measured by citations and incorporation
106 into rapidly produced national and international treatment guidelines. This allowed us to assess
107 how trial design features contribute to impact, reducing the differences in disease, drug, timing
108 or other factors that would be common in comparisons between RCTs. In this hypothesis
109 generating study, we analyze published RCTs of corticosteroid use for COVID-19 management
110 to identify design features that may explain their clinical and scientific impact, with the aim of
111 guiding future trial design.

112 We assessed study size, number of centres, single or multi-country conduct, dates of initiation
113 and of publication, known predictors of impact [27]. We also assessed two measures not
114 previously shown to predict impact, RoB-2 scores and PRECIS-2 scores. These are widely used
115 measures of internal and external validity respectively, the two main vulnerabilities in trial

design, which we believe may affect uptake of RCT findings. Each trial was assessed on its scientific impact and clinical impact, determined from the number of article citations generated and number of guideline citations respectively.

2. Materials and Methods

2.1 Methodology Overview

This study was based on a systematic search to ensure all relevant trials were included; the analysis of the included studies was qualitative, comparing characteristics of the highest impact trial with less influential trials.

2.2 Search Strategy

A systematic review of the literature was performed by searching the databases MEDLINE, EMBASE, Scopus, and Cochrane up until April 13, 2021. The search strategy included database-specific keywords and Medical Subject Headings (MeSH) terms (Appendix A1). Studies were limited to those performed on human subjects. No limitations were placed on the publication date, language or geographic location.

2.3 Inclusion/exclusion criteria

Studies had to employ an RCT design to investigate the use of one or more systemic corticosteroids to treat a COVID-19 infection or a COVID-19-induced condition in human

participants. Studies that were not randomized or involved non-human participants, were excluded. Platform trials testing several interventions were included if one or more of these was a systemic corticosteroid.

2.4 Study Selection and Screening

The studies obtained through database searches were imported into Covidence Systematic Review Software (Covidence, Veritas Health Innovation, Melbourne, Australia)[11]. The systematic screening process was performed independently by three reviewers (KDC, DD, and AL) which involved an initial title and abstract screen followed by full-text screen(see PRISMA flow diagram, Appendix B, Figure B.1). Disagreements were resolved by consensus meetings.

2.5 Data Extraction and Consensus Generation

Standardized data extraction was completed independently by three reviewers (KDC, DD, and AL). The extracted data was: author(s), dates of trial initiation, and of publication, the number of trial participants, the location and number of centers. Internal validity was assessed using the Cochrane RoB 2.0 tool[12] and external validity was quantified using the PRECIS-2 tool[13], by all three reviewers. Discrepancies were resolved through consensus.

2.6 PRECIS-2 (External Validity)

The second version of PRagmatic Explanatory Continuum Indicator Summary (PRECIS-2)[13] tool was used to assess the trial design. The scoring system of PRECIS-2 is composed of nine domains - Eligibility, Recruitment, Setting, Organization, Flexibility (delivery), Flexibility (adherence), Follow-up, Primary Outcome, Primary Analysis - scored from 1 (very explanatory)

to 5 (very pragmatic).. All three reviewers independently scored each included RCT using the PRECIS-2 tool. Any discrepancies in scores for a PRECIS-2 domain were resolved during a consensus meeting with the codeveloper of the tool (MZ). The final consensus score for each domain was used to generate the PRECIS-2 wheel for each included study using http://precis-2.org/.

2.7 Risk of Bias Assessment (Internal Validity)

The second version of the Cochrane risk-of-bias tool for randomized trials (RoB-2)[12] was used to assess the risk of bias in all studies analyzed in this paper. Included studies were assessed independently by all three reviewers (KDC, DD, and AL) and disagreements were resolved by consensus meetings. RoB-2 assesses bias in the following five domains: 1) risk of bias arising from the randomization process, 2) bias due to deviations from intended interventions, 3) bias due to missing outcome data, 4) bias in measurement of the outcome, and 5) bias in selection of the reported result[14].An algorithm consisting of a series of signalling questions leads to an assignment of "high risk," "some concerns," or "low risk," for each domain and an overall risk-of-bias judgment.

2.8 Assessment of Impact and Importance

Completed trials were ranked based on both their clinical impact and scientific impact. Scientific impact was indicated by the number of "cited by" articles on PubMed at time of data extraction, while clinical impact was defined by the number of major national and transnational clinical guidelines for COVID-19 that cited the trial findings. To obtain these rankings, information regarding the number of citations in Pubmed and citation in prominent national and

international clinical guidelines for COVID-19 (European Medicines Agency (EMA)[15], UK National Health Service (NHS)[16], UK National Institute for Health and Care Excellence (NICE)[17], US National Institute of Health (NIH)[18], World Health Organization (WHO)[19])) were extracted for each study.

3. Results

3.1 Search Results

Following the completion of the systematic search and the removal of duplicates, 443 unique articles were identified and screened. After level one screening, 58 studies progressed to a full text screen after which sixteen articles remained. Interobserver agreement was good. Cohen's kappa (K) coefficient for screening and full-text review of 0.72 and 0.87 respectively. Of these sixteen studies, only seven had been published in peer-reviewed journals and were eligible for this analysis. The literature search is summarized in the PRISMA flow diagram (see Appendix A, Figure A.1).

3.2 Trial Characteristics

We extracted data from seven studies: Horby et al, the RECOVERY trial[20], Angus et al, the REMAP-CAP trial[21], Tomazini et al, the CoDEX trial[22], Dequin et al, the CAPE COVID trial[23], trial by Edalatifard et al.[24], trial by Jamaati et al.[25], and trial by Tang et al.[26] The characteristics of the seven selected studies are provided in Table 1 (ordered by study size).

208 **Table 1.** Extracted characteristics for included articles.

Study, centres, country	Patient Recruitment	Publication Date	Size (N)	Pragmatism: 8 domains/ 40*	RoB **
Horby et al. 2021[20] 176 centres, UK	March 19 to June 8, 2020	July 17, 2020 (Preliminary report) Feb 25, 2021	6420	39	Low
Angus et al., 2020[21] 121 centres, International***	March 9 to June 17, 2020	Sept 2, 2020	614	30	Some
Tomazini et al. 2020[22] 41 centres, Brazil	April 17 to June 23, 2020	Oct 2, 2020	299	21	Some
Dequin et al. 2020[23] 33 centres, France	March 7 to July 3, 2020	Sept 2, 2020	149	28	Low
Edalatifard et al. 2020[24] 2 centres, Iran	April 20 to June 20, 2020	Sept 7, 2020	68	33	Low
Jamaati et al. 2021[25] 1 centre, Iran	March 2020, for 28 days	Feb 16, 2021	50	34	Some
Tang et al. 2021[26] 7 centres, China	Feb. 14 to March 31, 2020	Jan 22, 2021	86	29	Some

209 *3.2.1 Recruitment and publication dates*

210 Based on participant recruitment, Tang et al. was the earliest trial and began recruiting on
211 Feb 15, 2020. Jamaati et al., Horby et al. (the RECOVERY trial), Dequin et al., and Angus et al.,
212 all started recruitment in March of 2020 followed by Tomazini et al. and Edalatifard et al., in
213 April of 2020. The start dates, unsurprising given the pandemic, were within one or two months

of each other, but Horby et al. published their preliminary report first, in July 2020, followed by Angus et al., Dequin et al., and Edalatifard et al., in September 2020., Tomazini et al. in October 2020 and Tang et al., and Jamaati et al. in January and February of 2021 respectively. The final result from Horby et al. was also published February 2021.

3.2.2 Size and centres

Horby et al. (176 centres in the United Kingdom, n = 6420, 36 patients per centre) recruited by far the most participants (10 times more patients and seven times more patients per centre than the next largest trial, by Angus et al. (121 centres in Australia, Canada, France, Ireland, Netherlands, New Zealand, the United Kingdom, and the Unites States of America, n = 614). The remaining five trials were single country and much smaller; Dequin et al. (33 centres in France, n = 149), Edalatifard et al. (2 centres in Iran, n = 68), Jamaati et al. (1 centre in Iran, n = 50), Tang et al. (7 centres in China, n = 86), and Tomazini et al. (Brazil, n = 299).

3.2.3 Internal and external validity

Internal validity was assessed with the Cochrane RoB 2.0 tool. Dequin et al., Edalatifard et al., and Horby et al. each had a 'low' overall risk of bias score. The trials by Angus et al., Jamaati et al., Tang et al., and Tomazini et al. were classified as having 'some' risk of bias. See Appendix B, Table B.1.

Eight of the nine domains of PRECIS-2 were scored: flexibility (adherence) was not scored as these trials tested a hospital administered drug with no patient adherence component. Eligibility, recruitment, setting, organisation, flexibility (delivery), follow-up, primary outcome, and primary analysis were independently scored for each trial. The final consensus scores

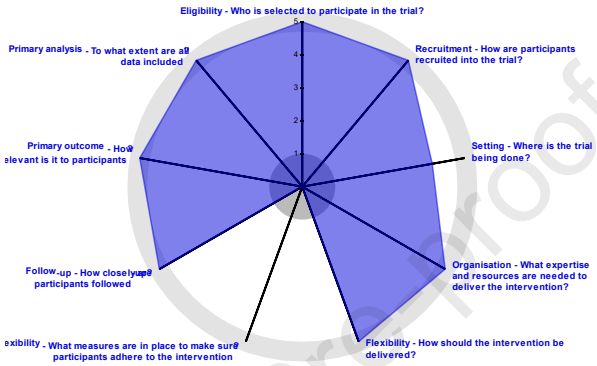
237 between all three reviewers are shown in Appendix B, Table B.2 and the associated PRECIS-2

238 wheels are shown in Figure 1.

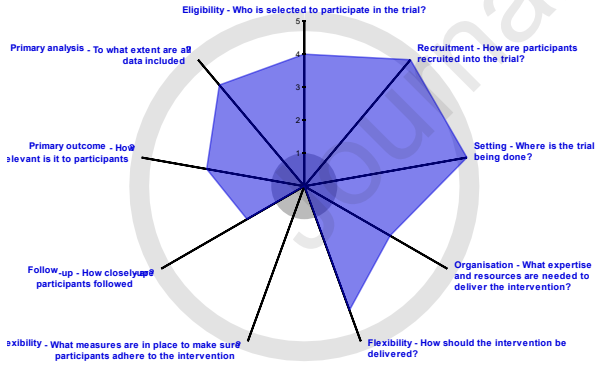
239

240 **Figure 1.** PRECIS-2 Score Wheel

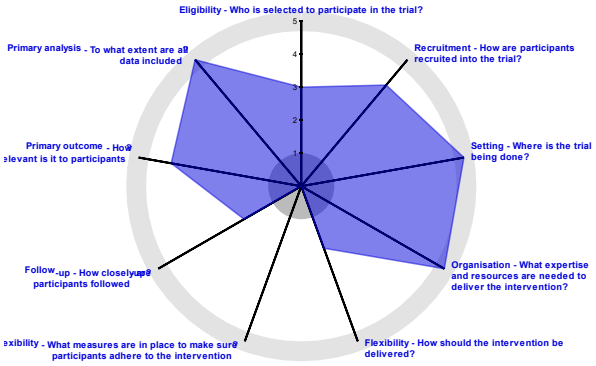
Horby et al. 2021[20]



Angus et al. 2020[21]

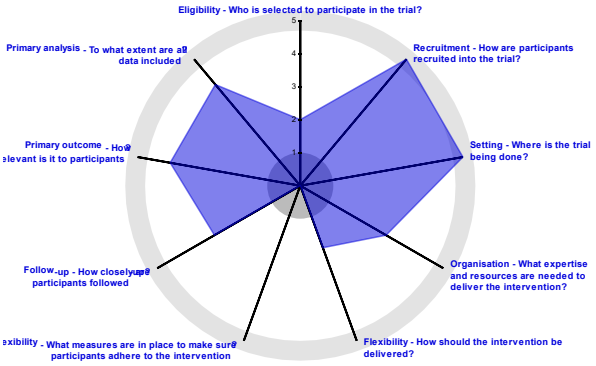


Tomazini et al. 2020[22]

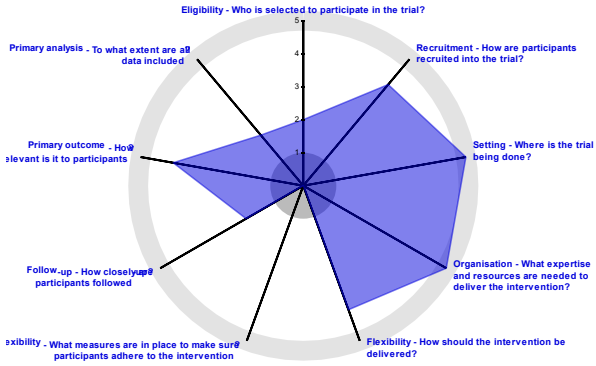


Dequin et al. 2020[23]

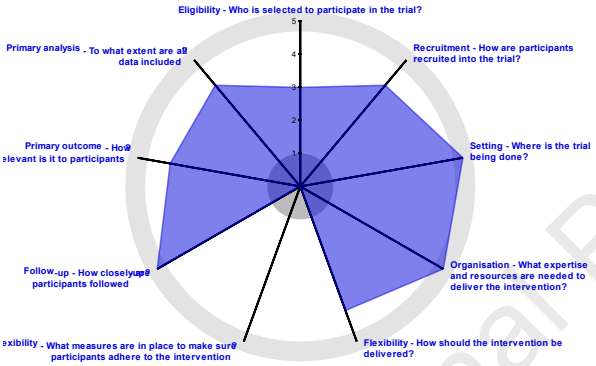
Edalatifard et al. 2020[24]

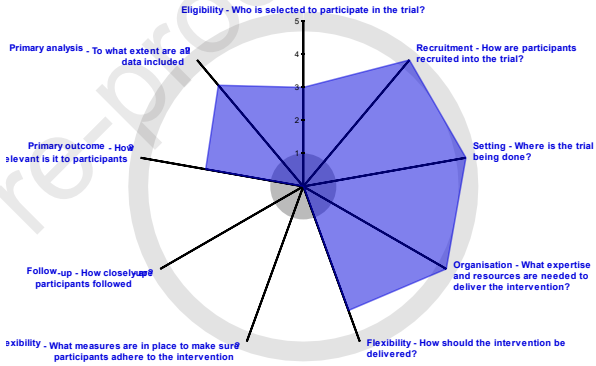


Jamaati et al. 2021[25]



Tang et al. 2021[26]





241 The RECOVERY trial (Horby et al.) had the highest overall PRECIS-2 score (most
242 pragmatic design features) with even distribution across all domains, suggesting that the
243 designers employed a consistently pragmatic approach to their trial design. In contrast, other
244 trials all had at least one domain that scored 3 or lower, suggesting a less consistent intention
245 towards pragmatism.

246

247 *3.3 Assessment of Impact*

248 The impact of each trial was assessed based on its scientific impact and clinical impact,
249 which were determined from the number of article citations generated and number of guideline
250 citations respectively. Extracted values can be found in Table 2.

Table 2. Extracted characteristics indicative of impact for included articles.

Study	Corticosteroid	Citations*	Influenced Guidelines
Horby et al. 2021[20]	Dexamethasone	1887	EMA, NHS, NICE, NIH, WHO
Angus et al., 2020[21]	Hydrocortisone	130	NHS, NICE, WHO
Tomazini et al. 2020[22]	Dexamethasone	188	NHS, NICE, WHO
Dequin et al. 2020[23]	Hydrocortisone	75	NHS, NICE, WHO
Edalatifard et al. 2020[24]	Methylprednisolone	35	NICE
Jamaati et al. 2021[25]	Dexamethasone	1	-
Tang et al. 2021[26]	Methylprednisolon	2	-

European Medicines Agency (EMA); National Health Service (NHS); National Institute for Health and Care Excellence (NICE); National Institute of Health (NIH); and the World Health Organization (WHO).
* Pubmed was used to identify citations for all trials.

As of March 27, 2021, the trial by Horby et al. was cited by 1887 other articles, which was over ten times as many citations as the next most cited RCT included in our study, thus establishing it as the most important paper based on this marker. The trials conducted by Tomazini et al. and Angus et al. were the second and third most cited trials respectively, each with over 100. Next, the trial by Dequin et al. and Edalatifard et al. had 75 and 35 citations respectively followed by Tang et al. and Jamaati et al. with 2 and 1 citations respectively.

The trial by Horby et al. was the only trial that influenced all five selected guidelines, hence its designation as the most impactful trial by this measure. The trials conducted by Tomazini et al., Sequin et al., and Angus et al. were referenced only by three treatment guidelines, published by the NHS, NICE, and WHO. The trial conducted by Edalatifard et al. influenced the NICE treatment guideline. Finally, the trials conducted by Jamaati et al. and Tang et al. 2021 were not cited by any of the selected treatment guidelines.

Based on these criteria, the RECOVERY trial (Horby et al.) was the most impactful (see Appendix B, Table B.3).

4. Discussion

4.1 Conclusion

We assessed start of recruitment, publication timing, and study size, all of which are known predictors of impact [27] and easy to extract. We also assessed RoB-2 scores and PRECIS-2 scores, not previously reported as being associated with impact. We used the number of major guidelines influenced (as a proxy for clinical impact) and number of literature citations (as a proxy for scientific impact) because they are straight-forward, easily accessible, and accurate measures of the influence of a trial on scientific and clinical decision making.

The principal finding of our study was that the RECOVERY trial by Horby et al was dominant in impact compared to any of the other six: RECOVERY was the only trial mentioned in all of the five guidelines reviewed and had ten-fold more scientific citations than the next most cited. It was uniquely strong in level of pragmatism, large sample size, number of centres, and early publication, but was similar to others in start date for recruitment and risk of bias.

Sample size is known to affect impact [28–30], also supported by our findings. Timing of publication is critical as early publication of RECOVERY's preliminary finding, resulted in loss of equipoise and early termination of the corticosteroid arm in Angus et al (REMAP-CAP) [21], Tommazini et al (CoDEX)[22], and Dequin et al (CAPE COVID) [23], with research waste, a problem in biomedical research[32–35].

All these trials started within weeks of each other, so early start-up could not have advantaged RECOVERY but is obviously necessary for early publication. Four trials were at "low" or (next category up) "some" risk (three trials) for bias, a truncated range suggesting that thorough understanding of bias prevention is now widespread and while necessary for interpretability, is not sufficient to ensure early publication and impact. Similarly, all but one trial had multiple centres, necessary but not sufficient to rapidly recruit large numbers of patients.

This suggests that co-occurrence of the unique features may be responsible for the higher impact of this trial. We therefore hypothesize that the differences between RECOVERY and the other trials that may have enabled rapid recruitment and its early publication lie in its pragmatic design features (inclusive recruitment, common clinical outcomes, no restrictions on usual clinical care other than randomization, minimal monitoring, usual care comparators, minimal extra data collection beyond what is needed for clinical care). These features underly a simple trial design, with no distortion of the usual clinical flow and needing few extra resources, which was easier to integrate into the everyday clinical context. This in turn allowed more centres to join the RECOVERY trial and improved recruitment yield per centre, leading to faster recruitment and thus early publication. High levels of pragmatism may also have improved confidence in the applicability of the trial findings and added to uptake of RECOVERY findings into clinical guidelines[31].

Of course other issues may also have contributed to the greater impact of RECOVERY. The value of large platform trials is supported by reviews analyzing existing RCTs for non-epidemic conditions such as vitamin D for infectious diseases[38] and can also be seen in the protocol of the BEAT-CF platform trial of multiple treatments for exacerbation of cystic

fibrosis[39]. The platform design of the RECOVERY trial allowed simultaneous testing of multiple other COVID-19 treatments, which eliminated the need to set up separate RCTs for each one [27] thus adding to the attractiveness of the unified single funding proposal. However, this design feature does not directly contribute to the impact of the trial in relation to corticosteroids themselves, although it undoubtedly reduces the cost and effort for testing each of the interventions in the platform. Platform trials should thus be used where possible, with the caution that they depend on sophisticated statistical skills in both design and analysis. Where the scarcity of this level of statistical skills makes this unfeasible, simpler parallel arm pragmatic, multicenter RCTs will still achieve worthwhile findings quickly, albeit one intervention at a time (or a few, in multi-arm trials).

4.2 Recommendations

It took four months after the first acknowledgement of the pandemic for RECOVERY to be launched and a further four months for that trial to report its initial findings, which changed medical care outcomes for hospitalized patients with COVID-19 around the world. Had this trial been launched on recognition of the pandemic, 4 months of death and disability for many patients could have been avoided. How might we eliminate that four month delay in providing evidence based care for the next pandemic illness?

Some might argue that for a question as important as care of patients in a pandemic, where many thousands of very ill patients would be cared for based on evidence from a single trial, confirmation in one or two other independent trials is reassuring, especially if conducted in different settings. We agree but note that even if equipoise is not lost, forcing their closure, these other trials might be poorly designed and contribute only low-quality evidence that

confuses clinical decision making [36], as seen with trials of ivermectin for COVID-19 whose results did not impact clinical management[37].

We therefore recommend that for epidemic and pandemic situations, where evidence based decision support is urgently needed, important choices between interventions should be made using multiple, simultaneous, large, multi-centre pragmatic randomized trials using shared protocols. Since ethics and logistics procedures are difficult to co-ordinate between countries, we recommend that several such trials should always be launched simultaneously, each conducted by a separate team in their own country. Organizations with international reach, such as WHO and the Gates foundation could facilitate the preparation of shared protocols with each participating country offering different socio- economic and health care delivery systems. Prospectively planned metanalyses combining these studies could produce evidence applicable to a wide range of settings, with subgroup analyses that would answer important secondary questions. These protocols should be developed co-operatively, with many shared elements, especially pragmatic features such as use of inclusive inclusion and subgroup definitions, common clinical case definitions, reduced data collection through reduced monitoring, clinically ascertainable and/or widely available test measures, simple primary outcomes, usual care comparators, and simplified trial procedures.

The protocols should be prepared in advance of any pandemic, with several generic protocols, each appropriate for one of the expected kinds and routes of spread of pandemic illness. These generic protocols, each specific to an expected type of pandemic, could be prepared, maintained and updated centrally, working closely with each national team. In the event of a pandemic, the most appropriate generic protocol could be centrally adapted to the specifics of the actual pandemic agents, adapted in each country to their own needs and launched in several countries

simultaneously, early in the pandemic. A pragmatic approach to design of these trials may avoid the usual ponderousness of research and rapidly inform global clinical practice in a pandemic [40].

4.3 Study strengths and limitations

When a pandemic arises, which interventions should be evaluated? The RECOVERY trial provided a good model by studying repurposing of widely used and readily available medications that could be easily accessed by most health systems all over the world[41].

There are several strengths to our study. First, our screen yielded a high Cohen's kappa coefficient suggesting high inter-rater reliability and minimal risk of selection bias[42]. Another strength is that we compared trials examining the same treatment for the same indication during the same time period, in a global pandemic panic. Therefore, we were able to hold constant many factors, such as whether or not the study yielded positive results, different treatments, different time periods, different health system and disease contexts, and focus only on trial design characteristics.

Our study has limitations. Due to the limited number of clinical trials examining the effect of corticosteroids in COVID-19 patients, our study is a small qualitative analysis rather than a large metanalysis, which limits us to hypothesis generation rather than causal attribution [43]. As well, we chose as our proxy for clinical impact the number of prominent national or international guidelines influenced but ideally we would have measured this directly with prescribing data or interviews with clinicians. Unfortunately this prescribing data is only available with a substantial time lag, and interviews were not within the scope of a student project.

Overall, our finding of the importance (and interaction) of pragmatism, size and timing points to trial design characteristics that future trial makers might find helpful to maximize the clinical and policy utility of their trials. Future research on this question should use more definitive outcomes for impact, should quantify the relationship between impact and potential study design correlates, and study this question also for non-epidemic situations[39].

Word Count: 3363 (excluding headings and figures)

Acknowledgements

Funding: This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Conflicts of interest/Competing interests: Not applicable

Compliance with Ethical Standards: All data was abstracted from the published literature and therefore ethics approval was unnecessary.

CRediT Author Statement

Aileen Liang: Conceptualization, Methodology, Investigation, Analysis, Interpretation, Data curation, Writing - Original draft preparation, Writing - Review & Editing, Visualization. **Katrina Domenica Cirone:** Conceptualization, Methodology, Investigation, Analysis, Interpretation, Data curation, Writing - Original draft preparation, Writing - Review & Editing, Visualization. **Xiaoxiao (Daisy) Deng:** Conceptualization, Methodology, Investigation, Analysis, Interpretation, Data curation, Writing- Original draft preparation, Writing - Review & Editing, Visualization. **Merrick Zwarenstein:** Conceptualization, Methodology, Investigation,

Analysis, Interpretation, Writing - Review & Editing, Visualization, Supervision.

Reference

[1] Velavan TP, Meyer CG. The COVID-19 epidemic. Trop Med Int Health 2020;25:278–80. https://doi.org/10.1111/tmi.13383.

[2] Cao X. COVID-19: immunopathology and its implications for therapy. Nat Rev Immunol 2020:1–2. https://doi.org/10.1038/s41577-020-0308-3.

[3] Li T, Lu H, Zhang W. Clinical observation and management of COVID-19 patients. Emerg Microbes Infect 2020;9:687–90. https://doi.org/10.1080/22221751.2020.1741327.

[4] Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. Lancet Lond Engl 2020;395:497–506. https://doi.org/10.1016/S0140-6736(20)30183-5.

[5] Cheng PK, Wong DA, Tong LK, Ip S-M, Lo AC, Lau C-S, et al. Viral shedding patterns of coronavirus in patients with probable severe acute respiratory syndrome. Lancet Lond Engl 2004;363:1699–700. https://doi.org/10.1016/S0140-6736(04)16255-7.

[6] Ye Z, Wang Y, Colunga-Lozano LE, Prasad M, Tangamornsuksan W, Rochwerg B, et al. Efficacy and safety of corticosteroids in COVID-19 based on evidence for COVID-19, other coronavirus infections, influenza, community-acquired pneumonia and acute respiratory distress syndrome: a systematic review and meta-analysis. CMAJ Can Med Assoc J J Assoc Medicale Can 2020;192:E756–67. https://doi.org/10.1503/cmaj.200645.

[7] Corral-Gudino L, Bahamonde A, Arnaiz-Revillas F, Gómez-Barquero J, Abadía-Otero J, García-Ibarbia C, et al. GLUCOCOVID: A controlled trial of methylprednisolone in adults hospitalized with COVID-19 pneumonia. 2020. https://doi.org/10.1101/2020.06.17.20133579.

[8] Russell CD, Millar JE, Baillie JK. Clinical evidence does not support corticosteroid treatment for 2019-nCoV lung injury. The Lancet 2020;395:473–5. https://doi.org/10.1016/S0140-6736(20)30317-2.

[9] Hariton E, Locascio JJ. Randomised controlled trials – the gold standard for effectiveness research. BJOG Int J Obstet Gynaecol 2018;125:1716–1716. https://doi.org/10.1111/1471-0528.15199.

[10] Jaillon P. [Controlled randomized clinical trials]. Bull Acad Natl Med 2007;191:739–56; discussion 756-758.

[11] Covidence - Better systematic review management n.d. https://www.covidence.org/ (accessed August 27, 2021).

[12] Sterne JAC, Savović J, Page MJ, Elbers RG, Blencowe NS, Boutron I, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. BMJ 2019;366:l4898. https://doi.org/10.1136/bmj.l4898.

[13] Loudon K, Treweek S, Sullivan F, Donnan P, Thorpe KE, Zwarenstein M. The PRECIS-2 tool: designing trials that are fit for purpose. BMJ 2015;350:h2147. https://doi.org/10.1136/bmj.h2147.

[14] Cochrane Handbook for Systematic Reviews of Interventions n.d. https://training.cochrane.org/handbook/current (accessed August 27, 2021).

[15] CZARSKA-THORLEY D. EMA endorses use of dexamethasone in COVID-19 patients on oxygen or mechanical ventilation. Eur Med Agency 2020. https://www.ema.europa.eu/en/news/ema-endorses-use-dexamethasone-covid-19-patients-oxygen-mechanical-ventilation (accessed August 27, 2021).
[16] NHS England » COVID-19 therapy: corticosteroids including dexamethasone and hydrocortisone n.d. https://www.england.nhs.uk/publication/covid-19-therapy-corticosteroids-including-dexamethasone-and-hydrocortisone/ (accessed August 27, 2021).
[17] COVID-19 rapid guideline: managing the long-term effects of COVID-19 (NG188): Evidence review 5: interventions. London: National Institute for Health and Care Excellence (UK); 2020.
[18] Corticosteroids. COVID-19 Treat Guidel n.d. https://www.covid19treatmentguidelines.nih.gov/therapies/immunomodulators/corticosteroids/ (accessed August 27, 2021).
[19] Corticosteroids for COVID-19 n.d. https://www.who.int/publications-detail-redirect/WHO-2019-nCoV-Corticosteroids-2020.1 (accessed August 27, 2021).
[20] RECOVERY Collaborative Group, Horby P, Lim WS, Emberson JR, Mafham M, Bell JL, et al. Dexamethasone in Hospitalized Patients with Covid-19. N Engl J Med 2021;384:693–704. https://doi.org/10.1056/NEJMoa2021436.
[21] Angus DC, Derde L, Al-Beidh F, Annane D, Arabi Y, Beane A, et al. Effect of Hydrocortisone on Mortality and Organ Support in Patients With Severe COVID-19: The REMAP-CAP COVID-19 Corticosteroid Domain Randomized Clinical Trial. JAMA 2020;324:1317–29. https://doi.org/10.1001/jama.2020.17022.
[22] Tomazini BM, Maia IS, Cavalcanti AB, Berwanger O, Rosa RG, Veiga VC, et al. Effect of Dexamethasone on Days Alive and Ventilator-Free in Patients With Moderate or Severe Acute Respiratory Distress Syndrome and COVID-19: The CoDEX Randomized Clinical Trial. JAMA 2020;324:1307–16. https://doi.org/10.1001/jama.2020.17021.
[23] Dequin P-F, Heming N, Meziani F, Plantefève G, Voiriot G, Badié J, et al. Effect of Hydrocortisone on 21-Day Mortality or Respiratory Support Among Critically Ill Patients With COVID-19: A Randomized Clinical Trial. JAMA 2020;324:1298–306. https://doi.org/10.1001/jama.2020.16761.
[24] Edalatifard M, Akhtari M, Salehi M, Naderi Z, Jamshidi A, Mostafaei S, et al. Intravenous methylprednisolone pulse as a treatment for hospitalised severe COVID-19 patients: results from a randomised controlled clinical trial. Eur Respir J 2020;56. https://doi.org/10.1183/13993003.02808-2020.
[25] Jamaati H, Hashemian SM, Farzanegan B, Malekmohammad M, Tabarsi P, Marjani M, et al. No clinical benefit of high dose corticosteroid administration in patients with COVID-19: A preliminary report of a randomized clinical trial. Eur J Pharmacol 2021;897:173947. https://doi.org/10.1016/j.ejphar.2021.173947.
[26] Tang X, Feng Y-M, Ni J-X, Zhang J-Y, Liu L-M, Hu K, et al. Early Use of Corticosteroid May Prolong SARS-CoV-2 Shedding in Non-Intensive Care Unit Patients with COVID-19 Pneumonia: A Multicenter, Single-Blind, Randomized Control Trial. Respir Int Rev Thorac Dis 2021;100:116–26. https://doi.org/10.1159/000512063.
[27] Saville BR, Berry SM. Efficiencies of platform clinical trials: A vision of the future. Clin Trials Lond Engl 2016;13:358–66. https://doi.org/10.1177/1740774515626362.
[28] Faber J, Fonseca LM. How sample size influences research outcomes. Dent Press J Orthod 2014;19:27–9. https://doi.org/10.1590/2176-9451.19.4.027-029.ebo.

[29] Biau DJ, Kernéis S, Porcher R. Statistics in Brief: The Importance of Sample Size in the Planning and Interpretation of Medical Research. Clin Orthop 2008;466:2282–8. https://doi.org/10.1007/s11999-008-0346-9.
[30] Andrade C. Sample Size and its Importance in Research. Indian J Psychol Med 2020;42:102–3. https://doi.org/10.4103/IJPSYM.IJPSYM_504_19.
[31] Mentz RJ, Hernandez AF, Berdan LG, Rorick T, O'Brien EC, Ibarra JC, et al. Good Clinical Practice Guidance and Pragmatic Clinical Trials. Circulation 2016;133:872–80. https://doi.org/10.1161/CIRCULATIONAHA.115.019902.
[32] Characteristics of registered clinical trials assessing treatments for COVID-19: a cross-sectional analysis | BMJ Open n.d. https://bmjopen.bmj.com/content/10/6/e039978 (accessed August 27, 2021).
[33] Hsiehchen D, Espinoza M, Hsieh A. Deficiencies in the Designs and Interventions of COVID-19 Clinical Trials. Med N Y N 2020;1:103–4. https://doi.org/10.1016/j.medj.2020.06.007.
[34] Glasziou PP, Sanders S, Hoffmann T. Waste in covid-19 research. BMJ 2020;369:m1847. https://doi.org/10.1136/bmj.m1847.
[35] Dal-Ré R, Mahillo-Fernández I. Waste in COVID-19 clinical trials conducted in western Europe. Eur J Intern Med 2020;81:91–3. https://doi.org/10.1016/j.ejim.2020.07.002.
[36] Honarmand K, Penn J, Agarwal A, Siemieniuk R, Brignardello-Petersen R, Bartoszko JJ, et al. Clinical trials in COVID-19 management & prevention: A meta-epidemiological study examining methodological quality. J Clin Epidemiol 2021;139:68–79. https://doi.org/10.1016/j.jclinepi.2021.07.002.
[37] Bias as a source of inconsistency in ivermectin trials for COVID-19: A systematic review. Ivermectin's suggested benefits are mainly based on potentially biased results. - Journal of Clinical Epidemiology n.d. https://www.jclinepi.com/article/S0895-4356(21)00422-4/fulltext (accessed January 4, 2022).
[38] Yamshchikov AV, Desai NS, Blumberg HM, Ziegler TR, Tangpricha V. Vitamin D for treatment and prevention of infectious diseases: a systematic review of randomized controlled trials. Endocr Pract Off J Am Coll Endocrinol Am Assoc Clin Endocrinol 2009;15:438–49. https://doi.org/10.4158/EP09101.ORR.
[39] Schultz A, Marsh JA, Saville BR, Norman R, Middleton PG, Greville HW, et al. Trial Refresh: A Case for an Adaptive Platform Trial for Pulmonary Exacerbations of Cystic Fibrosis. Front Pharmacol 2019;10:301. https://doi.org/10.3389/fphar.2019.00301.
[40] Institute of Medicine (US) Forum on Drug Discovery D. Challenges in Clinical Research. National Academies Press (US); 2010.
[41] Welcome — RECOVERY Trial n.d. https://www.recoverytrial.net/ (accessed January 4, 2022).
[42] Chan K, Bhandari M. Three-minute critical appraisal of a case series article. Indian J Orthop 2011;45:103–4. https://doi.org/10.4103/0019-5413.77126.
[43] Mathes T, Pieper D. Clarifying the distinction between case series and cohort studies in systematic reviews of comparative studies: potential impact on body of evidence and workload. BMC Med Res Methodol 2017;17:107. https://doi.org/10.1186/s12874-017-0391-8.

535

536

537

538

539

540

541

542

543

544

545

546

547

Declaration of Interest

Declarations of interest: none

The relationship between pragmatism, timing, and study size on impact of treatment trials: a qualitative, hypothesis generating study of systemic corticosteroids for COVID-19

CRediT Author Statement

Aileen Liang: Conceptualization, Methodology, Investigation, Analysis, Interpretation, Data curation, Writing - Original draft preparation, Writing - Review & Editing, Visualization.
Katrina Domenica Cirone: Conceptualization, Methodology, Investigation, Analysis, Interpretation, Data curation, Writing - Original draft preparation, Writing - Review & Editing, Visualization.
Xiaoxiao (Daisy) Deng: Conceptualization, Methodology, Investigation, Analysis, Interpretation, Data curation, Writing- Original draft preparation, Writing - Review & Editing, Visualization.
Merrick Zwarenstein: Conceptualization, Methodology, Investigation, Analysis, Interpretation, Writing - Review & Editing, Visualization, Supervision.

No specific funding was required for this study.

The relationship between pragmatism, timing, and study size on impact of treatment trials: a qualitative, hypothesis generating study of systemic corticosteroids for COVID-19

What is new?

- RECOVERY is a large multicenter, single country, platform randomized trial of several treatments repurposed for COVID-19 care. The sub-trial of corticosteroids had a low risk of bias and highly pragmatic design features that facilitated wide implementation and rapid recruitment. This trial was cited ten times more often than the next most cited trial and relied on in all the prominent guidelines we reviewed, changing clinical practice globally. It eliminated equipoise, rendering redundant the other simultaneous trials of steroids that lacked one or more of these features. This research waste should be reduced.
- Large, pragmatic, unbiassed, single country platform trials of repurposed drugs and interventions, covering different potential pandemic conditions and multiple treatments, for a range of sociodemographic situations and healthcare capacities could be a valuable investment in readiness for future pandemics, resulting in trials with greater scientific and clinical impact and less research waste.
- Generic protocols for trials aimed at each kind of pandemic threat could be prepared, in advance, by a multi-national consortium of research agencies and public health bodies. Each participating country team could adapt the generic protocol and maintain preapprovals from their ethics and health system committees in their own country. This could ensure rapid publication of a set of trials, tailored for local applicability, designed prospectively for metanalysis, and could speed study closure, saving lives and eliminating the research waste arising from a flurry of uncoordinated trials.