The Landscape of COVID-19 Research in the United States: a Cross-sectional Study of Randomized Trials Registered on ClinicalTrials.Gov

Chana A. Sacks, MD, MPH1,2, Crystal M. North, MD, MPH2,3,4, Molly Wolf, MD2,3, Michael Dougan, MD, PhD2,5, Kieran R Campbell, D. Phil6, Jason Moggridge, BSc6, and Michael Fralick, MD, PhD, SM, FRCP(C)6,7

1Division of General Internal Medicine and Mongan Institute, Massachusetts General Hospital, Boston, MA, USA; 2Division of Pulmonary and Critical Care Medicine, Massachusetts General Hospital, Boston, MA, USA; 3Medical Practice Evaluation Center, Massachusetts General Hospital, Boston, MA, USA; 4Division of Gastroenterology, Massachusetts General Hospital, Boston, MA, USA; 5Division of Cardiovascular Medicine, Massachusetts General Hospital, Boston, MA, USA; 6Lunenfeld Tanenbaum Research Institute, Sinai Health System, 60 Murray Street, Toronto, ON, Canada; 7Department of Medicine, University of Toronto, Toronto, Canada.

**Importance:** SARS-CoV-2 has infected over 200 million people worldwide, resulting in more than 4 million deaths. Randomized controlled trials are the single best tool to identify effective treatments against this novel pathogen.

**Objective:** To describe the characteristics of randomized controlled trials of treatments for COVID-19 in the United States launched in the first 9 months of the pandemic.

**Design, Setting, and Participants**
We conducted a cross-sectional study of all completed or actively enrolling randomized, interventional, clinical trials for the treatment of COVID-19 in the United States registered on www.clinicaltrials.gov as of August 10, 2020. We excluded trials of vaccines and other interventions intended to prevent COVID-19.

**Main Outcomes and Measures**

We used descriptive statistics to characterize the clinical trials and the statistical power for the available studies. For the late-phase trials (i.e., phase 3 and 2/3 studies), we compared the geographic distribution of the clinical trials with the geographic distribution of people diagnosed with COVID-19.

**Results:** We identified 200 randomized controlled trials of treatments for people with COVID-19. Across all trials, 87 (43.5%) were single-center, 64 (32.0%) were unblinded, and 80 (40.0%) were sponsored by industry. The most common treatments included monoclonal antibodies (N=46 trials), small molecule immunomodulators (N=28), antiviral medications (N=24 trials), and hydroxychloroquine (N=20 trials). Of the 9 trials completed by August 2020, the median sample size was 450 (IQR 67–1113); of the 191 ongoing trials, the median planned sample size was 150 (IQR 60–400). Of the late-phase trials (N=54), the most common primary outcome was a severity scale (N=23, 42.6%), followed by a composite of mortality and ventilation (N=10, 18.5%), and mortality alone (N=6, 11.1%). Among these late-phase trials, all trials of antivirals, monoclonal antibodies, or chloroquine/hydroxychloroquine had a power of less than 25% to detect a 20% relative risk reduction in mortality. Had the individual trials for a given class of treatments instead formed a single trial, the power to detect that same reduction in mortality would have been greater than 98%. There was large variability in access to trials with the highest number of trials per capita in the Northeast and the lowest in the Midwest.

**Conclusions and Relevance:** A large number of randomized controlled trials were launched early in the pandemic to evaluate treatments for COVID-19. However, many trials were underpowered for important clinical endpoints and substantial geographic disparities were observed, highlighting the importance of improving national clinical trial infrastructure.
INTRODUCTION

Since its emergence in December 2019, SARS-CoV-2 has infected more than 200 million people around the world.\(^1\) The United States (US) accounts for approximately 4% of the world’s population but in the first 9 months of the pandemic recorded approximately 26% of reported global cases and 23% of the world’s COVID-19-related deaths.\(^1,2\) At the start of the pandemic, no treatments beyond supportive care had been proven to be effective for COVID-19; therefore, there was an urgent need to test therapeutic strategies, including both repurposed medications and novel agents. As with any therapeutic intervention, randomized controlled trials are the single best tool to evaluate safety and efficacy of possible treatments against this novel pathogen.\(^3,4\)

The mechanisms that underlie the variable clinical syndrome caused by SARS-CoV-2 remain an area of active investigation.\(^3\) Many hypotheses have been suggested, which have received heightened publicity in both the scientific and lay press. At the beginning of the pandemic, in order to both maintain social distancing mandates as well as to divert crucial resources to hospitals’ efforts to develop SARS-CoV-2 testing and expand capacity for clinical care, the healthcare and biomedical research sectors temporarily stopped much of their non-COVID-19-related research. The urgency of the global threat, the temporary halting of existing research programs, and the availability of COVID-19-related research funding resulted in an unprecedented rise in COVID-19-related research.\(^5,6\) Interventional trials focused on identifying effective treatments were rapidly developed and launched, in some cases in a matter of weeks.

Several centrally coordinated national trials platforms exist across the US, many of which transitioned their entire infrastructure to running COVID-19-related clinical trials, but the need for trials was far greater than the capacity of these networks.\(^7,8\) Therefore, individual academic medical centers and pharmaceutical companies developed and carried out site-specific clinical trials. Such an uncoordinated approach risks conducting duplicative studies of the same interventions, underpowering studies, and extending the length of time needed to determine definitive answers. Having broad networks across the country is particularly critical in a pandemic during which cases geographically surge and recede, affecting an individual trial site’s ability to consistently enroll participants.

The landscape of pathophysiologic mechanisms under study, how the distribution of trial sites across the US compared to the national COVID-19 caseload, and the likelihood that the registered studies were powered to address the hypotheses under investigation remain unknown. To address this gap in knowledge, we leveraged data collected on the clinicaltrials.gov platform in the first 9 months of the global pandemic to characterize the landscape of clinical trials for treatments of COVID-19 across the US.

METHODS

Data Source

We conducted a cross-sectional study of all randomized, interventional trials for patients diagnosed with COVID-19 in the US that were registered on www.clinicaltrials.gov and started enrolling as of August 10, 2020 (approximately 9 months after the first cases of COVID-19 in the world were identified).\(^9\) We excluded clinical trials of vaccines and other interventions attempting to prevent (rather than treat) COVID-19, as well as studies that were registered but had not yet begun enrollment. Research ethics approval was not required because www.clinicaltrials.gov is a publicly available data repository.

Study Characteristics

We used natural language processing (i.e., regular expression) and manual data extraction to record information for each registered clinical trial. We extracted data related to trial-level characteristics including study design (e.g., blinding), study location, study setting (i.e., outpatient, inpatient, or both), funding source, patient population, planned sample size, intervention, comparator (i.e., placebo-controlled or active comparator(s)), and the trial’s primary outcome. Details related to the primary outcome included whether the outcome was a surrogate outcome or clinical outcome and whether it was a single or composite outcome. We refer to studies that were registered as phase 2/3 or phase 3 as “phase 3” studies.

We categorized the available interventions based on their mechanism of action (e.g., antiviral, antibiotic, anticoagulant). For clinical trials involving more than one intervention, we collected data on each planned intervention. In terms of geographic distribution, we identified all the participating study centers indicated on clinicaltrials.gov.

Statistical Analysis

To categorize the treatments studied in the selection of COVID-19 clinical trials, we performed keyword detection for drug or therapy names using regular expressions on the text descriptions given in the “Arms/Intervention” fields of the records from clinicaltrials.gov. Categorization of primary outcomes was carried out in a similar manner, with the descriptions given in the “Primary Outcome Measures” fields. Keyword detection was performed using the R programming language (RCore Team, 2020) with functions for regular expressions from the stringr package. Labeling of treatment and primary outcomes was then checked manually for each trial to ensure accurate labeling.

We used descriptive statistics to summarize the clinical trials’ characteristics. We then calculated the statistical power of the phase 3 trials in their current form compared to their statistical power if the patients from the individual trials for a given therapeutic intervention were pooled into a single trial. For phase 3 studies, we tabulated the sample size per study arm and calculated the statistical power for the trial as registered and again if all of the individual arms for a given
treatment were pooled (i.e., if all the patients contributed to one single trial rather than having many smaller individual trials). We assumed a 15% risk of mortality, an alpha of 0.05, and one-to-one randomization, and tested the robustness of our results by comparing model estimates to those with assumed mortality of 10% and 20%, respectively. We used the Miettinen formula to calculate power using the open-source EpiSheets software (R package episheet source). We used population census data from 2019 to calculate the standardized proportion of clinical trials by state per 100,000 people living within each state. We also reported the clinical trials by the number of COVID-19 cases per 100,000 people in that state.

RESULTS

We identified 200 interventional clinical trials for patients with COVID-19 (Figure 1). The median planned sample size was 150 patients (IQR 60–400), 9 (5%) had completed enrollment, 87 (44%) were single-center, 64 (32%) were unblinded, and 136 (68%) were placebo-controlled. Most studies (N=115, 58%) specified the sponsor as “Other.” The primary study sponsor was pharmaceutical companies for 80 trials (40%), with an additional 3% (N=5) of trials funded by the National Institutes of Health (NIH) or other sources of federal funding.

Nearly all studies required patients to be at least 18 years of age (N=188, 94%) and 26% (N=52) had an upper age limit for inclusion; 32 trials had upper age limits ranging between 50 and 85 years of age, with an additional 20 having upper age limits above this age. Clinical trials also commonly excluded pregnant women (73%); other populations excluded by multiple trials included people with HIV/AIDS (19%), people who were incarcerated (9%), and people with mental health diagnoses (8%) (Table 1).

The most common categories of treatments by mechanism of action included monoclonal antibodies (N=46), small molecule immunomodulators (N=28), antivirals (N=24), hydroxychloroquine (N=20), and polyclonal antibodies (N=15) (Table 1).

For phase 3 trials, the most common categories of treatments included monoclonal antibodies (N=13), antiviral medications (N=10), and chloroquine/hydroxychloroquine (N=9) (Table 2, Figure 2). Nearly all trials were of patients hospitalized with COVID-19 (N=45, 83%) and 80% (N=36) of the inpatient trials required patients to have hypoxia to be included. Of the inpatient studies, 27% of studies explicitly included patients who were critically ill (e.g., ICU, mechanical ventilation) (N=12 of 45). The median planned enrollment for the phase 3 trials was 465 (IQR 302–1029) (Table 2). The most common primary outcomes were symptom severity (N=23, 43%), a composite endpoint of mortality or ventilation (N=10, 19%), and mortality alone (N=6, 11%). For the three most common treatment categories (antivirals, monoclonal antibodies, chloroquine/hydroxychloroquine), the estimated power to...
Table 1 Trial Characteristics

| Trials | Median number of patients to be enrolled per trial (IQR) |
|--------|-------------------------------------------------------|
| All trials | 200 | 150 (60–400) |
| Completed | 9 | 450 (67–113) |
| Ongoing (anticipated numbers) | 191 | 150 (60–400) |

Table 2 Characteristics of Phase 2/3 and Phase 3 Trials (N=54)

| Trials | Median enrollment per trial | Overall planned enrollment |
|--------|----------------------------|---------------------------|
| Top ten study treatment arm categories | | |
| Antiviral | 10 | 782 | 23147 |
| Monoclonal antibodies | 13 | 450 | 16070 |
| Chloroquine/hydroxychloroquine | 9 | 500 | 9481 |
| Colchicine | 1 | 600 | 6000 |
| Antibiotic | 5 | 600 | 5365 |
| Anticoagulation | 4 | 762 | 4832 |
| ACE inhibitor | 2 | 2250 | 4500 |
| Polyclonal antibodies | 5 | 600 | 3720 |
| Other | 5 | 525 | 2919 |
| Small molecule immunomodulators | 7 | 358 | 2284 |

DISCUSSION

This study leverages the data infrastructure captured by clinicaltrials.gov to characterize the COVID-19 trials landscape approximately 9 months after the world’s first documented case of SARS-CoV-2. Of the 200 randomized trials identified, nearly half were single-center and almost one-third were not placebo-controlled. For a disease that surges and recedes in any one given location and without the possibility of historical controls given that this is a novel pathogen, these factors highlight important areas in which national clinical trial infrastructure and trial design might be optimized in this pandemic and beyond.

Perhaps the most crucial area in which the absence of a centralized national trials platform is most apparent is the proliferation of multiple underpowered trials instead of fewer, coordinated, multi-site studies. For example, there were four clinical trials of anticoagulation with a median planned sample size of about 760 patients each. Presuming a baseline mortality rate of 15%, the current trials each have 22% power to detect a 20% relative reduction in mortality. However, one centralized trial for anticoagulation would have had more than 4800 patients which could have provided 86% power to detect that 20% relative reduction in mortality (Table 3). While meta-analyses offer one strategy to pooling results across a number...
of smaller randomized trials in an attempt to draw broader conclusions, this approach can be challenging when trials use different endpoints. Such heterogeneity limits the ability of investigators to pool study results and to be able to rigorously account for disparate study designs and methodologies. Similar to the findings of Mathioudakis and colleagues, our analysis demonstrates that this was quite common: the primary outcomes across the registered clinical trials varied from changes in symptom severity to mortality. Utilization of pre-specified outcomes of interest that have been agreed upon by clinical and methodological experts and centralizing trials are two strategies to address this.

Our results also identify substantial geographic variations that could exacerbate the well-documented disparities within populations affected by COVID-19. Access to clinical trials may have only been available to patients in regions where academic medical centers had the resources to launch trials quickly or previously established interactions with pharmaceutical companies that would have facilitated rapid roll-out of industry-sponsored studies. Geographic disparities would be expected to widen without intentional effort to ensure access to trials in rural areas.

Commonly identified exclusion criteria documented here are also important to consider. More than one-quarter of all studies excluded older adults—the population with the highest COVID-19-associated mortality. The rate of exclusion of older adults in these trials mirrors exclusions that have been previously described in non-COVID-19 studies. Nearly three in four trials also excluded people who were pregnant, a common practice in non-COVID-19 trials as well. For example, early trials of the antiviral drug remdesivir excluded patients who were pregnant. However, the manufacturer

---

### Table 3: Statistical Power to Detect a Relative Risk Reduction in Mortality for the Most Common Treatment Arms Across Different Baseline Rates of Inpatient Mortality

|                     | Overall trial sample size | Assumed baseline risk of mortality | Estimated power across different potential relative reductions in mortality |
|---------------------|---------------------------|------------------------------------|------------------------------------------------------------------------------------------------|
|                     |                           |                                    | RR=0.60 | RR=0.70 | RR=0.80 | RR=0.90 |
| Antiviral           | Trials pooled             | 23,147                             | 15%      | 100%    | 100%    | 100%    | 90%     |
|                     | Current approach          | 782                                | 15%      | 74%     | 47%     | 23%     | 9%      |
| Monoclonal antibody | Trials pooled             | 16,070                             | 15%      | 100%    | 100%    | 100%    | 78%     |
|                     | Current approach          | 450                                | 15%      | 50%     | 29%     | 15%     | 7%      |
| Chloroquine/        | Trials pooled             | 9481                               | 15%      | 100%    | 100%    | 99%     | 55%     |
| hydroxychloroquine  | Current approach          | 500                                | 15%      | 54%     | 32%     | 16%     | 7%      |
| Antibiotic          | Trials pooled             | 5365                               | 15%      | 100%    | 100%    | 90%     | 35%     |
|                     | Current approach          | 600                                | 15%      | 62%     | 38%     | 19%     | 8%      |
| Anticoagulation     | Trials pooled             | 4832                               | 15%      | 100%    | 100%    | 86%     | 32%     |
|                     | Current approach          | 762                                | 15%      | 73%     | 46%     | 22%     | 9%      |

Legend: the calculations above assume 1 to 1 randomization (i.e., that half of the trial were randomized to the drug and half were randomized to standard of care or placebo) and an alpha of 0.05. RR, relative risk.
developed a so-called compassionate use program to make the drug available to pregnant people outside the context of a clinical trial. Excluding people who are pregnant from the trial while then administering the treatment to that population ensures that the generated data will be limited in their ability to inform future counseling on the risks and benefits to both the person who is pregnant and the fetus.

This study and similar work using data from clinicaltrials.gov is only possible because of the universal requirement that investigators pre-register clinical trials, a requirement that has only been enforced since 2007. Such transparency allows accountability and provides some critical metrics that can be used to assess improvement. Our analysis also suggests areas where increased focus on transparency and information gathering could be beneficial. For example, the funding source was listed on clinicaltrials.gov as “other” for more than half of the trials in this analysis, with no further detail available.
Additional details in this reporting may help promote funding transparency in biomedical research.

Our study has limitations. First, our results may not be generalizable to clinical trials outside of the US because our analysis did not include results from non-US registries. Second, our power calculations focused on planned rather than actual enrollment. Thus, our power calculations may overestimate the trial’s actual statistical power. Third, we based our power calculation on published event rates which is an imperfect approach representing population average estimates. As a result, individual studies may have had higher or lower event rates depending on the characteristics of the study populations. In addition, some trials did not perform 1 to 1 randomization and thus, some of our power estimates represent overestimates because other randomization schemes require larger sample sizes. Fourth, we utilized word tagging, a form of natural language processing, to identify various datapoints reported in our study which could introduce misclassification bias. We performed manual checking to minimize misclassification, and while a small degree of residual misclassification could still be present, we expect it to be random.

In conclusion, in the first 9 months of the global pandemic, there was an unprecedented rise in COVID-19-related research, including hundreds of clinical trials launched in the US. This study highlights an important opportunity to improve national clinical trial infrastructure, which would allow the biomedical community to better leverage the critical investment of trial participants who are putting themselves at risk to advance new knowledge in this pandemic.

Corresponding Author: Michael Fralick, MD, PhD, SM, FRCP; Lunenfeld-Tanenbaum Research Institute, Sinai Health System, 60 Murray Street, Toronto, ON M5T 3L9, Canada (e-mail: mike.fralick@mail.utoronto.ca).

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s11606-021-07167-9.

Author Contribution Study concept and design: All authors Acquisition of data: All authors Analysis/interpretation of data: All authors Drafting of the manuscript: Sacks, North, Fralick Critical revision of the manuscript: All authors Statistical analysis: All authors

Funding This study was funded by the Sinai Research fund.

Declarations:

Conflict of Interest: Drs. Sacks, North, Wolf, and Campbell report no conflicts of interest. Dr. Dougan has received consulting fees from Tillotts Pharma, Partner Therapeutics, ORIC pharmaceuticals, and Moderna; research funding from Novartis and Eli Lilly and is a member of the scientific advisory board of Neoleukin Therapeutics. Dr. Fralick is a consultant for ProofDx (previously Pine Trees Health), a start-up company developing a CRISPR-based diagnostic test for COVID-19.

1. Dong E, Du H, Gardner L. An interactive web-based dashboard to track COVID-19 in real time. Lancet Infect Dis. Published online 2020. doi:https://doi.org/10.1016/S1473-3099(20)30120-1
2. Andrew S. The US has 4% of the world’s population but 25% of its coronavirus cases. CNN. Published online June 2020. https://www.cnn.com/2020/06/30/health/us-coronavirus-toll-in-numbers-june-trnd/index.html
3. Bothwell LE, Greene JA, Podolsky SH, Jones DS. Assessing the Gold Standard — Lessons from the History of RCTs. N Engl J Med. 2016;375(22):2175-2181. doi:https://doi.org/10.1056/nejmsra1604593
4. Bothwell LE, Podolsky SH. The Emergence of the Randomized, Controlled Trial, N Engl J Med. 2016;375(6):501-504. doi:https://doi.org/10.1056/nejmg1604635
5. Teixeira da Silva JA, Tsigairis P, Ermansmanesh M. Publishing volumes in major databases related to Covid-19. Scientometrics. Published online 2020. doi:https://doi.org/10.1007/s11192-020-03675-3
6. Brainard J. Scientists are drowning in COVID-19 papers. Can new tools keep them afloat? Science. Published online 2020. doi:https://doi.org/10.1126/science.abc7839
7. Murray DD, Babiker AG, Baker JV, et al. Design and implementation of the multi-arm, multi-stage Therapeutics for Inpatients with COVID-19 (TICO) platform master protocol: An Accelerating COVID-19 Therapeutic Interventions and Vaccines (ACTIV) initiative. medRxiv. Published online 2020. doi:https://doi.org/10.1101/2020.11.08.20227876
8. Petal Network: Prevention and Early Treatment of Acute Lung Injury. PETAL Network. Published 2020. Accessed February 17, 2021. petalnet.org
9. Zhu N, Zhang D, Wang W, Li X, et al. A Novel Coronavirus from Patients with Pneumonia in China, 2019. New England Journal of Medicine. 2020 Feb 20;382(8):727–33.
10. Rothman KJ, Boice JD. Epidemiologic Analysis with a Programmable Calculator; 1979.
11. Miettinen OS. Individual Matching with Multiple Controls in the Case of All-or-None Responses. Biometrics. Published online 1969. doi:https://doi.org/10.2307/2528794
12. COVID-19 Data Repository by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University [Internet, accessed 2020 Aug 10]. Available from: https://raw.githubusercontent.com/CSSCEGISSandData/COVID-19-master/cese_covid_19_data/cese_covid_19_daily_reports_us/08-10-2020.csv
13. WHO Rapid Evidence Appraisal for COVID-19 Therapies (REACT) Working Group, Sterne JAC, Murthy S, Diaz JV, Shutsky AS, Villar J, et al. Association Between Administration of Systemic Corticosteroids and Mortality Among Critically Ill Patients With COVID-19: A Meta-analysis. JAMA. 2020;324(13):1330-41.
14. Mathioudakis AG, Fally M, Hashad R, Kouta A, Hadi AS, Knight SB, et al. Outcomes Evaluated in Controlled Clinical Trials on the Management of COVID-19: A Methodological Systematic Review. Life (Basel). 2020;10(12):E350.
15. The ‘meta-COS’ for research in COVID-19 hospitalised patients [Internet]. Available from: https://www.covidcometinitiative.org/assets/downloads/COVID-19%20meta-COS%20Table%2015th%20March%202021.pdf
16. Disparities in Deaths. Centers for Disease Control and Prevention. Published October 12, 2020. https://www.cdc.gov/coronavirus/2019-ncov/community/health-equity/racial-ethnic-disparities/disparities-deaths.html
17. Hamaker ME, Stauder R, Munster BC. Exclusion of Older Patients From Ongoing Clinical Trials for Hematological Malignancies: An Evaluation of the National Institutes of Health Clinical Trial Registry. The Oncologist. Published online 2014. doi:https://doi.org/10.1634/theoncologist.2014-0093
18. Nanna MG, Chen ST, Nelson AJ, Navar AM, Peterson ED. Representation of older adults in cardiovascular disease trials since the inclusion across the lifespan policy. JAMA Intern Med. Published online 2020. doi:https://doi.org/10.1001/jama.2020.2750
19. Blehar MC, Spong C, Grady C, Goldkind SF, Sahin L, Clayton JA. Enrolling Pregnant Women: Issues in Clinical Research. Women’s Health Issues. 2013;23(3):1-1. doi:https://doi.org/10.1016/j.whi.2012.10.003
20. Beigel JH, Tomashek KM, Dodd LE, et al. Remdesivir for the Treatment of Covid-19— Final Report. N Engl J Med. 2020;383(19):1813-1826. doi:https://doi.org/10.1056/nejmsra2007764
21. Goldman JD, Lye DGB, Hui DS, et al. Remdesivir for 5 or 10 Days in Patients with Severe Covid-19. N Engl J Med. 2020;383(19):1827-1837. doi:https://doi.org/10.1056/nejmoa2015301
22. Gilead. Gilead Sciences Statement on Access to Remdesivir Outside of Clinical Trials. Published online 2020. https://www.gilead.com/news-and-press/company-statements/gilead-sciences-statement-on-access-to-remdesivir-outside-of-clinical-trials

23. Wood AJJ. Progress and Deficiencies in the Registration of Clinical Trials. *N Engl J Med*. Published online 2009. doi:https://doi.org/10.1056/nejmsr0806582

**Publisher’s Note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.