Producing and using timely comparative evidence on drugs: lessons from clinical trials for covid-19

Huseyin Naci and colleagues call for greater collaboration among trialists, meta-analysts, and guidance developers to improve the evidence base for new treatments

Huseyin Naci, 1 Aaron S Kesselheim, 2 John-Arne Røttingen, 3, 4 Georgia Salanti, 5 Per O Vandvik, 6 Andrea Cipriani 1, 7

Since the early days of the novel coronavirus outbreak, a record number of studies have been launched to test several repurposed and new medicines as potential treatments for covid-19. 1 An analysis by the news organisation STAT identified over 1000 clinical trials registered on ClinicalTrials.gov between January and June 2020. 2 This is a testament to the research and clinical community’s commitment to identify effective treatments for covid-19. However, the large volume of studies may paradoxically limit the generation of robust evidence and complicate the formulation of trustworthy guidance and decisions related to drug use if the current research is replicative and redundant or produces conflicting data. 3–5 Indeed, the multiplicity of research on candidate therapeutics for covid-19 has exposed important flaws and failures in the current evidence ecosystem. 6–7 Crucially, these limitations also affect the full spectrum of research on new health technologies. 8 9

Users of evidence across the healthcare system (patients, clinicians, health technology assessment bodies, guideline developers, payers) need timely data on how different treatments compare with each other in terms of their benefits and harms—their comparative effectiveness. Producing comparative evidence and ensuring its rapid translation into trustworthy guidance requires extensive coordination and collaboration between the researchers conducting clinical trials, those conducting comparative effectiveness assessments, and those producing guidance. 8 9 The experience of covid-19 highlights the difficulties in making comparative assessments and suggests areas for improvement.

Limitations of covid-19 research

Three main limitations have characterised the system for evaluating repurposed or investigational therapeutics for covid-19. Firstly, global clinical research activity is fragmented. The drug trials rarely have similar design features. For example, study endpoints have been shown to be highly heterogeneous 10 and few of the late stage randomised trials measure all-cause mortality. 11 Even when randomised trials evaluate seemingly similar endpoints such as time to clinical recovery, outcome definitions and follow-up durations vary.

Secondly, the research agenda seems to be partly driven by hype and anecdote rather than informativeness and social value, 12 skewing the amount of available data. For example, a disproportionately large number of studies were launched to evaluate the antimalarial drugs hydroxychloroquine and chloroquine phosphate after the publication of a controversial uncontrolled study that received substantial attention. 13 About one in every six studies registered on ClinicalTrials.gov has focused on these antimalarial agents. 2

Thirdly, studies have not routinely adopted robust designs. We estimate that fewer than one third of studies evaluating covid-19 therapeutics on ClinicalTrials.gov are randomised controlled trials, which are the gold standard for evaluating treatments. 14 Many studies test investigational agents without a control group, 15 which can be misleading as they provide no data on what would have happened in the absence of the treatment.

The combination of these factors hasfuelled confusion and sensationalism. Psychological distress and anxiety have increased in the general population. 16 Findings of individual studies are watched closely and with suspense. Doing “science by press release”—publicising study findings before they are shared as preprints or published in peer reviewed journals—has become common. Healthcare professionals have not been immune to hype. During the early days of the pandemic, shortages of hydroxychloroquine were reported, driven by clinicians’ prescriptions after these products were hailed as potential breakthroughs.

Even regulators have been under pressure to act without sufficient evidence. 17 18 In the US, the Food and Drug Administration granted emergency use authorisation for hydroxychloroquine without any solid data suggesting that it was effective in covid-19. The FDA later revoked this authorisation when randomised trials found no benefits. The European Medicines Agency granted a conditional marketing authorisation for remdesivir on the basis of “non-comprehensive” data and without access to clinical study reports. 19

Progress on research coordination and collaboration

Mechanisms already exist for global research coordination during public health emergencies. Initiatives such as the Global Research Collaboration for Infectious Disease Preparedness (GloPID-R), 20 established in 2013 after agreement by the heads of international (biomedical) research funding organisations, and the World Health Organization’s research and development blueprint, 21 which was
developed after the Ebola outbreak in 2014-16, are platforms for collaboration. New models are also emerging. The G20 countries and WHO have established the Access to Covid-19 Tools (ACT) Accelerator, a global collaboration to accelerate the development, production, and equitable access to new diagnostics, therapeutics, and vaccines.22

These efforts have already paid off. Several large randomised trials have been launched in record speed. Many of these compare multiple treatments simultaneously. Three of the largest “mega”-trials—the Solidarity trial led by WHO, Discovery initiated by Inserm in France, and Recovery trial in the UK—have comparable protocols (including their simple, pragmatic, and adaptive designs) and collect data on similar endpoints (including death and need for ventilation). The Recovery trial has recruited over 13 500 patients, accounting for 15% of those admitted to hospital with covid-19 across the UK.23 Some of the most important insights about candidate therapeutics have emerged from Recovery, including the meaningful survival benefit associated with using dexamethasone in severely ill patients. The Solidarity trial, of which Discovery is an add-on, has included more than 7000 patients across more than 20 countries from different regions of the world and is the largest trial that can now follow the pandemic where it is globally most active.

However, efforts to date have not managed to avoid research waste and ensure that all relevant studies contribute to the formulation of guidance and decisions in practice and policy.24 Most studies on covid-19 treatments have methodological limitations (eg, small samples and diverse designs and outcomes).26 A sizeable portion of studies, collectively including thousands of patients, may therefore have little prospect of adding to the growing evidence base on efficacy.

### Areas for improvement

Determining the comparative effectiveness of drugs requires streamlining the design, analysis, reporting, and data sharing practices of clinical studies. These objectives are not new but progress towards achieving them has been slow.25-27,28 Despite several large multiarm trials, most research on covid-19 therapeutics is not fit for generating comparative evidence. We outline five priorities for greater collaboration and coordination among trialists, meta-analysts, guideline developers, and other stakeholders to facilitate producing and using trustworthy comparative evidence and guidance (table 1). These are also relevant to studies evaluating other types of interventions, including supportive care and non-drug interventions.

<table>
<thead>
<tr>
<th>Domain</th>
<th>Recommendation</th>
<th>Next steps</th>
<th>Target stakeholders</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Selecting treatments to include in large trials</strong></td>
<td>Use network meta-analyses to learn from the fast-evolving body of evidence and reach conclusions on which treatments to test in larger trials</td>
<td>• Complement WHO’s therapeutic landscape analysis with network meta-analyses of available data</td>
<td>WHO, Research funders (including industry) and meta-analysts</td>
</tr>
<tr>
<td><strong>Streamlining trial designs</strong></td>
<td>Ensure that future trials collect data on core outcome measures</td>
<td>• Streamline core outcome sets • Encourage (or mandate, where possible) use of core outcome sets in trials • Develop and communicate regulatory and health technology assessment guidance on preferred trial designs</td>
<td>WHO, Research funders (including industry), Ethics committees, Regulatory agencies, Health technology assessment bodies, (eg, NICE) Research community (trialists and meta-analysts)</td>
</tr>
<tr>
<td><strong>Sharing data</strong></td>
<td>Share individual participant data in a timely manner</td>
<td>• Monitor and publicly report data sharing practices of all trial sponsors • Incentivise (or mandate, where possible) data sharing • Prioritise sharing trial data with researchers planning to conduct comparative effectiveness assessments</td>
<td>WHO, Research funders (including industry sponsors), Ethics committees, Regulatory agencies, Academic institutions, Research community (trialists and meta-analysts)</td>
</tr>
<tr>
<td><strong>Assessing comparative effectiveness</strong></td>
<td>Pre-plan and conduct individual participant data network meta-analyses shortly after trial completion</td>
<td>• Minimise (or eliminate, if possible) duplication across multiple groups planning comparative effectiveness assessments • Develop a consortium of researchers, regulatory agencies, health technology assessment bodies, and payers, convened by WHO • Prospectively design network meta-analyses in collaboration with trialists to ensure timely availability of results shortly after trial completion</td>
<td>WHO, Research funders (including industry sponsors), Ethics committees, Regulatory agencies, Academic institutions, Non-profit organisations (eg, Cochrane), Research community (trialists and meta-analysts)</td>
</tr>
<tr>
<td><strong>Transforming data into guidance to inform policy and practice</strong></td>
<td>Use comparative evidence to generate rapid, living, and trustworthy guidance</td>
<td>• Use interoperable platforms to digitally structure comparative data for rapid dissemination of recommendations and development of interactive evidence summaries</td>
<td>WHO, Interoperable evidence platforms (eg, MAGICapp), Guideline developers, Health technology assessment bodies, Non-profit organisations (eg, Cochrane), Scientific journals and publishers</td>
</tr>
</tbody>
</table>
Selecting treatments to include in large trials

Key trials differ in which treatments they included (table 2), reflecting a lack of consensus on the most promising therapeutic candidates. Therefore, treatment selection even in large trials has not been fully complementary. For example, hydroxychloroquine was included in both Recovery and Solidarity. By contrast, dexamethasone, the first drug shown to improve survival in hospital patients in the Recovery trial, was not included in some other “mega” trials.

<table>
<thead>
<tr>
<th>Trial name</th>
<th>Primary sponsors</th>
<th>Comparators (up to 12 October 2020)</th>
<th>Target sample size</th>
</tr>
</thead>
<tbody>
<tr>
<td>GS-US-540-5774</td>
<td>Gilead</td>
<td>Remdesivir v standard of care</td>
<td>1600</td>
</tr>
<tr>
<td>GS-US-540-5773</td>
<td>Gilead</td>
<td>Remdesivir v standard of care</td>
<td>6000</td>
</tr>
<tr>
<td>Solidarity</td>
<td>WHO</td>
<td>Remdesivir Lopinavir/ritonavir (discontinued) Lopinavir/ritonavir+interferon beta-1a (discontinued) Hydroxychloroquine (discontinued) Standard of care</td>
<td>No specific sample size (target of several thousand participants)</td>
</tr>
<tr>
<td>Discovery</td>
<td>Institut National de la Santé Et de la Recherche Médicale, France</td>
<td>Remdesivir Lopinavir/ritonavir (discontinued) Lopinavir/ritonavir+interferon beta-1a (discontinued) Hydroxychloroquine (discontinued) Standard of care</td>
<td>3100</td>
</tr>
<tr>
<td>Recovery</td>
<td>UK Research and Innovation, UK National Institute for Health Research (NIHR)</td>
<td>Lopinavir-ritonavir (discontinued) Dexamethasone (continuing only in children) Hydroxychloroquine (discontinued) Azithromycin Tocilizumab Convalescent plasma REGN-COV2 Standard of care</td>
<td>13 500</td>
</tr>
<tr>
<td>Principle</td>
<td>UK Research and Innovation, UK NIHR</td>
<td>Azithromycin v standard of care Doxycycline v standard of care</td>
<td>3000</td>
</tr>
<tr>
<td>Remap-Cap</td>
<td>Canadian Institutes for Health Research, European Commission, UK NIHR, Health Research Council of New Zealand, Australian National Health and Medical Research Council</td>
<td>Antibiotics Antivirals Host immunomodulation with extended macrolide therapy Corticosteroid regimens (discontinued) Hydroxychloroquine Hydroxychloroquine+lopinavir/ritonavir Interferon beta-1a Anakinra Tocilizumab Sarilumab</td>
<td>7100</td>
</tr>
</tbody>
</table>

The Discovery trial is nested in Solidarity as a “daughter” trial, allowing for more detailed data collection than its more simple and larger “mother” trial.

Evidence based approaches to select treatments are emerging. For example, the UK has launched Accord (Accelerating Covid-19 Research and Development), which is an adaptive platform study comprising almost 50 small randomised trials of candidate drugs for further testing in Recovery. In addition to conducting such de novo trials, evidence synthesis methods would provide an opportunity to learn from a fast evolving body of research. Network meta-analyses could reach conclusions on which treatments to test in larger trials more efficiently than other approaches.

Streamlining trial designs

Harmonising the outcome measures used in different trials is a prerequisite for their inclusion in comparative effectiveness assessments. Users of evidence have a key role in defining and prioritising outcome measures. There is some consensus that all-cause mortality and respiratory support are the preferred core outcomes in the severe stages of covid-19. However, the availability of several core outcome sets has complicated efforts to streamline trial designs.

Ensuring that future trials collect data on one set of core outcomes will require collaboration from diverse stakeholders. WHO has convened experts for the development of model protocols, clinical reporting forms, and endorsing a set of core outcomes that are relevant to different stages of the disease (pre-exposure prophylaxis, post-exposure prophylaxis, early treatment, hospital admission, intensive care, post-hospital) and may span across different areas of medicine (for instance, long term effects of covid-19 include medical, psychological, and rehabilitation needs).
Research funders, ethics review boards, and clinical trial approval authorities should require inclusion of core outcomes in protocols. Streamlining regulatory and health technology assessment guidance across different settings would also help. In its conditional marketing authorisation of remdesivir in June 2020, EMA acknowledged the lack of “regulatory guidance or precedent specifying a particular preferred primary endpoint” for covid-19 therapeutics. The FDA, EMA, and health technology assessment bodies should produce joint guidance and provide parallel advice on the trial protocols of candidate therapeutics.

Sharing data
The benefits of timely access to data from clinical trials are widely accepted. Such data could be re-analysed and combined with data from other studies to determine comparative effectiveness. Individual participant data could also identify subgroups of patients with different responses to treatments, exploring characteristics that modify effectiveness and thus explain contradictory findings. While data sharing after trial completion is becoming more common, and several funders of health research are committed to this goal, data sharing is still not the norm. According to ClinicalTrials.gov, Gilead has no plans to release individual participant data from its phase III trials of remdesivir (NCT04292730 and NCT04292899).

Sponsors’ transparency and data sharing practices should be periodically monitored and publicly reported. Academic institutions should make data sharing an explicit criterion for promotion and tenure. All trial sponsors, including industry, should pledge to share data rapidly through one of the existing platforms (eg, Infectious Diseases Data Observatory). Requests for data after trial completion and publication are associated with poor retrieval rates in meta-analyses. Therefore, data sharing plans and agreements should be finalised in advance. Ideally, data sharing should accompany trial publication. When this is not feasible, data sharing should be prioritised for groups or institutions with plans to conduct comparative effectiveness assessments. New models of data sharing could also improve trial efficiency. For example, sharing real-time data across ongoing trials could allow early identification of efficacy and safety signals. However, such practices may override the integrity of individual trials and should therefore be agreed in advance and reflected in protocols.

Assessing comparative effectiveness
No single trial can compare the efficacy of all potential treatments for covid-19. Inevitably, indirect comparisons across trials will generate evidence on the comparative benefits and harms of different products. Several groups are working in parallel to identify trials and pool results in network meta-analyses as they emerge. Such “living” syntheses could provide useful evidence, but even small differences in study eligibility criteria and analytical strategies may yield conflicting results, which may delay the development of trustworthy guidance. It is therefore essential to coordinate ongoing activities, pool resources across groups, and minimise duplication.

A consortium should coordinate the design, implementation, and replication of comparative effectiveness assessments, ideally using individual participant data network meta-analyses. A network of leading independent research organisations, regulatory agencies, health technology assessment bodies, and payers could lead this effort in collaboration with WHO. A recent health technology assessment of biological agents for rheumatoid arthritis in Germany has shown the feasibility of this approach. The Institute for Quality and Efficiency in Health Care (IQWiG) requested re-analysis of individual participant data from several industry sponsored randomised trials to harmonise patient populations and primary endpoints before findings could be combined in network meta-analyses. Prompt access to comparative data is critical. As there is an ethical imperative for any treatment with promising results to immediately become the new standard of care (as occurred with dexamethasone, and to a lesser extent remdesivir, in patients with severe covid-19), comparative assessments should ideally accompany the publication of individual trial results. This would allow individual study results to be interpreted within their broader context and greatly increase speed in updating guidance for policy and practice.

Prospectively designing comparative effectiveness assessments would balance speed with rigour. Pre-planning of network meta-analyses requires close collaboration between trialists and meta-analysts. At a minimum, data from the trials with the most robust designs should be shared with third party researchers to conduct prospectively designed network meta-analyses. Such close collaboration would ensure that data completeness, standardisation, and quality issues are resolved promptly, and the results can be re-analysed and combined shortly after the database is locked.

Translating data into guidance
Covid-19 has created an unprecedented need for living and trustworthy guidance based on comparative evidence. Recent experience with Australia’s National Covid-19 Clinical Evidence Taskforce shows how a comprehensive set of recommendations can be dynamically updated based on new evidence, facilitated by innovative processes and digitally structured data in interoperable platforms (eg, MAGICapp). Such platforms allow for immediate global dissemination of recommendations, interactive evidence summaries, and decision aids that can be reused, adapted, or implemented. WHO and prominent guideline development organisations are now moving towards producing such living guidance for covid-19. Some are dedicated to sharing evidence and recommendations in a globally concerted effort, aiming for three weeks from evidence to publication.

The BMJ’s Rapid Recommendations entry on remdesivir shows how such global collaboration and iterative guidance development work, informed by living network meta-analysis. WHO living guidance on drugs for covid-19 is developed in a similar way, first published for corticosteroids. The guideline panel convened and created recommendations for corticosteroids two days after unpublished data from Recovery was made available through a prospective meta-analysis. This shows the value of close collaboration between trialists, meta-analysts, and guideline developers. Global dissemination of WHO guidance was delayed for six weeks, however, as it had to wait for publication of Recovery results in a scientific journal, underscoring remaining challenges.

Conclusions
The evidence based medicine movement has for decades challenged the primacy of individual studies. No single study can provide adequate evidence to inform the variety of therapeutic decisions in clinical practice. Information on the comparative benefits and harms of alternative treatments is imperative and is often best obtained from a synthesis of several studies. Producing and using timely, trustworthy, and actionable evidence and guidance requires designing, analysing, and reporting each study in a way that optimises its contribution to subsequent comparative effectiveness assessments. Progress to date has been too slow. However, covid-19 highlights the pressing need and the opportunity to harness new collaborations among relevant stakeholders, including trialists.
meta-analysts, guidance developers, regulatory agencies, health technology assessment bodies, and payers.

Key messages

- The record number of studies evaluating the effectiveness of repurposed and investigational drugs for covid-19 has exposed important shortcomings in the evidence ecosystem
- Despite the availability of several large multi-arm trials, evidence on the comparative effectiveness of potential therapeutic alternatives has been delayed
- Heterogeneity of trial design and outcomes makes comparison difficult
- Producing comparative evidence on covid-19 therapeutics and ensuring its rapid translation into trustworthy guidance will require greater coordination among trialists, meta-analysts, and other stakeholders

Contributors and sources. HN and AC devised the idea for this article. HN developed the first draft and all authors contributed to the writing of subsequent versions. HN is the guarantor.

Competing interests: POV is co-founder and chief executive of the non-profit MAGIC Evidence Ecosystem Foundation ([www.magicevidence.org](http://www.magicevidence.org)) and receives a fixed part-time salary for his work. AC has received research and consultancy fees from INCiPiT (Italian Network for Paediatric Trials), Cariplo Foundation, and Angelini Pharma.

Provenance and peer review: Not commissioned; externally peer reviewed.

AC is supported by the National Institute for Health Research (NIHR) Research Professorship (grant RP-2017-08-572-006), by the NIHR Oxford cognitve health Clinical Research Facility, by the NIHR Oxford and Thames Valley Applied Research Collaboration, and by the NIHR Oxford Health BioMedical Research Centre (grant BRC-1275-20055). ASK’s research is supported by Arnold Ventures and the Harvard-MIT Center for Regulatory Science. GS’s research is supported by the Swiss National Science Foundation (Grant No 203030_179518). The views expressed are those of the authors and not necessarily those of the UK National Health Service, the NIHR, or the UK Department of Health.
