From intent to impact: Investigating the effects of open sharing commitments

Prepared on behalf of Wellcome, UK Research and Innovation and the Bill & Melinda Gates Foundation

June 2022

www.research-consulting.com

From intent to impact: Investigating the effects of open sharing commitments

Commissioned by

Wellcome, UK Research and Innovation and the Bill & Melinda Gates Foundation

Contact

Hannah Hope

Open Research Lead, Research Environment

Wellcome

E: h.hope@wellcome.org

Authors

Andrea Chiarelli, Lucia Loffreda, Eleanor Cox, Rob Johnson (Research Consulting)

Christine Ferguson (Independent consultant)

W: www.research-consulting.com

E: andrea.chiarelli@research-consulting.com

Bibliometric analysis

Etienne Vignola-Gagné, David Campbell, Attila Emecz, on behalf of the Science Metrix/Elsevier team (Science Metrix, an Elsevier company)

W: https://science-metrix.com/

E: e.vignola-gagne@elsevier.com

Report dated: June 2022

Supporting data

Supporting data and code are available at: https://zenodo.org/communities/data-sharing-in-public-health-emergencies/

Contents

1. Background and methodology

1.1 Introduction

Background and rationale

In light of the emerging COVID-19 pandemic, in January 2020, Wellcome coordinated a Joint Statement which called on researchers, journal publishers, and funders to "ensure that research findings and data relevant to this outbreak are shared rapidly and openly to inform the public health response and help save lives".[1] This follows a long-standing ambition of Wellcome that 'knowledge and discoveries are shared, accessed and used in a manner that maximises health benefit', and built on previous efforts during the Ebola and Zika outbreaks.

Wellcome, UK Research and Innovation (UKRI) and The Bill & Melinda Gates Foundation commissioned this study to understand the impact of these statements and if the resulting increased requirements for rapid and open sharing during the COVID-19 pandemic could be measured and attributed.

The five commitments for signatories

The Joint Statement outlines five key requirements for signatories, who committed to work together to ensure that:

- all peer-reviewed research publications relevant to the outbreak are made immediately open access, or freely available at least for the duration of the outbreak
- research findings relevant to the outbreak are shared immediately with the WHO upon journal submission, by the journal and with author knowledge
- research findings are made available via preprint servers before journal publication, or via platforms that make papers openly accessible before peer review, with clear statements regarding the availability of underlying data
- researchers share interim and final research data relating to the outbreak, together with protocols and standards used to collect the data, as rapidly and widely as possible - including with public health and research communities and the WHO
- authors are clear that data or preprints shared ahead of submission will not pre-empt its publication in these journals

A parallel initiative was also supported by Wellcome, targeting publishers and focusing on a commitment to make COVID-19 and coronavirus-related publications, and the available data supporting them, immediately accessible in PubMed Central (PMC) or other public repositories.[2] This initiative statement followed a request from National Science and Technology Advisors from a dozen countries, who emphasised the value of scientific research and innovation in addressing global public health emergencies.[3] The present report did not directly investigate this publisher-focused initiative, as it was seen as one of several potential pathways to operationalise the broader commitments in the Joint Statement.

Overview of project objectives: Attribution, Impact and Longevity

The aim of this study was to understand the extent to which the commitments made by the various stakeholders were put into practice, and to determine the downstream impacts (if any) which arose as a consequence of the Joint Statement. This study focuses on the three objectives outlined in Figure 1.

Figure 1. Project objectives

Attribution	Collate the evidence on signatories honouring the commitments made in the 2020 COVID-19 results and data sharing statement
Impact	Determine the views of public health response organisations on the impact of the Joint Statement resulting in results and data being available
Longevity	Using the 2016 Zika Statement as an example, collate evidence on whether statements create long-term shifts in open research practices and behaviours

1.2 Methodology

Overview of methodology

RQ ##

We based our analysis on a theory of change (see Figure 2 on page 8) that investigates the objectives in Figure 1 and was built in collaboration with the project funders and advisory group. A theory of change approach to evaluation has been defined as 'a systematic and cumulative study of the links between activities, outcomes, and contexts of the initiative' (Connell and Kubisch, p.2, 1998).[4] This evaluative approach aims to understand not only *whether activities* produce effects but also *how* and *why* these effects occur.

The theory of change in Figure 2 provides a map of specific hypotheses that we sought to validate in the present report (see "RQ#" in Figure 2) via a mix of qualitative and bibliometric approaches and a process of evidence triangulation. Over the course of the report, references to the theory of change can be found in square boxes in the left-hand side column, under paragraph titles (see example under the present paragraph title); these boxes are shaded in red, amber or green to describe the estimated strength of the evidence we gathered (low, medium and high, respectively).

Furthermore, we followed the precepts of contribution analysis to assess whether any changes in open research behaviours during the pandemic could be attributed to the Joint Statement.[5] In particular, we sought to answer the following questions, with a particular focus on signatory organisations:

- Has the Joint Statement made an important contribution to the observed changes in open research behaviours and policies?
- Why have the changes in open research behaviours and policies occurred? What role did the Joint Statement play?
- Is it reasonable to conclude that the Joint Statement has made a difference?
- What conditions are needed to enable this type of Joint Statement to make a difference in future?

From a methodological standpoint, our qualitative analysis included:

- the review of over 130 literature sources and thematic coding of a subset of 70 sources;
- the delivery of an online survey of signatories, which received 55 responses (34% of signatories); and
- the delivery and thematic coding of 16 stakeholder interviews.

Literature for inclusion in this study was identified via web-based research, using targeted keyword searches. An initial set of search strings was developed to identify literature focusing on the signature of the Joint Statement:

- ("signs" OR "signed" OR "statement" OR "signatory" OR "signatories") AND ("Wellcome" AND "COVID")
- ("statement" AND "Wellcome" AND "COVID") AND ("signs" OR "signed" OR "signatory" OR "signatories")

Additional keywords that were deemed directly relevant to the five commitments outlined in the Joint Statement were appended to searches, including: "COVID-19", "data sharing", "open access publishing", "open science", "preprints". Searches were supplemented by snowball sampling,[1] particularly as relevant literature emerged rapidly as the pandemic unfolded.

Further information on the bibliometric analysis is available in the technical report prepared by Science-Metrix (Annex A) and available via Zenodo. Full details on the bibliometric results and methodology, including data and code, are available in the project's Zenodo community.[6]

Report structure

Following this introduction, this report is divided as follows:

- Part A: Developing an open sharing statement in a pandemic
  - Section 2. Making the Joint Statement a reality
  - Section 3. Gathering signatures and communicating alignment
- Part B: Assessing the implementation and impact of the open sharing statement
  - Section 4. Attribution: Honouring the commitments in the Joint Statement
  - Section 5. Impact: Assessing the effects of the Joint Statement
  - Section 6: Longevity: Assessing long-term effects
- Part C: Building on the Joint Statement to address future crises
  - Section 7: Conclusions
  - Section 8: Lessons learned and recommendations

Over the course of the report, bibliometric findings are presented in dedicated boxes shaded in light brown. Anonymised quotes from contributors to the present study are presented between paragraphs and in pink colour. Finally, a Glossary is presented at the end of the report, to clarify the language used and any technical terminology.

Limitations

The present report is subject to the following limitations:

- Our literature review was designed to provide an informed conclusion on the volume and characteristics of the evidence base and a synthesis of what that evidence indicates in relation to the research questions. It did not include a critical appraisal of that evidence.
- Study participants were recruited via convenience sampling, that is, we interviewed stakeholders who were both available and willing to communicate. A total of 16 interviewees contributed to this study (see Appendix A).

[1] Snowballing refers to using the references in publications or reports to identify additional documents that may be of interest.

- 160 Joint Statement signatories were contacted to contribute to our online survey, and 34% of these provided a response. It is possible that the views of non-responding signatories materially differ from those presented in our report.
- Our analysis of qualitative data (literature, interview transcripts and survey responses) is underpinned by thematic coding, which relies on an extent of subjective interpretation.

In addition, we note the following limitation to the bibliometric analysis, which are explored in more detail in Annex A:

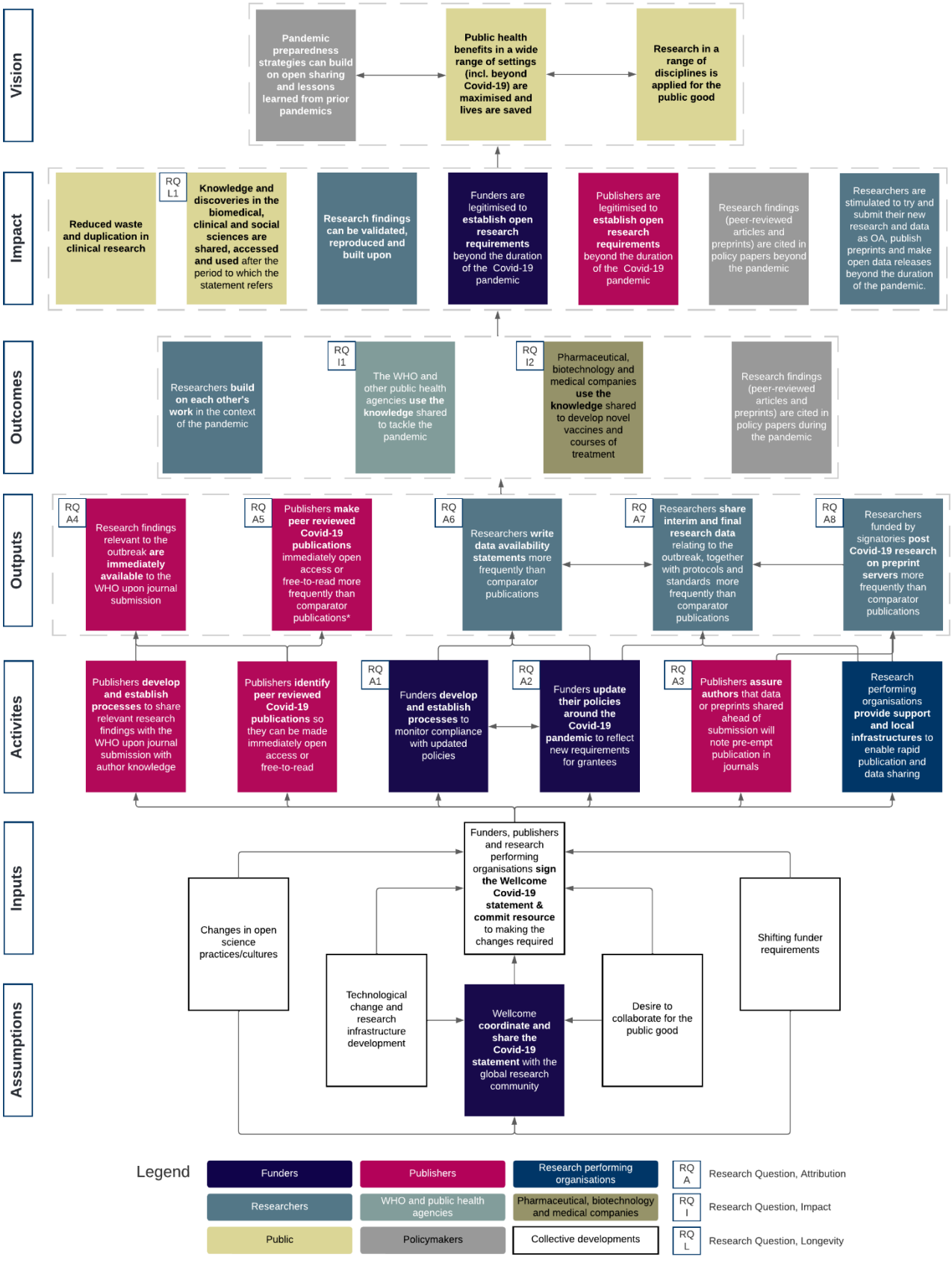
- The difference-in-differences design can control for many disciplinary, cultural, and author-level factors that might affect findings. It can also control for durable features or practices of signatory organisations and the researchers they support. Yet it cannot differentiate between the specific effects of the Joint Statement and new practices or initiatives by signatory organisations taken in response to the pandemic but independently of the Statement.
- In multiple metadata and full text processing steps such as isolating thematic sets of publications and preprints or retrieving mentions towards data sharing repositories, the study relied on manually curated lists of keywords and expressions. These queries are characterised by high precision (low share of false positives) but somewhat lower recall (somewhat higher share of false negatives, that is, imperfect coverage and representativity).
- Gaps in metadata and full text records in the datasets meant that signatory or non-signatory status could not be inferred for large portions of COVID-19 journal publications and preprints. In turn, this diminishes the representativity of our findings. To mitigate this issue, findings have been computed and provided separately for these publications and preprints of unknown signatory status (see Annex A).
- Availability of full text records, in a format workable for text mining to identify data availability statement and data deposition mentions within, varied between publishers, including due to licensing issues (see Annex A).
- Control groups for COVID-19 research, made up of "human viral respiratory diseases" (HVRD) journal publications and preprints, often contained significantly lower numbers of available observations, which has restricted the availability of robust difference-in-differences findings for some indicators.
- Coverage of preprint servers was limited to arXiv, bioRxiv, medrxiv, and SSRN.

Acknowledgements

This report was commissioned by Wellcome, UK Research and Innovation and the Bill & Melinda Gates Foundation. We gratefully acknowledge:

- the guidance and support received from Hannah Hope, Sabah Adams (Wellcome), Rachel Bruce, Callum Day (UKRI), and Ashley Farley (Bill & Melinda Gates Foundation);
- the expert advice of an Advisory Group including Alice Norton (University of Oxford), Elizabeth Gadd (University of Glasgow), Katharina Lauer (ELIXIR Europe), Ludo Waltman (Leiden University) and Robert Terry (TDR - World Health Organization);
- the 16 contributors listed in Appendix A and the 55 signatory organisations that submitted a response to our online survey; and
- the support and input, including provisions for an extraordinary open data release, provided by Euan Adie (Overton) to assess citations of journal publications and preprints in policy documents.

Figure 2. Theory of change and research questions



Part A. Developing an open sharing statement in a pandemic

Research Consulting
Science-Metrix

2. Making the Joint Statement a reality

2.1 Developing the Joint Statement

The Joint Statement was developed by Wellcome by updating the text of previous statements

The Joint Statement was developed internally by Wellcome, by updating the principles set out in the 2016 Statement on data sharing in public health emergencies, prompted by the Zika Virus outbreak (see Table 1), and the later 2018 Statement in response to the Ebola Virus.[7] We note that the 2016 Zika statement was informed by the WHO Consensus statement, while the principles developed by the GloPID-R working group on data sharing during public health emergencies influenced the Joint Statement.[8]

The commitments of the Joint Statement sought to reaffirm and amplify requirements around open sharing during previous public health emergencies and to reinforce the importance of sharing timely and relevant data with the World Health Organization to support the global public health response.

Interviewees recognised that the prompt release of the Joint Statement at the beginning of the pandemic was a key driver of impact. If the Joint Statement had been published later on, its influence on organisational responses to the pandemic would have been limited: higher education and research stakeholders might have chosen to align with another initiative (see Table 2 and Figure 3 below, for examples) or simply made independent decisions based on internal drivers.

"I was very much aware of the fact that Wellcome had put out the Joint Statement at the beginning of the pandemic, as a reminder that we all needed to share what we had. And a good case in point, of course, was the speed with which the coronavirus original sequence needed to be published, because without that we couldn't have got any of the vaccine production underway." – Academic

Table 1. Comparison between the COVID-19 Statement and the Zika Statement (key difference on the time horizon of the commitments in **gold colour**)

Commitments in the COVID-19 Statement	Commitments in the Zika Statement
• all peer-reviewed research publications relevant to the outbreak are made immediately open access, or freely available at least for the duration of the outbreak • research findings relevant to the outbreak are shared immediately with the WHO upon journal submission, by the journal and with author knowledge • research findings are made available via preprint servers before journal publication, or via platforms that make papers openly accessible before peer review, with clear statements regarding the availability of underlying data • researchers share interim and final research data relating to the outbreak, together with protocols and standards used to collect the	• journal signatories will make all content concerning the Zika virus free to access. Any data or preprint deposited for unrestricted dissemination ahead of submission of any paper will not pre-empt its publication in these journals • funder signatories will require researchers undertaking work relevant to public health emergencies to set in place mechanisms to share quality-assured interim and final data as rapidly and widely as possible, including with public health and research communities and the World Health Organization.

data, as rapidly and widely as possible – including with public health and research communities and the WHO
- authors are clear that data or preprints shared ahead of submission will not pre-empt its publication in these journals

Wellcome led the development of the Joint Statement alongside prospective signatories

The Joint Statement was developed rapidly within Wellcome and refined collaboratively with prospective signatories. Upon sharing the Joint Statement with organisations who previously signed the Zika Statement, the commitments for the COVID-19 version were met with broadly positive responses. In a few cases, changes to the wording of some commitments were suggested by signatories. Overall, the principles in the Joint Statement were shared by prospective signatories, with just one organisation that was approached by Wellcome declining to sign.

2.2 The broader landscape of initiatives

The Joint Statement was not alone in encouraging open sharing behaviours

Another important factor that is likely to have affected the research landscape during the pandemic is the release of other open sharing statements or commitments. Table 2 and Figure 3 provide examples of these, including cases where signatories would have overlapped to an extent with those of the Joint Statement. It should be noted that Wellcome's initiative to support open access and data sharing via PubMed Central (see section 1.1) is not reflected in Table 2, as it overlaps with the Public Health Emergency COVID-19 Initiative.

The efforts in Table 2 are unlikely to have presented a confounding effect in our stakeholder interviews, as our focus on the Joint Statement was clearly outlined to all contributors; however, a minority of these might have had an impact on research and publishing practices, and therefore on the bibliometric analysis. We note that it was not possible to control for the impacts of these concurrent statements or initiatives in our bibliometric analysis.

Overall, we note that the attitudes of signatories of the Joint Statement may have been affected by these other initiatives, including through peer pressure, marketing efforts or other types of exposure.

Table 2. Examples of other initiatives focusing on open sharing and collaboration during the COVID-19 pandemic

Initiative	Audience	Focus
Public Health Emergency COVID-19 Initiative (March 2020)[9]	Publishers	• Open access and open data sharing via PubMed Central
Open COVID Pledge (April 2020)[10]	Private companies	• Making intellectual property free of charge for use during the pandemic
Open letter on COVID-19 data (April 2020)[11]	Data submitters, data users, policy makers and the wider research community	• Open data sharing

WHO Solidarity Call to Action (May 2020)[12]	WHO Member States, intergovernmental organisations, nongovernmental organisations	• Sharing intellectual property through public health-driven voluntary, non-exclusive and transparent licences
UNESCO Joint Appeal for Open Science (October 2020)[13]	Member States, policymakers, civil society representatives, youth networks and the scientific community	• Access to research and its applications • Research collaboration • Quality assurance of research • Equal opportunities and scientific literacy
Carbis Bay G7 Summit Communiqué (June 2021)[14] and G7 Research Compact (July 2021)[15]	Group of Seven (Canada, France, Germany, Italy, Japan, United Kingdom, United States)	• Research collaboration • International cooperation • Open data sharing • Assessment and rewards for collaboration and knowledge sharing
Shared commitment to public involvement (March 2022)[16]	Research funders, regulators and research performing organisations involved in UK health and social care research	• Improving public involvement in research • Opening up access to, and promoting use of, current evidence

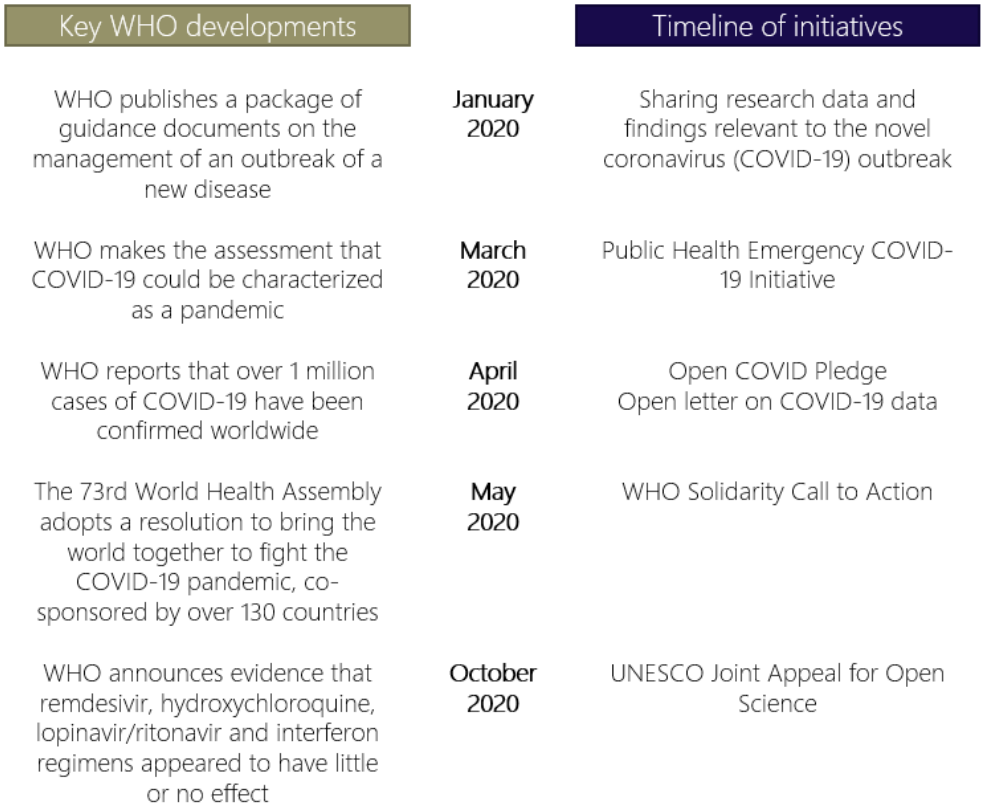


Figure 3. Timeline of key WHO developments during 2020 and examples of parallel initiatives focusing on open sharing and collaboration

3. Gathering signatures and communicating alignment

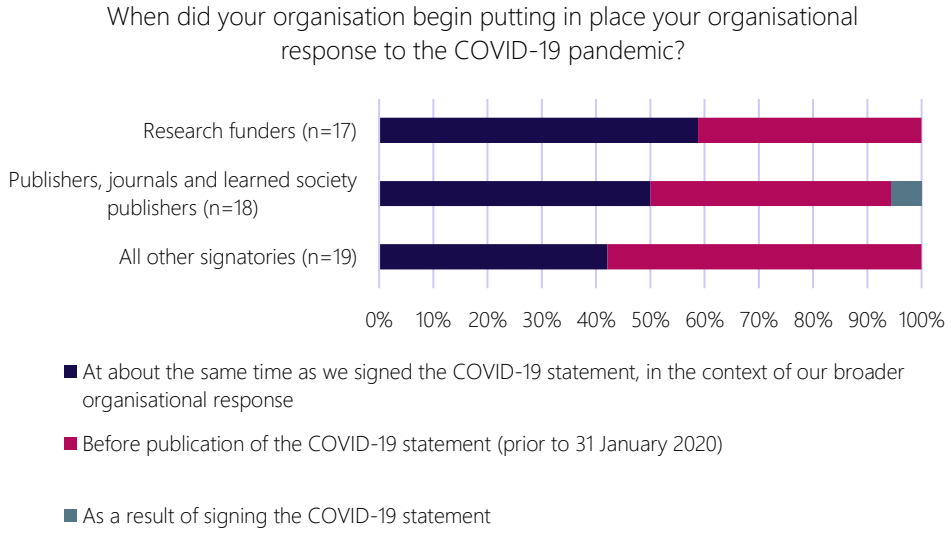
3.1 Signing the Joint Statement

The Joint Statement was signed by a large number of prominent international organisations

A total of 160 organisations signed the Joint Statement: signatories include research funders (n=29, 18%), research performing organisations (n=13, 8%), publishers and journals (n=40, 25%, including the 'big five', i.e. Elsevier, Sage, Springer Nature, Taylor & Francis and Wiley), learned societies and learned society publishers (n=14, 8%), preprint servers (n=5, 3%) and more (see Appendix B).[1] Signatories represent organisations from Europe, Africa, Asia and North and South America.

Notably, a wide range of signatories began putting in place their organisational response to the pandemic prior to signing the Joint Statement (see Figure 4). This suggests that many organisations were already considering how to make a difference, and that the Joint Statement was not the sole reason for their pandemic-related activities.

Figure 4. Timing of organisational responses to the pandemic



"We already complied with the call from Wellcome at the outbreak of the pandemic. Our policies are designed to accelerate research and make it available and useful to all: it was an opportunity for us to show how these principles are relevant and essential to address a global pandemic." – Publisher

Many organisations signed the Joint Statement as it was in line with their

According to our survey of signatories, the top three organisational drivers for signing the Joint Statement were that it was in line with existing organisational efforts around COVID-19 (87%); it helped signal alignment with open sharing practices to peer organisations

organisational strategy and ethos

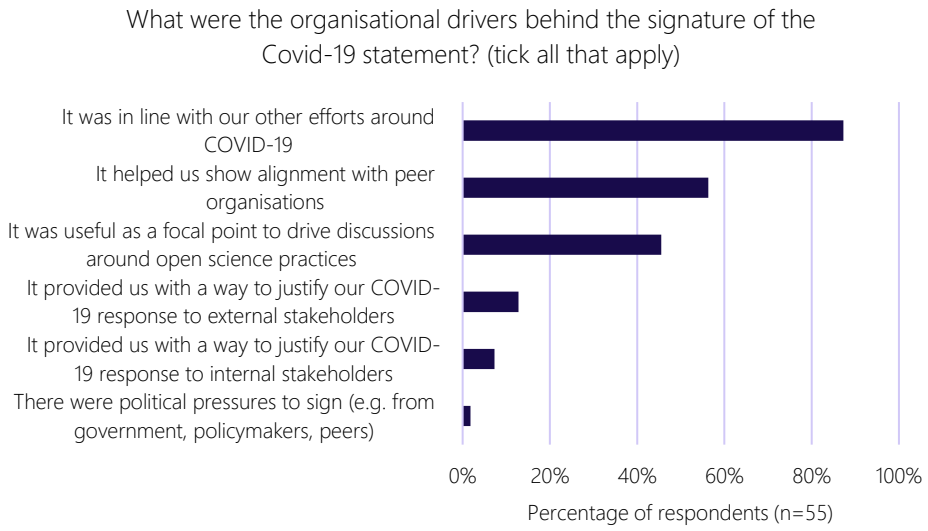
(57%); and it was useful as a focal point to drive discussions around open research practices (45%) (see Figure 5).

Qualitative responses to the survey also indicate that some organisations viewed the Joint Statement as in line with their organisational mission. A handful of survey responses noted that the commitments outlined in the Joint Statement were "the right thing to do" in order to save lives in the context of a global pandemic, and one interviewee described the Joint Statement as "preaching to the choir". This is not to say that the statement had no impact in these cases (as demonstrated by the difference-in-differences analysis in section 5.2 and Annex A), but that for some signatories there were very limited barriers to the decision to sign. In these cases, we expect that the Joint Statement helped shape the specific actions taken by signatories to respond to statement commitments, rather than affecting their overall attitudes to open sharing/open research.

Notably, although the present report describes the Joint Statement in a neutral manner to reflect its collective nature (i.e. we use the wording "the Joint Statement"), signatories attributed it to Wellcome to different extents in their public communications (e.g. "the Wellcome Statement",[17] "the Wellcome Joint Statement",[18] "the Wellcome coordinated statement").[19] Our study did not investigate whether the Wellcome brand had an impact on the uptake of the Joint Statement, but it is reasonable to hypothesise that a less recognised organisation may not have led to equally high recognition and widespread adoption of the statement.

"From its founding, [Organisation] has been a deep believer in data and tool sharing. Similarly, preprint posting has really taken off within our community, even before the pandemic. So, COVID has further reinforced trends and beliefs that were already in place in our community." – Research performing organisation

Figure 5. Organisational drivers for signing the Joint Statement



3.2 Communicating the signature of the Joint Statement

There was variation in the ways signatories communicated their signature

Following the release of the Joint Statement, Wellcome encouraged, but did not require, organisations to share their signature of the Joint Statement publicly, for example through press releases or social media communications. Many organisations did so and shared an announcement of their signature with both internal and external stakeholders (see Figure 6). Notably, 11% of respondents did not communicate their signature or commitment to the statement with stakeholders.

Figure 6. Communication of the statement to relevant stakeholders

Signatories actively communicated policy and process changes

RQ A2 | RQ A3

A number of changes to policies and processes were made during the pandemic, but, needless to say, these were not only caused by an organisation's signature of the Joint Statement (see Section 3.1). According to our survey findings, both funders and publishers actively communicated policy changes linked to open sharing. For example, we found that publishers communicated policy and process changes to existing and new authors as well as editorial staff, while peer reviewers were less commonly notified of changes. For funders, survey responses show that new applicants and new grant holders were most commonly notified of policy changes resulting from the Joint Statement before the signature of a grant agreement. A funder's signature of the Joint Statement was expected to directly affect funding calls related to COVID-19 research only. Other concurrent or future calls may be affected by wider changes in open sharing practices arising in a COVID-19 context, but we considered such cases to be beyond the scope of the present report.

Part B. Assessing the implementation and impact of the open sharing statement

Research Consulting
Science-Metrix

4. Attribution: Honouring the commitments in the Joint Statement

4.1 Operationalising the commitments in the Joint Statement

Publishers implemented the statement via new workflows to identify and share COVID-19 literature

RQ A4

The Joint Statement required signatories to make relevant publications open access or free to read and share these with the WHO. This presented a challenge as these seemingly simple actions are based on the assumption that COVID-19-related works can be easily identified.

Our survey of signatories highlighted that publishers large and small had to design and implement new workflows to make the above a reality, for example by adding specific tags or keywords in their manuscript tracking systems. In some cases, this approach also fed into dedicated COVID-19 portals hosted by publishers or journals seeking to provide a single access point to all online and relevant literature they published after peer review. Joint Statement signatories that hosted portals focusing on freely available peer-reviewed materials include: BMJ,[20] Elsevier[21] (including Cell Press[22] and The Lancet as journal signatories),[23] JAMA,[24] Journal of Hospital Infection - Healthcare Infection Society,[25] NEJM,[26] Oxford University Press,[27] PLOS,[28] SAGE,[29] Springer Nature,[30] Taylor & Francis,[31] Wiley.[32] Preprint servers such as medRxiv and bioRxiv also created a dedicated collection of COVID-19-related preprints.[33]

It is important to note that portals were, in some cases, set up before or at the same time as signing the 2020 Joint Statement (e.g. the Elsevier COVID-19 Information Center[34] and the Springer Nature COVID-19 Research Highlights were both launched in January 2020).[35] In practice, publishers and preprint servers recognised the importance of opening up information based on earlier events, external pressures or possibly in acknowledgement of earlier commitments made as part of signing the statements issued during the Zika or Ebola virus outbreaks.

We also highlight that, going beyond the Joint Statement, some publishers (e.g. Hindawi[17] and SAGE)[36] offered article processing charge waivers to authors: this is significant as the publishers enabled authors (irrespective of their geographical location) to publish via open access without paying the usual charges. In other cases, publishers only went as far as offering free access to publications on COVID-19 for the duration of the pandemic, which is a time-limited option under which access may be withdrawn in future.

Most research funders and publishers pointed to existing open sharing policies and procedures that were in place prior to the pandemic

RQ A2 | RQ A3

Our survey investigated the extent to which signatory publishers and funders made changes to their policies and procedures upon signing the Joint Statement, particularly with regard to preprint posting and open data sharing. In most cases, publishers and funders referred to existing policies (e.g. assuring authors that preprint posting is allowed), though some dedicated efforts were also reported. Our review did not lead to the identification of cases where preprint posting was *mandated*, however, which would have ideally been required to fully meet the commitments in the Joint Statement.

Notably, some funders updated their calls for proposals to *encourage* the sharing of open access or free-to-read versions as soon as possible (see, for example, the additional terms and conditions by UKRI for COVID-19 research).[37] This was one of the most significant process changes reported, but, as noted below, grantee compliance was generally difficult to measure.

Case study 1

Supporting the implementation of the Joint Statement: Health Research Board Ireland's COVID-19 Pandemic Rapid Response Funding Call

In 2020, the Health Research Board Ireland (HRB) funded a programme of research to provide evidence to inform the national and global efforts to deal with the coronavirus outbreak.[38] Among the criteria for successful applicants (see p. 14 of the guidance notes)[39] was a mandate to ensure that research data produced via this funding be managed responsibly and shared rapidly to ensure public health action: prospective grant holders were required to submit a data management plan to the HRB and strongly encouraged to publish their study protocols via the HRB Open Research publishing platform.[40]

The above policy position, including a commitment to open access publishing, enables HRB to meet some of the commitments made as a signatory of the Joint Statement and showcases their aim to recognise and value a range of research outputs, including datasets, software and materials, intellectual property, policy influences as well as publications (in the context of their signature of the San Francisco Declaration on Research Assessment).[41]

Case study 2

Sharing preprints in the biomedical sciences to accelerate research

Various players have sought to enhance the sharing and discoverability of preprints during the pandemic, leveraging existing scholarly infrastructure. For example, Europe PMC has been indexing preprints since 2018, but made additional efforts in 2020 to index "the full text of COVID-19 related preprints to make them searchable, alongside peer reviewed articles".[42] In practice, this allowed researchers to identify a broader range of relevant literature and also supports text mining approaches where permitted by licensing terms.

Another example is provided by the National Library of Medicine (NLM) in the US, which launched a preprint pilot in 2020 to make National Institutes of Health (NIH)-funded preprints accessible through major services such as PubMed for preprint citations and PubMed Central for full preprints.[43] The pilot was inspired a) by a vision to foster open science practices, and followed a 2017 policy change through which NIH incentivised the use of preprints and other interim research products to disseminate research findings more rapidly;[44] b) by a desire to foster the growing interest in preprints by the research community and provide education to distinguish preprints from journal-organised peer-reviewed research articles; and c) to focus on preprints resulting from NIH-funded COVID-19 research, and thereby accelerate sharing of research findings during the pandemic. The pilot has run for over a year (since June 2020) and resulted in the addition of more than 3,200 NIH-funded preprint records to PubMed Central.[45]

Efforts to monitor compliance with the Joint Statement were limited, but more common among publishers

RQ A1

Our survey and interviews investigated the extent to which signatories monitored compliance with the commitments in the Joint Statement (Figure 7). In principle, one may expect monitoring by publishers to be more quantitative, e.g. in the form of shares of articles published via open access or shares of articles including a data availability statement, as publishers may hold this information in their databases. However, this is not always possible: for example, the information available to publishers may be too partial to provide comprehensive assessments on the topic of data availability statements.

The situation is even more complex for funders, as they are one step removed from information on publishing venues and research data sharing behaviours: they are more likely to gather insights in these areas via grantee reports or as part of survey exercises (the latter being the case for UKRI during the COVID-19 pandemic). In addition, funders are often affected by resourcing limitations, as has been found to be the case with regard to other policy requirements: for example, funders are often unable to closely monitor the extent to which their grantees share data according to their policies, as this tends to require disproportionate efforts and resources.[46]

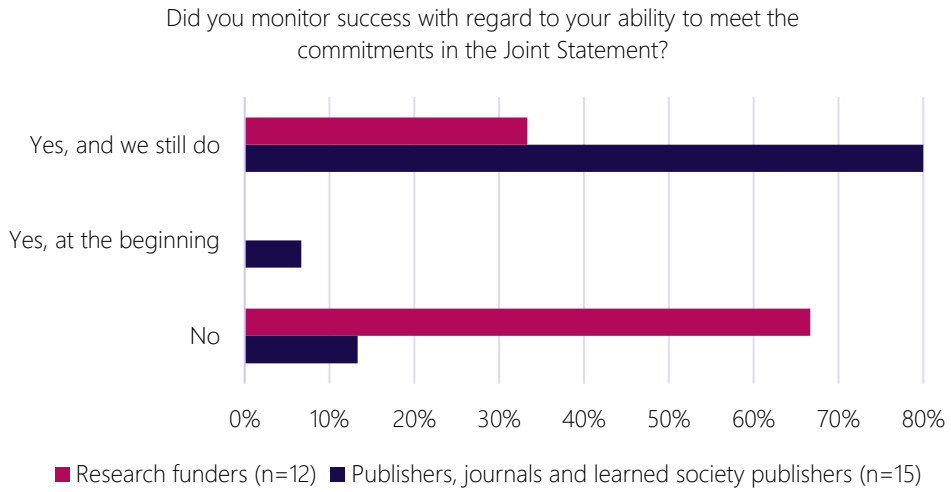
In our survey, we found that publishers did and still do monitor compliance in this respect, for instance by enhancing their submission systems to track COVID-19-related literature or by monitoring the share of published articles with an associated preprint and/or a linked dataset in a repository. Systematically monitoring author compliance was highlighted as a more challenging task, particularly for publishers with large portfolios. Most funders reported that they did not monitor success with regard to their ability to meet the commitments in the Joint Statement, partly in acknowledgement of the fact that, if publisher signatories are compliant, then funded authors would also be compliant in turn.

We also found that neither publisher nor funder signatories put in place measures to address non-compliance with Joint Statement commitments, likely in recognition that the initiative sought to foster and enable open behaviours rather than set up a formal monitoring framework (for instance, the Joint Statement says: "...we commit to work together to help ensure...").

Wellcome's efforts to understand the impact of the Joint Statement on Open Practices has taken the form of the present report, which was commissioned in partnership with UKRI and the Bill & Melinda Gates Foundation. However, we note that this study is not meant to formally assess the performance or compliance of individual signatories.

> *"We did not monitor the success [of our efforts to meet the Joint Statement's commitments] as we do not have the instruments to do this. Monitoring of data sharing is even harder. We did however evaluate the openness of preprints and data of the projects that we funded in our [COVID-19 call for proposals]".* – Research funder

Figure 7. Internal monitoring efforts undertaken by research funders and publishers/journals



Research performing organisations supported the global pandemic response

We note that only a minority of research performing organisations signed the Joint Statement, and this may have arisen from the fact that the document did not address them specifically and only referred to "researchers, journals and funders".

Although our research didn't investigate research performing organisations in detail, we have captured a range of relevant contributions in our survey. Examples of efforts to support implementation of open sharing by research performing organisations included changing a press release policy to allow the discussion of preprints (as opposed to journal content only); requiring all new COVID-related projects to share data and results as early as possible during ongoing projects; providing support via data stewards and developers; and setting up an institutional feed to highlight links to relevant preprints and journal content. These examples show that research performing organisations could, indeed, support implementation of some elements proposed by the Joint Statement.

Future statements may consider opportunities to either define the types of organisations that are eligible to sign or seek to formulate broader or additional requirements by stakeholder groups to enhance clarity and streamline the operationalisation of the commitments.

4.2 Comparing signatories and non-signatories with regard to open access, data sharing and policy citations

The global research community rose to tackle the pandemic, rapidly affecting funding, publishing and collaboration

Before considering the impacts of the Joint Statement, it is essential to acknowledge that open research practices have been increasing in prevalence during the pandemic, driven by a desire to work together against a shared challenge and save lives.[47] The changes in open sharing behaviours during COVID-19 have been described as "a completely new culture of doing research" by some, whereas others consider prior commitments to open

sharing and investment in data sharing infrastructures as having played a role in shaping the behaviours seen during the pandemic.[48]

An analysis by Nature reports that, during 2020, around 4% of the world's research output was devoted to COVID-19.[49] The above is not surprising: not only did the global research community wish to collaborate to tackle a shared challenge, but also funding opportunities rapidly emerged to cover the costs of COVID-19 research, including via reshaping and redirection of available funds. For example, the COVID-19 research project tracker, a live database of funded research projects across the world related to the COVID-19 pandemic, which is run by the UK Collaborative on Development Research (UKCDR) and GloPID-R, tracks upwards of 15,600 research projects, covering all 9 of the WHO priority research areas and with a total estimated funding of $7,364.6 million.[50] Our bibliometric analysis has identified close to 200,000 research papers on the topic of COVID-19 in 2020 and 2021, including about 160,000 journal publications and close to 40,000 preprints (covering medRxiv, bioRxiv, arXiv and SSRN).[ii] Of the journal publications, 46% are open access and a further 40% are free to read. Preprints, on the other hand, are all freely available online, leading to an unprecedented amount of literature being available to any interested stakeholder.

"I think that COVID-19 really influenced the importance of rapidly disseminating information because it was a rapidly evolving pandemic… that was something that was very unique that I think was almost a silver lining to help modernise and enhance the way we do things." – Public health organisation

Preprinting and rapid peer review increased the pace of research during the pandemic

The large number of preprints mentioned above is significant: during the pandemic, preprints received an unusual amount of attention from scientists, news organisations, the general public, and policymakers, representing a noticeable departure from previous behaviours. Recent research noted that preprint servers "have been legitimised for their speed and agility as a result of the coronavirus", and our interviews provided additional evidence of this.[51]

We also note that the Rapid Review Collaboration Initiative, which was started in April 2020 by a group of publishers, contributed to enhancing the speed of scholarly communication during the pandemic.[52] The overall faster speed of peer review of COVID-19-related outputs[53] has been described as a shift "from prioritising novelty towards a focus on clinical or societal relevance" and has benefits beyond the faster speed of publication e.g. the ability to reference published works in fast-moving policies;[54] but also drawbacks (e.g. retractions arising from the publication of incorrect findings in prominent journals).[55]

[ii] The total of almost 200,000 research articles can be split out as follows: 153,421 journal publications (with a DOI), 6,703 journal publications without a DOI (excluded in the analysis that follows) and 38,122 preprints. This total includes preprints that have been subsequently shared as publications. The deduplicated total, only including items once, is 186,450 research articles, including journal publications and preprints.

"The next time around... more peer reviewed papers are not the answer. We need to have preprint servers. We need to have high velocity means of sharing information around the world. I would say if there's any medical publisher right now who is on the fence about preprint servers, Coronavirus finished that off." – Publisher

The global research community joined forces to tackle a shared challenge

Furthermore, as the COVID-19 pandemic was perceived as a shared challenge by the global research community, it is not surprising that about 34% of scientific articles on COVID-19 in early 2020 involved international collaborations, compared with 28% of non-COVID-19 scientific articles in the same period, and 23% of all articles published between 2015 and 2019.[56] It is noteworthy that a wide range of collaborations were established between different stakeholder groups, including universities, research institutes, pharmaceutical players, policymakers, public health bodies and more. These took various forms, including, but not limited to new initiatives, networks and working groups, such as COVID CIRCLE,[57] the COVID-19 Clinical Research Coalition,[58] the COVID R&D Alliance,[59] the COVID-19 Taskforce,[60] GloPID-R[61] and VODAN Africa and Asia.[62]

Importantly, organisations involved in the pandemic response appreciated the importance of transparency and openness: it was understood that collaboration would speed up COVID-19 research, for example in terms of more rapidly developing promising treatments, jump-starting and optimising studies or identifying novel uses for repurposed drugs, even though some limitations did apply in terms of intellectual property (particularly in the case of pharmaceutical companies).[63]

"Everything that we did, we did it with the ethos that this is open innovation. We wouldn't have done anything else because the whole effort right from the start of getting the government funding was purely to provide societal and scientific benefits." - Pharmaceutical company

The evolving research landscape must be considered when assessing the impact of the Joint Statement

The evolving research landscape should be kept in mind when viewing the bibliometrics boxes that follow (see light brown shading): shifting research cultures mean that it is more difficult to assess whether signatories have made changes in their processes, procedures or behaviours due to the Joint Statement itself or because of other phenomena in the research landscape. The difference-in-differences approach used by Science-Metrix is key to complement our qualitative findings and address the extent to which any changes seen in signatory organisations may be tied to the Joint Statement.

A key feature of the bibliometrics analysis presented here has been the differentiation of journal publications and preprints by signatory status (whether they were published in a signatory journal, financially supported by a signatory funder, or co-authored by a researcher affiliated with a signatory research performing organisation). The efforts to differentiate between signatory journal publications and preprints and non-signatory ones were successful for 37% of COVID-19 journal publications and 13% of preprints (with the main limitation to this operation being low availability of funding data, see Annex A). Within the roughly 60,000 COVID-19 journal publications and preprints where it is possible to determine signatory status, just under 85% were found to have signatory status and about 15% to hold non-signatory status. Among close to 4,400 arXiv, bioRxiv, medRxiv

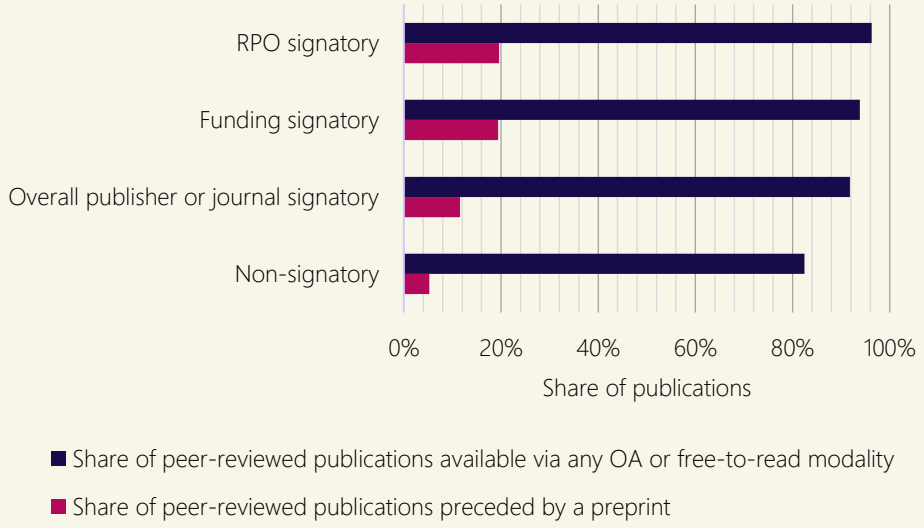
and SSRN COVID-19 preprints where signatory status could be determined, 13% were attributed non-signatory status, and the remaining proportion had signatory status.

Despite the low proportions of journal publications and preprints found not to have benefitted from signatory organizations' support in any way, identification of signatory status has proven interesting for a second reason. As will be seen below, it was found that preprinting, data sharing and policy-related uptake measurements were arguably as distinctive within the three different signatory layers as between signatory and non-signatory groups.

Signatories made journal publications available to read freely more than non-signatories

Descriptive findings show that a greater number of COVID-19 journal publications have been made available to read freely (gold, hybrid or green open access or free to read) by signatories (92%) than by non-signatories (82%; see Figure 8).

Figure 8. Access to journal publications[iii] and usage of preprinting[iv]



Publisher-based free-to-read and hybrid open access are more common among signatories than non-signatories

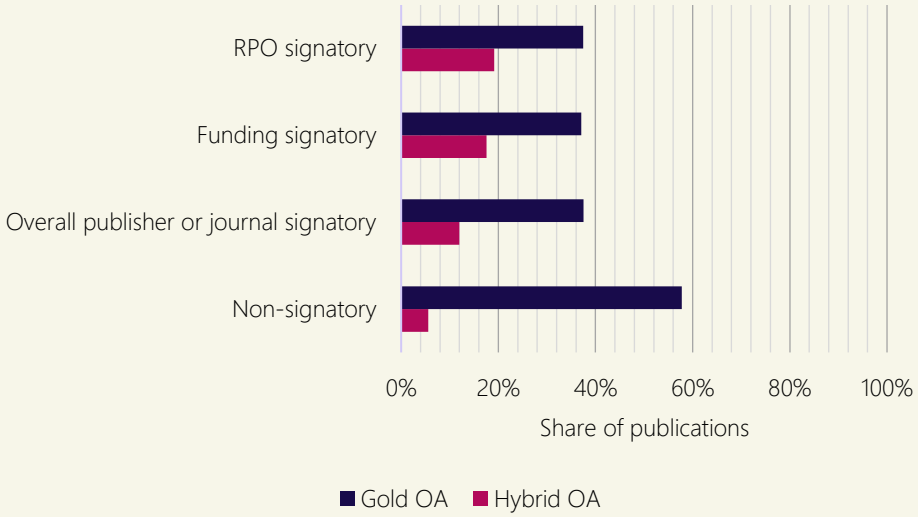
35% of journal publications in signatory journals were publisher-based free-to-read (but not open access, see Glossary), compared to 14% for non-signatory articles. If the five major publishers are excluded from calculations, however, only 7% of signatory articles are publisher-based free-to-read.

In addition, 12% of articles were published via Hybrid OA in signatory journals, while the figure is only about 6% for non-signatories. Non-signatory articles had much higher shares of gold OA publications (58%) compared to Joint Statement signatories (38%, see Figure 9).

[iii] RPO = Research performing organisation; Note: Publications for which Unpaywall records were incomplete or missing were considered null and removed from the computation of OA/free-to-read figures. Source: Scopus, Unpaywall; processed by Science-Metrix

[iv] Source: arXiv, bioRxiv, medRxiv, SSRN, Scopus; processed by Science-Metrix

Figure 9. Shares of Covid-19 journal publications available under Gold or Hybrid open access[v]



Preprint posting is more common among signatories than non-signatories

Non-signatory COVID-19 journal publications were preceded by preprints in only 5% of cases. For publications with signatory status through funding or institutional affiliation, this share jumps to almost 20%. Publications published with signatory journals were preceded by preprints in 12% of cases. Flipping the perspective and starting from preprints themselves, we found that 47% of COVID-19 bioRxiv preprints, 39% of medRxiv preprints, 32% of arXiv and 18% of SSRN preprints were subsequently published in journals (these figures have been computed independently of signatory status).

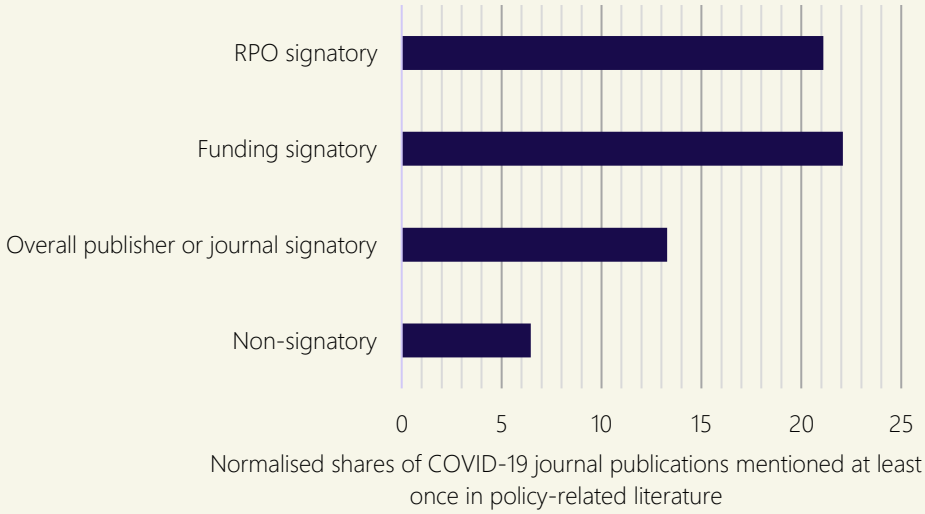
Articles and preprints associated with signatories were more commonly cited in policy documents

Descriptive findings for journal publications show stronger performances on policy-related uptake for affiliation- and funding-based signatory papers, followed by journal signatories (see Figure 10). Non-signatory publications fared below the signatory groups on policy-related uptake in the descriptive results.

Signatory COVID-19 preprints saw much higher policy-related uptake (22% with at least one policy-related citation) than non-signatory preprints (6%). This lead for signatory preprints was reduced when not controlled for disciplinary differences and other author-related factors (i.e. when the set of authors contributing to each group is not kept constant; 12% to 5%).

[v] Note: Publications for which Unpaywall records were incomplete or missing were considered null and removed from the computation of OA/free-to-read figures. Source: Scopus, Unpaywall; processed by Science-Metrix

Figure 10. Policy citation of journal publications[vi]



Case study 3

The explosion of preprint posting: sharing COVID-19 research via MedrXiv and biorXiv

Racing to get pandemic-related research results out in the public domain has led to a surge since 2020 in submissions to preprint servers. These are online platforms where researchers can post their results before journal-organised peer review, which is typically time-consuming, delaying the appearance of scientific results. The CSHL-hosted preprint sites bioRxiv and medRxiv are amongst the top six servers for hosting coronavirus research[42] - to date these servers host 22,600 COVID-19 SARS-CoV-2 preprints (17,260 medRxiv, 5,340 bioRxiv).[33] For researchers who preprint, the advantages are that results are released as soon as they are written up, meaning they are available within days of this, for evaluation by experts and thereafter used to readily guide health-related behaviour. The downside is that preprints are not peer-reviewed before posting and cannot be considered conclusive on appearance in the public domain. (Although reputable journals conduct peer review before publishing scientific work, this time-consuming step is of course not necessarily a badge of perfection.) To combat this, enhanced screening of coronavirus research in preprint submissions has been implemented.[64] Baseline screening of preprints on these servers has always been in place prior to posting: initially to ensure completeness of the manuscripts, avoid plagiarism, and by expert volunteers, to detect non-scientific assertions or biosecurity threats. Enhanced screening at bioRxiv and medRxiv serves to prevent posting results that can potentially cause harm; and notably studies that involve making predictions about treatments for COVID-19 solely based on computational results.[64]

Although open sharing behaviours increased

Our interviewees agreed that data sharing at the beginning of the pandemic has enabled the prompt development of vaccines and therapeutics. In particular, after the COVID-19 virus was identified in China in January 2020, the scientific community moved swiftly: "the virus's entire genetic makeup was published online within days"[47] via The Lancet and

[vi] Note: Publications with no metadata in PlumX were considered null and removed from the computation of this indicator. Source: Scopus, Overton and PlumX; processed by Science-Metrix

during the pandemic, data sharing remains immature

across two repositories (China National Microbiological Data Center and China National GeneBank).[65] This is very fast - notwithstanding the advance in identification and sequencing technology since the SARS coronavirus outbreak in 2003, where the process took almost three months.

Our interviewees pointed out that, to tackle an evolving pandemic, several types of data are required (see Case study 4). For example, genomic viral data and clinical data at the individual (anonymised) and aggregated level are key to plan and execute an effective public health response: clinical data may be used in combination with genomic viral data to predict the severity of COVID-19 infections[66] or to assess the impact of identified mutations on vaccine effectiveness.[67] Data from public services (sometimes referred to as open government data)[68] may also be of help, for example wastewater was analysed in both the UK[69] and the United States (via the National Wastewater Surveillance System) to monitor infection levels.[70] Finally, economic data as well as information on ethnic groups can help protect the least advantaged portions of a country's population.[71]

Nevertheless, barriers to data sharing remain, as noted in The State of Open Data 2021 report,[72] in recent research on data-sharing in preprints, and as confirmed by the limited levels of data sharing identified by our bibliometric analysis (see below).[73]

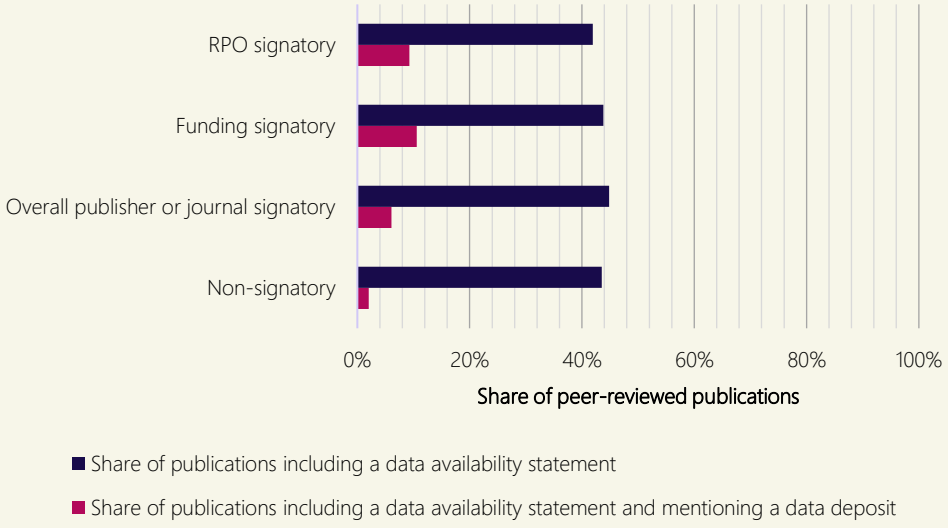
Less than half of COVID-19 publications have data availability statements

Descriptive statistics show that the shares of COVID-19 publications where authors have written a data availability statement ranged between 42% and 45%, both for signatory and non-signatory groups of journal publications.

Non-signatory publications were less likely to mention repositories as part of their data availability statements

Our findings show greatly varying degrees of data sharing by signatory status. Within the subset of publications with a data availability statement, non-signatory publications mentioned selected repositories (see Annex A) in slightly less than 5% of cases. On the other hand, in the case of publications associated with Joint Statement signatories, this figure was as high as 22%-24%.

Figure 11. Data sharing behaviours in COVID-19 journal publications



Case study 4

The International COVID-19 Data Alliance (ICODA) has emerged since the pandemic started. It was convened by Health Data Research UK in 2020, who identified close to 100

Elevating data analysis: ICODA's efforts to combine data from different sources

high quality yet isolated data repositories housing COVID-19 data.[74] ICODA aims to increase the co-ordination across these so that relevant health information from different sources across institutional and geographic boundaries can be easily found and aggregated for analysis. Funded by the Bill and Melinda Gates Foundation, Wellcome and others, this initiative[75] incorporates world-renowned repositories such as the COVID-19 data portal[76] from EMBL-EBI and Vivli's COVID-19 portal for clinical research.[77] It is modelled on other global initiatives such as the Global Alliance for Genomics and Health and has gained traction across countries who recognize that an effective pandemic response requires a combination of diverse data types.[78]

ICODA aims to build a global alliance of data partners and to build trustworthiness and transparency in their policies, data governance, privacy and usage.[79] A secure trusted research environment is provided through the ICODA workbench (note that researchers need to be accredited to participate).[80] A wide range of research outputs supported by ICODA have emerged since 2020.[81]

5. Impact: Assessing the effects of the Joint Statement

5.1 The benefits of openness during a pandemic

The sharing of the viral genome at the beginning of the pandemic was key to promptly developing vaccines and therapeutics

RQ I2

Our stakeholder interviewees clearly described the free and open sharing of viral genomes in early 2020 as the core enabler of the "rapid development of diagnostic kits and vaccines".[82] However, due to the emergence of new strains, this was only a starting point. The subsequent submission of genome sequences to online databases (see Case study 5) has been instrumental in developing and expanding pathogen genomic surveillance initiatives: genomic sequencing, particularly when results are shared in the public domain, can inform our understanding of how COVID-19 is evolving, so that the global health response can be tailored accordingly.[83]

"Obviously, the really important thing was the sharing of the original SARS-Cov-2 sequence from China. And from a vaccine development point of view, we didn't need the virus, we just needed the sequence." – Academic

Case study 5

Sharing and analysing viral genomic data: the role of GISAID during the COVID-19 pandemic

Since January 2020, GISAID's data sharing platform has been the most popular primary source of genomic and associated data from SARS-CoV-2 cases. GISAID (the Global Initiative on Sharing Avian Influenza Data) is so called because it originally launched in 2008 to promote data sharing during the avian influenza epidemic.

Usage of the viral genomic data is granted free provided that users sign a database access agreement that confirms their identity and prevents republishing the site's data unless permission is granted from the data provider.[84] This ensures that researchers depositing sequences retain their rights to viral data and are acknowledged in publications,[85] a feature welcomed by many as promoting equity and sovereignty.[86] The GISAID platform and database, officially hosted by Germany, has enabled web efforts that aid, for example, analyses of genomic characteristics and virus evolution during the COVID-19 outbreak.

GISAID curators ensure improvement of deposits or withdrawal if duplicates sequences are identified; sequences are made available to registered users after curation with immediate access.[87] Despite the equity and openness promoted by GISAID, concerns have been raised about this restricted data sharing,[86] which isn't applied to other databases (e.g. those of the International Nucleotide Sequence database Collaboration (INSDC)) and does hinder some studies (e.g. it may prevent authors from re-sharing data arising from the mixing or combination of GISAID data with data from other sources).[88]

The number of articles, datasets and resources on COVID-19 significantly increased on a daily basis

The core objective of the Joint Statement was to enhance the open sharing of a wide range of research outputs, including articles, preprints, data, protocols and more. Due to the increased sharing of knowledge observed during the pandemic, some raised concerns around the possible issue of information overload, i.e. a "situation that arises when there is so much relevant and potentially useful information available that it becomes a hindrance rather than a help".[89] In practice, however, this wasn't seen as a major drawback by our interviewees: organisations relied on internal and external networks (including Twitter and newsletters) to sift through the vast number of materials shared daily. Some agencies such as the US CDC devoted internal staff time to reviewing, summarising, and publicly sharing key materials in the form of "Science updates" (see case study 6), which made it easier for others to access the latest credible science on COVID-19.

Case study 6

Keeping on top of COVID-19 research: the CDC COVID-19 Science updates

To help health researchers, policymakers and practitioners keep on top of the emerging scientific information relating to COVID-19, staff from the Centers for Disease control and prevention (CDC) in the United States have produced a weekly digest of the previous week's scientific literature known as the CDC COVID-19 Science updates.[90] A typical weekly digest contains summaries of between 11 and 19 articles taken from preprint servers or scientific journals that are featured in the WHO COVID-19 database.[91] As part of an effort to produce an evidence base for public health action to combat SARS-CoV-2, the articles are selected by relevance to public health priority topics in the CDC Science Agenda for COVID-19. Each article summary includes a link to the article and brief information relating to 4 headings: Key findings, Methods, Implications and Figures.

These digests were produced and made available weekly from April 2020[92] through 17 December 2021 with access to an overall listing of featured articles.[90]

Timely access to research findings and public health data was invaluable to track, monitor and respond to the virus

RQ I1

The prompt sharing of viral genomes over the course of 2020 and 2021 has helped policymakers, public health agencies and scientists understand how the Delta and Omicron variants replaced Alpha in turn. The sharing of public health data has also contributed to the global response to the pandemic. This has been pursued by individual countries via portals of varying sophistication but also through central visualisations such as the prominent Johns Hopkins COVID-19 Dashboard (see case study 7). Only by analysing this wealth of datasets and combining this with potentially sensitive patient data can a country form a clear picture of COVID-19 infections, which highlights the pressures on a country's technological capabilities and digital infrastructure. This also has implications for LMICs, as technological infrastructures and governmental funding programmes may be less advanced and therefore lead to sub-optimal knowledge of the infection at the national and local level.

Research data and public health data were also interpreted by expert government advisors to generate insights and recommendations for policymakers, including around mask and vaccine mandates and national or local lockdowns. Examples of such advisors include the Scientific Advisory Group for Emergencies (SAGE)[93] in the UK and the Chief Medical Advisor to the President in the United States.[94]

With Omicron, the WHO had a very detailed conversation with South Africa, in which they shared all of the findings that they had well before any publication. – Public health agency

Open sharing enabled timely information transfer to inform public health responses

RQ I1

Public health agencies have been able to benefit from the increased availability of information and data, including via trusted shared infrastructure and the growing availability of preprints. Policymakers across the world started benefiting from the openness and speed of communication seen during the pandemic. For example, an interviewee noted that preprints were regularly reviewed by government advisors and evidence considered, where appropriate, when providing insights to the government. Recent research confirms the use of preprints in policy documents by the European Centre for Disease Prevention and Control, United Kingdom Parliamentary Office of Science and Technology, and World Health Organization Scientific Briefs,[95] and the bibliometric analysis by Science-Metrix provides further detail on the topic. We note that the benefits of open sharing were reaped asymmetrically across the globe, with high income countries being better placed to leverage the body of information and guidance emerging throughout the pandemic thanks to more significant resources, equipment and funding.[96] Initiatives such as The Global Health Network's COVID Hub sought to support countries that are less well placed to achieve this, seeking to work collaboratively and share information resources, data and know-how between countries in LMICs.[97]

Through scientific advisory groups, you take your synthesis to the group and people kick that around and have their own inputs, too. And at the end of the day, you end up quite quickly with a consensus view which is pretty powerful, and has actually integrated a huge amount of data, thought and opinion, and you produce a synthesis that provides practical answers for governments who are asking questions of their scientists. – Academic

Case study 7

The COVID-19 Dashboard by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University (JHU)

A health surveillance dashboard was launched at the outset of the pandemic by researchers at the John Hopkins University (JHU), who later comprised the John Hopkins Coronavirus Resource Center (CRC).[98] The dashboard allows users to visualise and track confirmed cases of COVID-19 in real time for all affected countries across the world. Complex data (location and number of COVID-19 cases) is aggregated from thousands of government and (social) media resources around the world via a publicly accessible GitHub server scheduled to pull and evaluate data at certain times of the day.[99] Help with aggregating the huge dataset and massive demand for visualisations of this data has been provided by the Applied Physics Laboratory at the JHU's Center for Systems Science and Engineering. The CRC has proved enormously valuable: the dashboard, which began receiving over 1 billion hits per day within weeks of its creation, now provides a range of visualisations, from global variant surveillance to worldwide vaccination status to US state data and US hospital capacity.[100] Data supporting the visualisations have been used by policy makers and the research community for purposes of modelling and planning – leading to 1,200 citations within the first 4 months of the data's publication. The disjointed public health response to the pandemic, and the gargantuan effort required to assemble accurate global data for this dashboard, have reinforced the need for open public data standards and data sharing.[100]

5.2 The impact of the Joint Statement on open sharing behaviours

Controlling for behaviours prior to the signature of the Joint Statement

In cases where results point to a positive impact of the Joint Statement, we could not fully control for signatories that have implemented changes in response to the pandemic prior to signature. We also note that the previous statements on Zika and Ebola included a call for open sharing in future public health emergencies, which may have affected responses prior to the release of the COVID-19 Joint Statement. Therefore, our assessment of impact might be more limited than results suggest. For this reason, we refer to a "likely" impact of the statement.

The Joint Statement likely led to an increase in publisher-based free-to-read journal articles, but a net differential loss in Gold OA

RQ A5

The impact of the Joint Statement on the overall level of open availability of journal articles is uncertain. The reason for this appears to be that prior signatory research recorded high shares of OA and free-to-read publications even before the pandemic, with negative or inconclusive difference-in-differences simply capturing the fact that non-signatory research had more space available to improve in this dimension. The share of publications accessible via Gold OA, by contrast, has remained stable for signatory publications. Given that non-signatory publications have seen an increase in Gold OA over the same period, the net result is a differential loss in Gold OA in signatory publications. That said, a consistent differential gain was observed for signatory publications in terms of the share of publications accessible via publisher-based free-to-read arrangements. This is explained by the fact that some publishers have made COVID-19-related contents in hybrid or subscription-only journals free-to-read as part of their Joint Statement commitments, waiving article processing charges (where applicable) for submitting authors.

The Joint Statement likely had a positive impact on preprint posting

RQ A8

Controlled differential findings all converge on preprint posting and show a clear increase in this practice for signatory publications between 2018-19 and 2020-21. This increase is above an increase also recorded for non-signatory publications. The differential gain ranges between +5 percentage points and + 29 percentage points, depending on the difference-in-differences model considered. On this basis, it is possible to conclude a likely contribution of the Joint Statement to increasing the share of journal publications that are preceded by a preprint. However, the majority of COVID-19 publications were not shared as preprints prior to publication, indicating that the practice is still not widespread.

The Joint Statement likely had a positive impact on policy citations

Controlled, differential findings showed that signatory COVID-19 journal publications benefitted from a great increase in policy-related uptake compared with signatory publications from 2018-19. Non-signatory publications also tended to record increases, but these were smaller than for signatory publications. Note that not all difference-in-differences models have reached definitive statistical certainty on policy citations, although even for those that don't, results are still positive and point towards a much higher likelihood of differential gain rather than differential decrease. Therefore, it is possible to conclude a likely positive contribution of the Joint Statement towards policy-related uptake. Again, it should be noted that the analysis leading to this finding was able to control for factors such as the potentially higher authority of signatory sources from

the perspective of policymakers, government scientists or public health scientists, insofar as these haven't greatly changed between 2018-19 and 2020-21.

The specific impact of the Joint Statement on data sharing behaviours is difficult to assess

RQ A6 | RQ A7

The controlled experiments did not find a differential gain in data availability statement writing to be traced back to the Joint Statement, and also provide inconclusive results as to whether data availability statement writing has increased with the pandemic (regardless of signatory status). In practice, we found that the general shift in data-availability-statement-related practices appears stronger than the Joint Statement effect.

5.3 Benchmarking COVID-19 research against comparable prior research

Identifying "comparable prior research"

As part of our investigation of open research trends, we considered descriptive statistics that contrast sharing of research on COVID-19 with sharing of research on during prior outbreaks of comparable diseases. To achieve this, we designed a thematic data set of publications on the topic of "human viral respiratory diseases" (HVRD) that include research on prior coronaviruses, SARS and MERS syndromes, among others. The key limitation to this approach is that HVRD research is expected to be made up mostly of biomedical research, whereas COVID-19 research was characterised by unprecedented diversity in topics.

COVID-19 research is characterised by a higher level of open sharing compared to prior HVRD research

In one difference-in-differences model (out of 4 main models), the overall OA and free-to-read score goes from 94% (HVRD) to 99% (Covid-19) for signatories, and from 84% (HVRD) to 100% (Covid-19) for non-signatories. While the differential gain for signatory COVID-19 research is of "only" 5 percentage points compared to 16 for non-signatory COVID-19 research, this result should be assessed positively, indicating that COVID-19 research reached fully OA or free-to-read status for a selection of publications from core researchers in the field.

Data availability statement writing and preprinting are more common in COVID-19 research than prior HVRD research

Multiple findings show increases in preprinting for COVID-19 against comparable HVRD research from the same period. In a conservative model, the practice of preprinting before publishing in a journal increased from 2% to 11% for signatory journal publications, and from 1% to 5% for non-signatory journal publications. Signatory preprints were seeing data availability statement levels at 59% against 38% for signatory HVRD preprints from 2020-2021.

Policy uptake is more common in COVID-19 research than prior HVRD research

In the most conservative model, policy-related uptake (shares of articles: expressed as subfield- and year-normalised shares of articles cited at least once in the policy-related literature) increased from 3.6 (HVRD) to 13.0 (Covid-19) for signatory journal publications, compared to 3.5 (HVRD) to 6.6 (Covid-19) for non-signatory publications.

For preprints, the share cited by one or more policy-related documents went up from 3% (HVRD) to 12% (Covid-19) for signatory preprints, compared to going from 0 (HVRD) to 5% (Covid-19) for non-signatory preprints.

6. Longevity: Assessing long-term effects

6.1 Reflecting on the longevity of changes introduced during the pandemic

It is expected that some exceptional behaviours seen during the pandemic will return to normal...

RQ L1

It is clearly too soon to assess whether the open sharing landscape has been radically changed by the past two years. In the fast-evolving context of the pandemic, exceptional behaviours emerged, and the global research community refocused its efforts (see section 2.2). Some of these behaviours are likely to "return to normal" after the pandemic, but practices such as preprint posting may well be here to stay given the extent of growth seen by this pathway to information sharing.[101]

For example, rapid peer review is very likely to return to its pre-pandemic form: journals have chosen various pathways to increase the pace of peer review, and the research community (peer reviewers, in particular) has shown their support in the interest of tackling COVID-19. However, this level of pressure is unlikely to be sustainable. Initiatives such as Review Commons[102] or peer review "outside journal-organised mechanisms" may continue to develop, however, in is the spirit of continued experimentation with new approaches to academic publishing.[101]

Efforts such as the CDC COVID-19 Science updates are likely to stop, too, as their website currently notes that "as of December 18, 2021, CDC has paused production of the weekly COVID-19 Science Update" (case study 6). Finally, the Joint Statement asked publishers to make relevant publications open access or free to read at least for the duration of the pandemic. It is expected that most free-to-read materials for which publishers retain copyright, will go back to their standard subscription-only status at some point in the future (and some already have), unless they have been published via gold open access.[103]

Overall, given research actors around the world joined forces to resolve an unusually complex threat of the pandemic, at least a partial return to pre-pandemic business models and pressures is expected, once the immediate threat has passed.

> *"We opened up our Coronavirus papers during the pandemic. Does that mean that we're now going to open up everything else? It doesn't, it's not that simple. We will continue our transition towards open publishing as all publishers are... in a way that we feel we can navigate responsibly."*
> – Publisher

... but the end of the pandemic will not mark an end to commitments supporting open research

Signatory contributors acknowledged that, although many pandemic behaviours were exceptional, their commitment to open research and open sharing is likely to continue in the future. It was recognised that any efforts to pursue open sharing will only work if the broader research landscape is aligned, for example incentives for researchers are in place and the funding is available. These considerations will also affect the extent to which the world is prepared for a future pandemic (see case study 8).

Discussion in the literature supports this. The pandemic did, indeed, nudge the sector further along the path to recognising open sharing behaviours, but practices such as the widespread use of preprint servers or comprehensive data sharing are likely to flourish only if these are rewarded or mandated by research performing organisations and research funders.[101]

In addition, a range of research collaborations and new or existing research infrastructures (see Section 4.2) have proved their worth to the global research community, and there is a desire for these to continue to operate in the future. Again, funding and incentives are key to their survival post-pandemic yet these may take a different shape over the next few years.

"Especially working on infectious diseases, there will be much more openness. I don't think it will be transformational, but I think it can help to catalyse some changes. We have seen some open drug discovery projects and consortia that have grown during the pandemic. And I hope that some of those will continue and develop post-pandemic doing other things. But again, that will depend on their ability to pay for what they're doing and attract the support for what they do." – Academic

The pandemic has given rise to new organisational workflows that can be reused in the future

Joint Statement signatories have devised a wide range of strategic and operational changes to implement their responses to COVID-19. In many cases, these may have crystallised in the form of new procedures that organisations may more easily deploy in the future should similar emergencies arise. Although we are unable to provide evidence of future behaviours, we argue that the Joint Statement may positively affect a range of scenarios – not necessarily limited to the biomedical field – where signatories may re-deploy such procedures to ramp up open sharing efforts and tackle a shared challenge (e.g. an environmental crisis or a natural disaster of unusually large proportions).

Case study 8

Preparing for the future: The GloPID-R Principles of Data Sharing in Public Health Emergencies

The GloPID-R Principles of Data Sharing in Public Health Emergencies (2018) provide a framework for timely data sharing during an outbreak.[8] They can be used to support data sharing during such emergencies as part of the research response, from preparedness and public health response to the development of vaccines, diagnostics and therapeutics. GloPID-R is a network of research funding organisations investing in infectious disease preparedness research.[104] To accompany this, data sharing practices were assessed during past outbreaks (Ebola outbreak in West Africa; Middle East Respiratory Syndrome (MERS), Cholera and Yellow Fever outbreaks in Africa and China, and Zika in Latin America).[105] Barriers were identified and recommendations to overcome these were formulated in a report commissioned by Wellcome in December 2018. A roadmap of five recommendations and priorities for GloPID-R funders was published for GloPID-R in 2019 including: 1. To improve funder policies, 2. align tools and strengthen capacity, 3. to build trust and advance equity in data sharing not just utility of data, 4. to influence stakeholders beyond funders, i.e. ministries of health, commercial publishers and research companies, universities, etc; 5. to strengthen data sharing platforms and unify governance structures internationally.[106] Barriers to data sharing, addressed by these recommendations for regional outbreaks, are relevant to the COVID-19 pandemic and living systematic reviews such as this show that these are exacerbated where resources are lacking.[107]

6.2 Assessing longevity in comparison with the Zika Statement

Identifying suitable comparators for Zika research

Similar to the comparison with HVRD in the context of COVID-19, we have sought to benchmark Zika research against other relevant topics: in this case, our analysis considers "human vector-borne viral diseases" (HVVD). Here again, combining a comparison of disease groups with signatory and non-signatory research in each disease group results in four different breakdowns (signatory Zika research, signatory HVVD research, non-signatory Zika research and non-signatory HVVD research). These are shown in Figures 12 and 13 below, to provide context as to whether the Zika Statement has affected behaviours at the time of its release as well as over the following few years.

The Zika Statement had an impact on preprinting and data deposit as part of a data availability statement

Zika research published by or supported by signatories to the 2016 Zika Statement did see higher levels of open access or free-to-read status, preprinting, and data sharing practices than non-signatory Zika research.

That said, only for two dimensions were signatory Zika scores higher than HVVD signatory research:

- Figure 12 shows that 7% of signatory Zika journal publications were preceded by preprints overall between 2016 and 2020, versus 2% for non-signatory Zika research; and
- Figure 13 shows that 7% of signatory Zika journal publications included a data deposition mention as part of data availability statement, overall, between 2016 and 2020. This figure was 3% for non-signatory Zika journal publications over the same period. Data depositions as part of a data availability statement reached a high point of 10% of Zika signatory articles in 2019 (versus 5% for Zika non-signatory research that year).

The scores of HVVD signatory research had reached the same levels as those of signatory Zika research by 2020 for both preprinting and data deposition, suggesting that the Zika Statement's specific impacts had ended by that year. In other dimensions, such as OA and free-to-read status and data availability statement writing, there was no clear Zika Statement effect; as a result, the question of longevity does not apply in these cases.

The impact of the Zika statement on signatory behaviours is unlikely to have extended beyond 2018

In summary, a contribution of the Zika Statement to higher signatory preprinting and data deposition scores was likely for the years 2016 and 2017, but subsequent increases roughly mirrored those of the other analytical groups and are therefore likely to be less or not related to the Statement. Therefore, it appears the Zika Statement's effects were either of a small magnitude to start with or, where they reached greater magnitude, they did not persist in time after 2018.

Given our findings that the COVID-19 Statement signatories encompassed a much larger proportion of the corresponding research than the Zika Statement did, and also led to impacts of larger magnitudes than those of the Zika Statement, it now appears uncertain whether the latter can provide insight into the longevity of the former. While the Zika

Statement has seen limited effects and longevity, we note that this does not preclude that the COVID-19 Joint Statement will have a different outcome in the medium- or long-term.

Figure 12. Impact of the Zika statement on preprinting

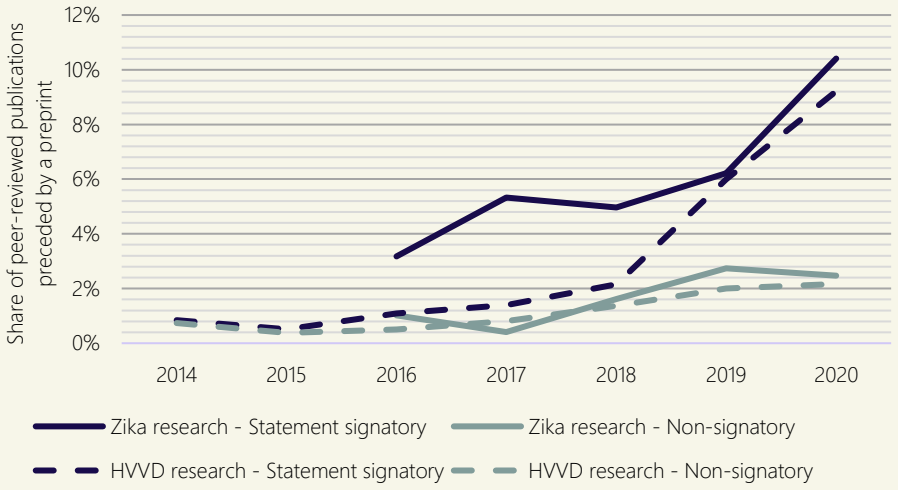
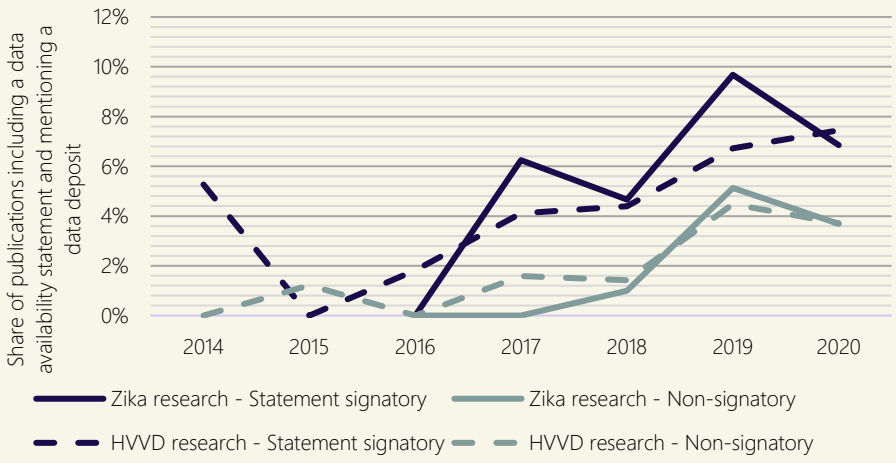


Figure 13. Impact of the Zika statement on data sharing behaviours



Part C. Building on the Joint Statement to address future crises

Research Consulting
Science-Metrix

7. Conclusions

7.1 The impact of openness during the pandemic

Background

Before outlining our conclusions, it is useful to go back to the project objectives noted at the beginning of this report. Our study sought to answer three key questions:

- Attribution: What is the evidence that signatories honoured the commitments made?
- Impact: What impact did open sharing have on public health response organisations, and to what extent is this seen as being linked to the Joint Statement?
- Longevity: What evidence is available that open sharing statements create long-lasting shifts in open research practices and behaviours?

The remainder of this section provides a summary of our findings and implications arising from these, mapping them to the three questions noted above via coloured boxes on the left-hand side.

The global health community greatly benefited from open sharing during the pandemic

Impact

Our consultation indicated that open and rapid sharing was a key success factor in the global pandemic response, alongside efforts to collaborate internationally and the availability of advanced research infrastructures. In particular, open sharing of genomic viral data allowed the prompt development of vaccines, and the broader availability of research articles and preprints helped inform policy in real time.

However, our findings show that Joint Statement signatories fell short of their commitments to help ensure that "research findings are made available via preprint servers before journal publication" and that "researchers share interim and final research data relating to the outbreak". Although descriptive bibliometric findings paint a positive picture of open sharing practices, particularly in comparison with the past, there is still significant room for improvement and closer alignment with the behaviours that signatories committed to.

In addition, the specific link between any improvements in open sharing and the Joint Statement was difficult to address, because (i) there was no pre-existing monitoring and evaluation framework for signatories; and (ii) it is complex to disentangle closely related phenomena and trends in the context of a fast-evolving pandemic (see the Assumptions and Inputs in Figure 1, p.8). Point (ii) above was clearly communicated by the public health response stakeholders we interviewed, as they noted that the Joint Statement was one of several initiatives and operated in a broader context of unusually high levels of collaboration and open research.

"I'm hoping that the research community has once again learned how important open science is... Whether this increased awareness will translate into action and have a lasting effect remains to be seen. At [funding organisation] we think changing the reward system such that engagement in open science is rewarded is key to the future success of open science." – Research funder

The Joint Statement has been a useful tool to foster short-term change, but its longevity may be limited

Impact
Longevity

The Joint Statement was released at the beginning of the pandemic, which helped shape the response of 160 national and international signatories at the right point in time. Many organisations signed the Joint Statement as it was in line with their organisational strategy and ethos, but this doesn't mean that the Joint Statement did not have an impact. Our interviews highlighted that the Joint Statement was useful to:

- align the efforts of key stakeholders;
- influence their policy requirements and organisational targets during the pandemic (and, potentially, beyond); and
- shape the long-term vision for open research.

Our comparison with the Zika statement indicates that open sharing statements may not have an impact beyond the horizon of the emergency they refer to. Nevertheless, open sharing trends have significantly shifted from the days of the Zika outbreak due to wider cultural shifts. While the pandemic has brought about extraordinary positive behaviours, embedding these in regular working practices depends on efforts to shift the wider research culture in favour of greater openness.

"Swift movement and greater acceptance of early sharing of data enabled [funding organisation] to progress work on the open research agenda, and bring other national funders along. The need for open science policies during the COVID-19 pandemic to remove obstacles to the free flow of research data and ideas has highlighted the benefits but also the challenges and gaps that exist in the national system and need to be addressed in the longer-term." – Research funder

7.2 The Joint Statement's ability to affect signatory behaviours

The Joint Statement operated in a broader context of culture change, but is likely to have had a distinct impact on signatory behaviours

Attribution

Over the course of this report, we have commented on the significant changes in the publishing and open science landscape brought by the COVID-19 pandemic. These changes were embraced by a wide range of stakeholders who wished to contribute to identifying a solution to the ongoing public health crisis and are in many cases independent of the Joint Statement. Nevertheless, there is evidence that it helped to shaped the response of individual signatories in a positive and consistent way.

Interviewees were somewhat sceptical about the socio-economic impact of the Joint Statement, as this acted in a much broader context of open sharing. On a similar note, the difference-in-differences analysis deployed by Science-Metrix could only conclude a *likely* contribution of the Joint Statement to the changes in practices and behaviours observed during the pandemic.

Bibliometric findings highlight changes in preprinting, policy-related

Controlled differential findings show that (difference-in-differences analysis, 2020-2021 vs 2018-2019):

- There was a clear increase in preprinting for signatory publications between 2018-19 and 2020-21. This increase is above an increase also recorded for non-signatory

uptake and publisher-based free-to-read

Attribution

publications. However, we acknowledge that the majority of COVID-19 publications were not shared as preprints prior to publication, indicating that the practice is still not widespread.

- Signatory COVID-19 journal publications benefitted from a greater increase in policy-related uptake compared against comparable signatory publications from 2018-19. Non-signatory publications also tended to record increases, but these were smaller than for signatory publications.
- Publisher-based free-to-read journal publications have seen an increase that can be linked back to one of the Joint Statement commitments.

On the other hand, the share of publications accessible via Gold OA has remained stable between periods for signatory publications. Given that non-signatory publications have seen an increase in Gold OA over the same period, the net result is a differential loss in Gold OA in signatory publications. Finally, the impact of the Joint Statement on data availability statements and data deposit could not be estimated. Although descriptive statistics did show increases in these areas, they could not be clearly attributed to the Joint Statement itself, and we note that overall levels of data sharing remain low across the board. Nevertheless, signatory organisations, and particularly research performing organisations and research funders, do appear more adept at fostering data sharing by researchers than non-signatory organisations, even if it is uncertain whether this ability has improved with the Joint Statement or has remained unchanged between 2018 and 2021.

Publishers were successful in implementing certain commitments in the Joint Statement, but compliance mechanisms were overall limited

Attribution

Our study focused on research funder and publisher signatories. Both stakeholder groups put in place mechanisms to operationalise the commitments in the Joint Statement, but publishers were particularly well-placed to do so given the close focus on research outputs. As publishers are most often responsible for hosting and preserving the scholarly record, they could more directly change whether this is accessible to the broader public or paywalled. Funders sought to influence behaviours via their grant requirements, but it is to be expected that they cannot address the openness commitment as directly as by removing a paywall. However, funders might have been well-positioned to address preprinting and data sharing commitments instead, with bibliometric findings showing research performing organisations and funder signatory publications to have achieved higher levels in these dimensions than signatory publications based solely on journal of publication.

The monitoring of the commitments was limited across the board, and compliance mechanisms weren't in place for either research funders or publishers: there were no sanctions from these organisations if relevant researchers, for example, did not share their data, include appropriate data availability statements or did not preprint their research. The lack of monitoring and compliance mechanisms also constrained this study's ability to fully assess the impact of the Joint Statement, and is a key opportunity for improvement for future statements.

Signatories are now well-placed to deploy prompt responses in future crises

Longevity

Signatories have built and deployed a wide range of workflows and mechanisms to align with Joint Statement commitments. Should future crises arise, regardless of whether they are in the biomedical field, we argue that research funders and publishers in particular, are well-placed to re-deploy these solutions. At the same time, we note that the operationalisation of Joint Statement commitments was ad hoc and varied between each

signatory organisation. There was no formal knowledge or good practice sharing among signatories, nor were there mechanisms in place for these organisations to learn from one another (although such exchanges did take place in other fora, such as the COVID-19 Rapid Review Initiative). A set of examples of how commitments can be operationalised by different signatory organisations (e.g. a research funder vs a publisher) would have been a valuable tool in enhancing alignment with the Joint Statement and deliver a more concerted response.

Table 3. Findings in response to the contribution analysis framework

Has the Joint Statement made an important contribution to the observed changes in open research behaviours and policies?	Why have the changes in open research behaviours and policies occurred? What role did the Joint Statement play?
Several Joint Statement signatories made policy and operational changes to align with the Statement's commitments. These included opening up access to research outputs as well as amending the wording of grant agreements for new awards. Research performing organisations also made some changes to internal practices, for example to provide extra support on data preparation/sharing or to allow preprint citation in press releases. Signatories highlighted that the Joint Statement was useful to: • align the efforts of key stakeholders; • influence their policy requirements and organisational targets during the pandemic (and, potentially, beyond); and • shape the long-term vision for open science.	This question is difficult to answer definitively, as the desire of signatory organisations to support the global pandemic response is a major confounding factor. Almost 90% of signatories were motivated to sign the Joint Statement as it was in line with existing organisational efforts around COVID-19, and some indicate that the commitments outlined in the Joint Statement were "the right thing to do". The Joint Statement, however, did provide a focal point for organisations to shape their response: signatories aligned their behaviours to support or meet the commitments in the Joint Statement. It may be hypothesised that, without the Joint Statement, the response of the signatories would have been overall less coherent and focused on facets of open sharing perceived as important by each individual organisation.
Is it reasonable to conclude that the Joint Statement has made a difference?	**What conditions are needed to make this type of Joint Statement succeed?**
Our evidence indicates that the Joint Statement did have an impact, but its implementation fell short of the ambitious objectives in some dimensions (chiefly preprinting and data sharing). The need to establish an assessment approach post-hoc also led to evaluation challenges, such as the difficulty of attributing some changes observed to Joint Statement commitments. The range of high-profile signatories, alongside the evidence we gathered, suggests that similar statements may be used in future outbreaks, pandemics or other large-scale emergencies as a tool to align stakeholders and support international government bodies. Nevertheless, it is difficult, if not impossible, to reliably attribute specific socio-economic impacts to the Joint Statement.	Our research indicates that several elements contributed to the impact of the Joint Statement: it was aligned with plans that signatory organisations were considering in the first place; it was coordinated by a highly recognised and credible organisation at the global level, with previous experience with a similar statement; and it was released at the right point in time, very early on in the pandemic and before other initiatives emerged. Our study has also identified key pre-conditions for the success of future statements: the inclusion of guidance on the operationalisation of the commitments; and the establishment of a monitoring and evaluation framework to assess the impact of the statement.

8. Lessons learned and recommendations

The elements of a successful open sharing statement

The evidence we gathered suggests that similar statements could helpfully be developed in the future: we have gathered both qualitative and bibliometric evidence supporting the value of the COVID-19 Joint Statement in aligning stakeholders and influencing present and future open science requirements and policy objectives.

We argue that the above was made possible by some specific features of the Joint Statement:

- the organisation behind the development of the statement – Wellcome – is a highly recognised and credible organisation at the global level, which conferred an extent of authority and prestige on the Joint Statement;
- Wellcome was behind a previous statement released in the context of the Zika virus outbreak, showcasing the execution of a similar initiative in the past and providing a starting cohort of prospective signatories for the COVID-19 iteration; and
- the Joint Statement was released at the right point in time, very early on in the pandemic and before other initiatives emerged, giving it more visibility and impetus.

In addition, one should not forget that signatories, at least in most cases, already agreed with the ethos of the Joint Statement, and therefore had limited reasons not to sign it. This is a reminder that the Joint Statement had a positive focus on joining forces to address a global pandemic and was framed as a voluntary commitment rather than a formal requirement for behavioural or policy changes. By signing, organisations agreed to short-to-medium action that, in many cases, they might have pursued either way, and were not subjected to monitoring or enforcement of the commitments made.

In practice, the prestige and high profile of the Joint Statement also made it a positive news story for signatories, which allowed an extent of marketing and branding via organisational websites. The fact that there was something for organisations to gain by signing the Joint Statement is not a drawback: the Joint Statement's value proposition was clear, which is likely part of the reason why it attracted support from prominent international stakeholders.

However, as noted above, some of the objectives in the Joint Statement were not achieved by signatories, chiefly with regard to the expected penetration of open research practices such as preprint posting and data sharing. This indicates that, while successful in several aspects, the Joint Statement could have been more effective in communicating clear operational targets to the cohort of signatories. To enhance the odds of success of future initiatives, the next section summarises a series of recommendations and lessons learned.

Recommendations for future statements

Building on the findings of this report, we recommend that organisations developing open sharing statements in the future should (see Table 4 for more information):

1. Strengthen expectations in relation to the sharing of research data and preprints: While the Joint Statement was effective in increasing access to journal publications and ensuring they were shared with the WHO upon journal submission, low levels of data sharing and preprint posting indicate that these commitments remained largely unmet. Given the demonstrable value of these practices in responding to public health emergencies, future statements should

include strengthened expectations on data sharing and preprint posting, with clear guidance on how these should be achieved.

2. **Establish a monitoring and evaluation framework to assess the level of success of the initiative:** The present report required the development of a post-hoc assessment framework to review of the impact of the Joint Statement. The publication of future statements should be accompanied (either immediately or within a matter of weeks) by a clear overview of their intended impact in the short, medium and long term (theory of change) and a range of success factors that could be used for self-auditing purposes. This should ideally include a list of indicators or data points that signatories are expected to collect based on the type of organisation. It is expected that a focus on specific objectives or targets may dissuade some organisations from signing, but that increased signatory alignment with statement commitments would deliver a greater overall impact.
3. **Present commitments for signatories in a more granular way, including guidance on their operationalisation:** The Joint Statement has been signed by a very diverse cohort, and its commitments are described broadly and not tailored to different actors in the higher education and research landscape. We recognise that this can enable stakeholders with less established open sharing practices to sign the statement, as it helps keep the text brief and highly accessible. However, the lack of clarity may also create confusion among signatories, so additional guidance is seen as beneficial. To support signatories, an accompanying guidance document should be provided with more detailed suggestions on the operationalisation of commitments, potentially building on the experiences of signatories of the COVID-19 and Zika Statements and including tailored advice for different types of organisations.
4. **Build knowledge sharing mechanisms to foster learning and collaboration between signatories:** To support the effective operationalisation of future statements, knowledge sharing mechanisms should be established to share examples and enable signatories to learn from each other's experiences and challenges. This would build on the operational guidance noted in recommendation 3 above and would help signatories to more easily align their commitments. It is essential that any examples and discussions between signatories are contextualised, and that there is no implication that one-size-fits-all solutions are appropriate.
5. **Carefully consider the long-term policy impact of specific statement commitments:** The Joint Statement led to an increase in free-to-read materials and a net differential loss of share in Gold OA by signatories. Although this has been hugely beneficial in the context of the pandemic, free-to-read materials may be placed behind paywalls at any point in time. As a result, the impact of the Joint Statement on open sharing cultures may remain limited to the context of the pandemic as opposed to being cemented in the behaviours of relevant stakeholder groups.
6. **Assess the impact of statement commitments on low- and middle-income countries:** Open research (including articles, data, preprints and more) can be more easily leveraged by high-income countries, given the more significant levels of investment, infrastructure and resources available. As a result, the benefits arising from the implementation of the Joint Statement are liable to be reaped asymmetrically, with LMICs not feeling equally incentivised to share information.

Future statements should explicitly promote equitable access to and re-use of research outputs from low- and middle-income countries.

We expect that, by considering the above recommendations and building on the experiences of the Zika and COVID-19 statements, the impact, reach and inclusiveness of future initiatives will be even higher. This will help the higher education and research community achieve more concerted and equitable outcomes at the international level and enhance the chances of a prompt resolution to future global crises.

Table 4. Examples of areas where future open sharing statements could be improved

Recommendation	Opportunities for improvement
1. Strengthen expectations in relation to the sharing of research data and preprints	Publishers and funders could be asked to implement more formal expectations around data sharing and the posting of preprints. This could be achieved by requiring a digital object identifier and/or accession number for data, and stipulating the posting of preprints as a condition of funding or article acceptance, where the funding scheme or article subject matter relate to the emergency response.
2. Establish a monitoring and evaluation framework to assess the level of success of the initiative	The present report is a post-hoc evaluation of the Joint Statement. As such, it built on an evidence framework (chiefly in the form of the theory of change and contribution analysis) that was assembled after the launch of the Joint Statement. Although this was partly due to the pace of the COVID-19 pandemic, it is expected that clearer target setting a priori would lead to more straightforward data collection and monitoring exercises but also to higher signatory alignment with the commitments made. An audit framework, for example, could have led to increased success in areas where the Joint Statement failed to fully deliver, e.g. preprinting and data sharing.
3. Present commitments for signatories in a more granular way, including guidance on their operationalisation	A research performing organisation may wonder what their role is with regard to sharing relevant outputs with WHO upon journal submission, as they typically have limited control over, and awareness of, what their researchers submit for publication. Where an organisation's role is clear, implementation mechanisms may vary based on stakeholder type: for example, a research performing organisation and a publisher may both play a role with regard to making all journal publications immediately open access or free-to-read. However, the former may rely on dedicated funding pots or pathways such as green open access, while the latter may consider removing paywalls from an appropriate portion of their contents or waiving article processing charges.
4. Build knowledge sharing mechanisms to foster learning and collaboration between signatories	Signatory organisations may require further guidance and support to build on the operational guidance described above. For example, a publisher with a more advanced and automated submission system may be better placed to operationalise a given commitment. Sharing their knowledge with peer signatories may help these strengthen their systems and more

	easily implement a commitment (as opposed to potentially reinventing the wheel, with significant investment of time and resources and a delay to their contribution to the initiative).
5. Carefully consider the long-term policy impact of specific statement commitments	Many publishers chose the free-to-read pathway rather than open access (and openly licensed) options during the COVID-19 pandemic. Several have already started re-introducing paywalls.
6. Assess the impact of statement commitments on low- and middle-income countries	Joint Statement commitments expected researchers from low- and middle-income countries to share data and findings openly. While the academic and policy-related benefits of open sharing and big data approaches could be immediately leveraged by high income countries, the Joint Statement did not address how low- and middle-income countries where digital infrastructure is typically limited could achieve similar gains. Future statements should include specific commitments to ensure appropriate credit for researchers whose work has been re-used, including those in low- and middle-income countries.

Glossary

COVID-19	Coronavirus disease (COVID-19), an infectious disease caused by the SARS-CoV-2 virus.
Data availability statement	A statement specifying whether data behind a research project can be accessed and, if so, where and how.
Difference-in-differences	Data analysis approach comparing the changes in outcomes over time between a population enrolled in a program (the treatment group) and a population that is not (the comparison group).
Publisher-based free-to-read	Published version of record or manuscript accepted for publication, for which the publisher has chosen to provide temporary or permanent free access, without the use of open licensing.
Intellectual property	Intangible property that is the result of creativity, such as patents, copyrights and trademarks.
Joint Statement	The Joint Statement is the statement published as a press release on the Wellcome website and titled "Sharing research data and findings relevant to the novel coronavirus (COVID-19) outbreak". It is the main focus of the present work.
Open access	A publication is considered in open access (OA) if: • its content is universally and freely accessible, at no cost to the reader, via the Internet or otherwise; • the author or copyright owner irrevocably grants to all users, for an unlimited period, the right to use, copy, or distribute the article, on condition that proper attribution is given; • it is deposited, immediately, in full and in a suitable electronic form, in at least one widely and internationally recognized open access repository committed to open access. There are different pathways to OA, including: • green OA: published version or manuscript accepted for publication, available via a repository. Other OA versions may also be available via gold or free-to-read on the publisher platform. • gold OA: published version, typically with a Creative Commons license, available on a publisher platform. The document is in a journal that only publishes via open access. • hybrid OA: published version, typically with a Creative Commons license, available on a publisher platform. The document is in a journal that provides authors the choice of publishing open access or via the subscription model.
Peer review	The process of subjecting an author's manuscript to the scrutiny of others who are experts in the same field, prior to publication in a journal, conference proceedings or a book.
Journal publications	Journal articles, conference papers and reviews.
Preprint	A public draft of a scientific document, posted publicly before peer review.
Protocols and standards	Written procedures or guidelines that outline the methodology of the research study, as well as good and ethical practices that should be observed when conducting research.
Research articles	Sum of journal publications and preprints.

Research data	Information that has been collected, observed, generated or created to reach or validate original research findings.
Repository	An archive for collecting, preserving, and disseminating digital copies of the intellectual output of researchers.

References

1. Wellcome Trust. Coronavirus (COVID-19): sharing research data. *Wellcome* https://wellcome.org/press-release/sharing-research-data-and-findings-relevant-novel-coronavirus-ncov-outbreak (2020).
2. Wellcome Trust. Publishers make coronavirus (COVID-19) content freely available and reusable. *Wellcome* https://wellcome.org/press-release/publishers-make-coronavirus-covid-19-content-freely-available-and-reusable (2020).
3. National Science and Technology Advisors. National Science and Technology Advisors COVID-19 Request. (2020).
4. Connell, J. & Kubisch, A. Applying a Theory of Change Approach to the Evaluation of Comprehensive Community Initiatives_ Progress, Prospects, and Problems.pdf. https://www.eccnetwork.net/sites/default/files/media/file/Applying%20a%20Theory%20of%20Change%20Approach%20to%20the%20Evaluation%20of%20Comprehensive%20Community%20Initiatives_%20Progress%2C%20Prospects%2C%20and%20Problems.pdf (1998).
5. Mayne, J. *Contribution Analysis: An approach to exploring cause and effect*. https://web.archive.org/web/20150226022328/http:/www.cgiar-ilac.org/files/ILAC_Brief16_Contribution_Analysis_0.pdf (2015).
6. Loffreda, L., Chiarelli, A. & Vignola-Gagné, E. Sharing research data and findings relevant to the novel coronavirus (COVID-19) outbreak - Data Management Plan. *Zenodo* (2022) doi:10.5281/zenodo.5848642.
7. Wellcome Trust. Statement on data sharing in public health emergencies. *Wellcome* https://wellcome.org/press-release/statement-data-sharing-public-health-emergencies (2016).
8. GLOPID-R. *Principles of Data Sharing in Public Health Emergencies*. 160018 http://www.nature.com/articles/sdata201618 (2018).
9. National Library of Medicine. Public Health Emergency COVID-19 Initiative - PMC. *National Institutes of Health* https://www.ncbi.nlm.nih.gov/pmc/about/covid-19/?cmp=1 (2021).
10. Open COVID Pledge. Open COVID Pledge. *Open Covid Pledge* https://opencovidpledge.org/ (2020).
11. COVID-19 Data Portal. Open letter: Support data sharing for COVID-19. *COVID-19 Data Portal* https://www.covid19dataportal.org/support-data-sharing-covid19 (2020).

12. World Health Organization. Solidarity Call to Action. *World Health Organization* https://www.who.int/initiatives/covid-19-technology-access-pool/solidarity-call-to-action (2022).
13. UNESCO. Joint Appeal for Open Science. *UNESCO* https://en.unesco.org/news/joint-appeal-open-science (2020).
14. The White House. CARBIS BAY G7 SUMMIT COMMUNIQUÉ. *The White House* https://www.whitehouse.gov/briefing-room/statements-releases/2021/06/13/carbis-bay-g7-summit-communique/ (2021).
15. G7. G7 Research Compact. https://www.g7uk.org/wp-content/uploads/2021/06/G7-2021-Research-Compact-PDF-356KB-2-pages-1.pdf (2021).
16. National Institute for Health and Care Research. Shared commitment to public involvement. https://www.nihr.ac.uk/documents/shared-commitment-to-public-involvement/30134 (2022).
17. Astell, M. Sharing findings related to COVID-19 in these extraordinary times. *Hindawi* https://www.hindawi.com/post/sharing-findings-related-covid-19-these-extraordinary-times/ (2020).
18. Health Infection Society. Novel Coronavirus: HIS statement on sharing research data and findings - Healthcare Infection Society. *Healthcare Infection Society* https://www.his.org.uk/resources-guidelines/covid-19/ (2020).
19. SAGE Publishing. SAGE Publishing Statement on the COVID-19 Pandemic. *SAGE Publishing* https://group.sagepub.com/press-releases/sage-publishing-statement-on-the-covid-19-pandemic (2020).
20. The BMJ. Coronavirus (covid-19) Hub - Latest News & Research. *The BMJ* https://www.bmj.com/coronavirus (2022).
21. Elsevier. Novel Coronavirus Information Center. *Elsevier* https://www.elsevier.com/connect/coronavirus-information-center (2020).
22. Cell Press. Cell Press Coronavirus Resource Hub. *Cell Press* https://www.cell.com/COVID-19, https://www.cell.com/COVID-19 (2022).
23. The Lancet. COVID-19 Resource Centre. *The Lancet* https://www.thelancet.com/coronavirus (2022).
24. American Medical Association. About Coronavirus Disease 2019 (COVID-19) - Information from JAMA Network, the CDC, and WHO. *JAMA Network* https://jamanetwork.com/journals/jama/pages/coronavirus-alert (2022).
25. Healthcare Infection Society. Home Page: Journal of Hospital Infection. *Healthcare Infection Society* https://www.journalofhospitalinfection.com/ (2022).
26. The New England Journal of Medicine. Coronavirus (Covid-19) — NEJM. *The New England Journal of Medicine* https://www.nejm.org/coronavirus?query=RP (2022).

27. Oxford University Press. Access to OUP resources on COVID-19, other coronaviruses, and related topics. *Oxford Academic* https://academic.oup.com/journals/pages/coronavirus (2020).
28. PLOS. COVID-19 pandemic (2019-21) - PLOS Collections. *PLOS* https://collections.plos.org/collection/covid-19/ (2020).
29. SAGE Journals. Coronavirus (COVID-19) research. *SAGE Journals* https://journals.sagepub.com/coronavirus (2022).
30. Springer Nature. Coronavirus (COVID-19) Research Highlights | For Researchers | Springer Nature. *Springer Nature* https://www.springernature.com/gp/researchers/campaigns/coronavirus (2022).
31. Taylor & Francis. COVID-19: Novel Coronavirus Content Free to Access. *Taylor & Francis* https://taylorandfrancis.com/coronavirus/ (2020).
32. Wiley Online Library. Novel Coronavirus - Wiley Online Library. *Wiley Online Library - Coronavirus Resources & News* https://novel-coronavirus.onlinelibrary.wiley.com/ (2020).
33. Cold Spring Harbor Laboratory. COVID-19 SARS-CoV-2 preprints from medRxiv and bioRxiv. *bioRxiv* https://connect.biorxiv.org/relate/content/181 (2022).
34. Elsevier. Elsevier gives full access to its content on its COVID-19 Information Center for PubMed Central and other public health databases to accelerate fight against coronavirus. *Elsevier* https://www.elsevier.com/about/press-releases/corporate/elsevier-gives-full-access-to-its-content-on-its-covid-19-information-center-for-pubmed-central-and-other-public-health-databases-to-accelerate-fight-against-coronavirus (2020).
35. Springer Nature. Novel Coronavirus - (2019-nCoV): Live resource hub | Corporate Affairs Homepage | Springer Nature. *Springer Nature Group* https://group.springernature.com/gp/group/media/press-releases/archive-2020/springer-nature-coronavirus/17607604 (2020).
36. SAGE Publishing. SAGE Waives Article Processing Charges for Research Related to COVID-19. *SAGE Publications Ltd* https://uk.sagepub.com/en-gb/eur/press/sage-waives-article-processing-charges-for-research-related-to-covid-19 (2020).
37. UKRI. Call Conditions - to be added by Calls and schemed team will pull through to each award. (2020).
38. Health Research Board. COVID-19 Pandemic Rapid Response Funding Call (COV19 2020). https://web.archive.org/web/20200404052431/https:/www.hrb.ie/funding/funding-schemes/all-funding-schemes/grant/covid-19-pandemic-rapid-response-funding-call-cov19-2020/ (2020).

39. Health Research Board & Irish Research Council. COVID-19 Pandemic Rapid Response Funding Opportunity (COV19) 2020. https://web.archive.org/web/20211211093245/https:/www.hrb.ie/fileadmin/2._Plugin_related_files/Funding_schemes/HRB-IRC_COV19-2020_Guidance_Notes_-_update_1-4-2020.pdf (2021).
40. HRB Open Research. Open Access Publishing Platform | Beyond a Research Journal. https://www.hrbopenresearch.org/ (2020).
41. Health Research Board. Management and Sharing of Research Data. https://www.hrb.ie/funding/funding-schemes/before-you-apply/all-grant-policies/management-and-sharing-of-research-data/ (2019).
42. Europe PMC. Preprints in Europe PMC. *Europe PMC* https://europepmc.org/Preprints#about-including-preprints (2022).
43. National Library of Medicine. NIH Preprint Pilot - PMC. https://www.ncbi.nlm.nih.gov/pmc/about/nihpreprints/ (2020).
44. Funk, K. The NIH Preprint Pilot: A New Experiment for a New Era. *NLM Musings from the Mezzanine* https://nlmdirector.nlm.nih.gov/2020/06/09/the-nih-preprint-pilot-a-new-experiment-for-a-new-era/ (2020).
45. National Library of Medicine. 'Preprint[filter]' - PMC - NCBI. https://www.ncbi.nlm.nih.gov/pmc/?term=preprint%5Bfilter%5D (2022).
46. Couture, J. L., Blake, R. E., McDonald, G. & Ward, C. L. A funder-imposed data publication requirement seldom inspired data sharing. *PLOS ONE* **13**, e0199789 (2018).
47. Le Guillou, I. Covid-19: How unprecedented data sharing has led to faster-than-ever outbreak research | Research and Innovation. *Horizon: The EU Research & Innovation Magazine* https://ec.europa.eu/research-and-innovation/en/horizon-magazine/covid-19-how-unprecedented-data-sharing-has-led-faster-ever-outbreak-research (2020).
48. Kupferschmidt, K. 'A completely new culture of doing research.' Coronavirus outbreak changes how scientists communicate. https://www.science.org/content/article/completely-new-culture-doing-research-coronavirus-outbreak-changes-how-scientists (2020).
49. Else, H. How a torrent of COVID science changed research publishing — in seven charts. *Nature* **588**, 553–553 (2020).
50. UKCDR & GLOPID-R. COVID-19 Research Project Tracker by UKCDR & GloPID-R. https://www.ukcdr.org.uk/covid-circle/covid-19-research-project-tracker/.

51. Bell, K. & Green, J. Premature evaluation? Some cautionary thoughts on global pandemics and scholarly publishing. *Crit. Public Health* **30**, 379–383 (2020).
52. Redhead, C. Scholarly publishers are working together to maximize efficiency during COVID-19 pandemic. *OASPA* https://oaspa.org/scholarly-publishers-working-together-during-covid-19-pandemic/ (2020).
53. Horbach, S. P. J. M. Pandemic publishing: Medical journals strongly speed up their publication process for COVID-19. *Quant. Sci. Stud.* **1**, 1056–1067 (2020).
54. Horbach, S. How the pandemic changed editorial peer review – and why we should wonder whether that's desirable. *Impact of Social Sciences* https://blogs.lse.ac.uk/impactofsocialsciences/2021/02/10/how-the-pandemic-changed-editorial-peer-review-and-why-we-should-wonder-whether-thats-desirable/ (2021).
55. Piller, C. Many scientists citing two scandalous COVID-19 papers ignore their retractions. *Science* https://www.science.org/content/article/many-scientists-citing-two-scandalous-covid-19-papers-ignore-their-retractions (2021).
56. Lee, J. J. & Haupt, J. P. Scientific globalism during a global crisis: research collaboration and open access publications on COVID-19. *High. Educ.* **81**, 949–966 (2021).
57. UKCDR & GLOPID-R. COVID CIRCLE. *UKCDR* https://www.ukcdr.org.uk/covid-circle/ (2021).
58. COVID-19 Clinical Research Coalition. Home | COVID-19 Clinical Research Coalition. *COVID-19 Clinical Research Coalition* https://covid19crc.org/ (2022).
59. COVID R&D Alliance. COVID-19 Research and Development Alliance. *Covid R&D Alliance* https://www.covidrdalliance.com (2022).
60. ECRIN. COVID-19 Taskforce | ECRIN. *ECRIN - European Clinical Research Infrastructure Network* https://ecrin.org/covid-19-taskforce (2022).
61. Norton, A. J. *et al.* Priorities for COVID-19 research response and preparedness in low-resource settings. *The Lancet* **397**, 1866–1868 (2021).
62. VODAN Africa. Home | VODAN Africa & Asia. *VODAN Africa* https://www.vodan-totafrica.info/ (2022).
63. National Center for Advancing Translational Sciences. Collaboration Speeds COVID-19 Research | National Center for Advancing Translational Sciences. *National Institutes of Health*

https://web.archive.org/web/20220316143120/https:/ncats.nih.gov/covid19-translational-approach/collaboration (2021).

64. Kwon, D. How swamped preprint servers are blocking bad coronavirus research. *Nature* **581**, 130–131 (2020).
65. Lu, R. *et al.* Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. *The Lancet* **395**, 565–574 (2020).
66. Dey, S. *et al.* Impact of Clinical and Genomic Factors on SARS-CoV2 Disease Severity. 2021.03.15.21253549 (2021) doi:10.1101/2021.03.15.21253549.
67. Cook, S. Genomic surveillance in the roll out of vaccines. *University of Cambridge* https://www.phgfoundation.org/blog/genomic-surveillance-in-the-roll-out-of-vaccines (2021).
68. González-Zapata, F. *et al.* Open Data in Action: Initiatives during the Initial Stage of the COVID-19 Pandemic. *SSRN Electron. J.* (2021) doi:10.2139/ssrn.3937613.
69. Hillary, L. S. *et al.* SARS-CoV-2 qPCR and water chemistry data from wastewater at six sites in England and Wales between March and July 2020, with associated Covid-19 positive test and related-death data collated from other sources. *UK CEH* https://catalogue.ceh.ac.uk/id/ce40e62a-21ae-45b9-ba5b-031639a504f7 (2021).
70. CDC. National Wastewater Surveillance System. *Centers for Disease Control and Prevention* https://www.cdc.gov/healthywater/surveillance/wastewater-surveillance/wastewater-surveillance.html (2022).
71. GOV UK. Ethnicity and COVID-19 from Race Disparity Unit. *GOV UK* https://www.ethnicity-facts-figures.service.gov.uk/covid-19 (2022).
72. Science, D. *et al. The State of Open Data 2021*. https://digitalscience.figshare.com/articles/report/The_State_of_Open_Data_2021/17061347/1 (2021) doi:10.6084/m9.figshare.17061347.v1.
73. Strcic, J. *et al.* Open data and data sharing in articles about COVID-19 published in preprint servers medRxiv and bioRxiv. *Scientometrics* (2022) doi:10.1007/s11192-022-04346-1.
74. ICODA. ICODA Home. *ICODA - A globally coordinated, health data-led research response to tackle the COVID-19 pandemic*. https://icoda-research.org/ (2022).

75. ICODA. The urgent need for better coordination across the global data sharing landscape. *ICODA - A globally coordinated, health data-led research response to tackle the COVID-19 pandemic*. https://icoda-research.org/the-urgent-need-for-better-coordination-across-the-global-data-sharing-landscape/ (2021).
76. COVID-19 Data Portal. COVID-19 Data Portal - accelerating scientific research through data. https://www.covid19dataportal.org/ (2022).
77. Vivli. Vivli - Center for Global Clinical Research Data. *Vivli Center for Global Clinical Research Data* https://vivli.org/ (2022).
78. Global Alliance for Genomics & Health. Enabling responsible genomic data sharing for the benefit of human health. *Global Alliance for Genomics & Health* https://www.ga4gh.org/ (2022).
79. ICODA. ICODA Partners. *ICODA - A globally coordinated, health data-led research response to tackle the COVID-19 pandemic*. https://icoda-research.org/partners/ (2022).
80. ICODA. ICODA - Portal. *ICODA* https://portal.covid-19.aridhia.io/ (2020).
81. ICODA. Publications. *ICODA - A globally coordinated, health data-led research response to tackle the COVID-19 pandemic*. https://icoda-research.org/research/publications/ (2020).
82. Scholz, A. H. *et al.* Multilateral benefit-sharing from digital sequence information will support both science and biodiversity conservation. *Nat. Commun.* **13**, 1086 (2022).
83. UK Health Security Agency. UK completes over 2 million SARS-CoV-2 whole genome sequences. *GOV.UK* https://www.gov.uk/government/news/uk-completes-over-2-million-sars-cov-2-whole-genome-sequences (2022).
84. GISAID. GISAID EpiFlu Database Access Agreement. *GISAID* https://www.gisaid.org/registration/terms-of-use/ (2011).
85. GISAID. GISAID - Publish with data from GISAID. *GISAID* https://www.gisaid.org/help/publish-with-data-from-gisaid/ (2022).
86. Van Noorden, R. Scientists call for fully open sharing of coronavirus genome data. *Nature* **590**, 195–196 (2021).
87. GISAID. GISAID - Acknowledgements. *GISAID* https://www.gisaid.org/about-us/acknowledgements/ (2022).
88. International Nucleotide Sequence Database Collaboration. International Nucleotide Sequence Database Collaboration | INSDC. *International Nucleotide Sequence Database Collaboration* https://www.insdc.org/ (2022).

89. Bawden, D. & Robinson, L. Information Overload: An Introduction. *Oxford Research Encyclopedia of Politics* https://oxfordre.com/politics/view/10.1093/acrefore/9780190228637.001.0001/acrefore-9780190228637-e-1360 (2020) doi:10.1093/acrefore/9780190228637.013.1360.

90. CDC. CDC Library: COVID-19 Science Update: 12/17/2021. *Centers for Disease Control and Prevention* https://www.cdc.gov/library/covid19/12172021_covidupdate.html (2021).

91. World Health Organization. Global research on coronavirus disease (COVID-19). *World Health Organization* https://www.who.int/emergencies/diseases/novel-coronavirus-2019/global-research-on-novel-coronavirus-2019-ncov (2022).

92. CDC. CDC Library: COVID-19 Science Update: 04/01/2020. *CDC* https://www.cdc.gov/library/covid19/040120_covidupdate.html (2021).

93. Scientific Advisory Group for Emergencies. List of participants of SAGE and related sub-groups. *GOV.UK* https://www.gov.uk/government/publications/scientific-advisory-group-for-emergencies-sage-coronavirus-covid-19-response-membership/list-of-participants-of-sage-and-related-sub-groups (2022).

94. The White House. Statement by President Joe Biden on the Omicron COVID-19 Variant. *The White House* https://www.whitehouse.gov/briefing-room/statements-releases/2021/11/26/statement-by-president-joe-biden-on-the-omicron-covid-19-variant/ (2021).

95. Fraser, N. *et al.* The evolving role of preprints in the dissemination of COVID-19 research and their impact on the science communication landscape. *PLOS Biol.* **19**, e3000959 (2021).

96. Cole, N. L., Reichmann, S. & Ross-Hellauer, T. Global Thinking. ON-MERRIT recommendations for maximising equity in open and responsible research. (2022) doi:10.5281/zenodo.6276753.

97. Global Health Network. Home • Coronavirus. https://coronavirus.tghn.org/ (2022).

98. John Hopkins University & Medicine. COVID-19 Dashboard Map. *Johns Hopkins Coronavirus Resource Center* https://coronavirus.jhu.edu/map.html.

99. Dong, E., Du, H. & Gardner, L. An interactive web-based dashboard to track COVID-19 in real time. *Lancet Infect. Dis.* **20**, 533–534 (2020).

100. Gardner, L., Ratcliff, J., Dong, E. & Katz, A. A need for open public data standards and sharing in light of COVID-19. *Lancet Infect. Dis.* **21**, e80 (2021).

101. Callaway, E. Will the pandemic permanently alter scientific publishing? *Nature* **582**, 167–168 (2020).

102. Review Commons. Review Commons Homepage. *Review Commons* https://www.reviewcommons.org/.

103. Barbour, V. Openness, Integrity, Inclusion, and Innovation in Scholarly Communication: Competing or Complementary Forces? *Front. Res. Metr. Anal.* **6**, (2022).

104. GLOPID-R. GloPID-R - Global Research Collaboration for Infectious Disease Preparedness. *GLOPID-R* https://www.glopid-r.org/ (2022).

105. GLOPID-R, Wellcome Trust, & UKaid. *Data Sharing in Public Health Emergencies: Learning from Past Outbreaks*. 1–33 https://www.glopid-r.org/wp-content/uploads/2017/02/data-sharing-in-public-health-emergencies-case-studies-workshop-reportv2.pdf (2018).

106. Norton, A., Pardinaz-Solis, R. & Carson, G. *GLOPID-R Roadmap for Data Sharing in Public Health Emergencies*. https://www.glopid-r.org/wp-content/uploads/2019/06/glopid-r-roadmap-for-data-sharing.pdf (2020).

107. Pratt, B. & Bull, S. Equitable data sharing in epidemics and pandemics. *BMC Med. Ethics* **22**, 136 (2021).

Appendix A. Project contributors

The stakeholders listed in Table A1 contributed to our project via interviews.

Table A1. Project contributors.

Name	Organisation	Role
Anthony Philippakis	Broad Institute of MIT and Harvard	Chief Data Officer
Ben Morton	Malawi-Liverpool Wellcome Trust	Clinical Lead
Brian A. King	Centers for Disease Control and Prevention, USA	Chief Science Officer
Caroline Barelle	Elasmogen Ltd	CEO
Charles Mowbray	Drugs for Neglected Diseases Institute	Discovery Director
Filip Pattyn	Ontoforce	Scientific Advisor
Linus Grabenhenrich	Robert Koch Institut	Head of Research Data and Senior Epidemiologist
Matthew Todd	University College London	Professor of Drug Discovery
Michael Ferguson	Wellcome Centre for Anti-Infectives Research, University of Dundee	Regius Professor of Life Sciences
Oliver Morgan	World Health Organization	Director of Health Emergency Information and Risk Assessment
Rajeev Venkayya	Takeda Pharmaceutical Company Limited	MD, President, Global Vaccine Business Unit
Sarah Gilbert	Jenner Institute, Oxford	Saïd Professorship of Vaccinology (DBE)
Sofonias Kifle Tessema	Africa Centers of Disease Control and Prevention	Implementation Science Expert
Trudie Lang	The Global Health Network, Oxford	Professor of Global Health Research
Vasee Moorthy	World Health Organization	Senior Advisor, Research and Development
Vikram Savkar	Wolters Kluwer	Senior Vice President and General Manager

Appendix B. Joint Statement signatories

The organisations listed in Table B1 signed the Joint Statement. The 'Stakeholder type' column has been added for the purposes of this study to enable further analysis and is not provided via the list of signatories on the Wellcome website.

Table B1. Joint Statement signatories.

Organisation	Stakeholder type
Abasyn University Peshawar, Pakistan	University
Academy of Medical Sciences, UK	Research funder
Africa Centres for Disease Control and Prevention (Africa CDC)	Public health agency
African Academy of Sciences	Learned society
African Journal of Laboratory Medicine	Journal
American Chemical Society (ACS)	Learned society
American Physical Society (APS)	Learned society
American Society for Microbiology	Learned society
Annals of Internal Medicine	Journal
ARTiFACTS	Service provider
arXiv	Preprint server
Assistant Secretary for Preparedness and Response / Biodefense Advanced Research and Development Authority, USA	Government agency or department
Australian and New Zealand Intensive Care Society (ANZICS)	Research performing organisation
Austrian Science Fund (FWF)	Research funder
Ayass Bioscience LLC	Biotech industry company
BenevolentAI	Biotech industry company
Bernhard Nocht Institute for Tropical Medicine	Research performing organisation
Bill & Melinda Gates Foundation	Research funder
Biochemical Society & Portland Press	Society publisher
BioRxiv	Preprint server
Broad Institute of MIT and Harvard	Research performing organisation
Brunel University	University
Bulletin of the World Health Organization	Journal
CABI	Not-for-profit
Calouste Gulbenkian Foundation – Gulbenkian Science Institute	Research performing organisation
Cambridge University Press (CUP)	University press
Canada Foundation for Innovation	Research funder
Canadian Institutes of Health Research	Research funder
Canadian Science Publishing	Publisher
Cancer Research UK	Research funder
Cell Press	Publisher
Center for Biomedical Research Transparency (CBMRT)	Not-for-profit
Centers for Disease Prevention and Control (CDC)	Public health agency
CEPI	Research funder

Organisation	Stakeholder type
Certara	Pharmaceutical industry company
Chinese Centre for Disease Control and Prevention	Public health agency
Chinese Journal of Lung Cancer	Journal
CIFAR	Research performing organisation
Cochrane	Not-for-profit
Cold Spring Harbor Laboratory Press	Publisher
Company of Biologists	Publisher
CONCYTEC – National Council for Science, Technology and Technological Innovation, Peru	Research funder
Department of Biotechnology, Ministry of Science & Technology, Government of India	Government agency or department
Dutch Research Council (NWO)	Research funder
EATRIS – European Infrastructure for Translational Medicine	Research performing organisation
EcoHealth Alliance	Public health agency
EDP Sciences	Publisher
eLife	Publisher
Elsevier	Publisher
EMBO	Learned society
EMBO Press	Society publisher
Emerald Publishing	Publisher
Epicentre – MSF	Research performing organisation
European Commission	Research funder
European Institute for Systems Biology & Medicine (EISBM)	Research performing organisation
European Respiratory Society	Society publisher
European University Association (EUA)	University association
F1000 Research Limited	Publisher
Fondation Mérieux	Not-for-profit
Fonds de recherche du Québec, Canada	Research funder
Food & Drug Administration, USA	Public health agency
French National Research Agency (ANR)	Research funder
Frontiers	Publisher
Future Science Group (FSG)	Publisher
Genome Canada	Not-for-profit
Génome Québec	Not-for-profit
GeoVax	Biotech industry company
Global Alliance for Genomics and Health (GA4GH)	Non-governmental organisation
Global Virus Network	Non-governmental organisation
GloPID-R	Research funder
Good Clinical Practice Alliance – Europe	Non-governmental organisation
Health Research Board, Ireland	Research funder
Healthcare Infection Society	Public health agency
Hindawi	Publisher

Organisation	Stakeholder type
Indiana University	University
Indiana University – Purdue University Indianapolis (IUPUI)	University
Infectious Diseases Data Observatory (IDDO)	Research network
Inserm (Institut national de la santé et de la recherche médicale)	Research performing organisation
Institute of Tropical Medicine, Antwerp, Belgium	Research performing organisation
International Forum for Acute Care Trialists (InFACT)	Research network
International Severe Respiratory and emerging Infection Consortium (ISARIC)	Research network
International Society for Infectious Diseases (ISID)	Learned society
International Society on Thrombosis and Haemostasis (ISTH)	Learned society
Inter-university Consortium for Political and Social Research (ICPSR)	University association
IOP Publishing	Publisher
Jain Hospital & Research Center, India	Research performing organisation
Japan Agency for Medical Research and Development (AMED)	Research funder
JMIR Publications	Publisher
Johnson & Johnson	Pharmaceutical industry company
Kaohsiung Journal of Medical Sciences	Journal
Karger Publishers	Publisher
Kent Ridge Health Singapore	Healthcare service provider
Life Science Alliance	Journal
Luxembourg National Research Fund (FNR)	Research funder
MDPI	Publisher
Médecins Sans Frontières/Doctors Without Borders (MSF)	Non-governmental organisation
MedRxiv	Preprint server
Merck Research Laboratories	Pharmaceutical industry company
Michael Smith Foundation for Health Research	Research funder
Microbide Limited	Biotech industry company
Microbiology Society	Learned society
National Academy of Medicine	Learned society
National Institute for Health Research (NIHR)	Research funder
National Institute for Infectious Diseases Lazzaro Spallanzani, Italy	Public health agency
National Institutes of Health (NIH)	Research funder
National Research Foundation Singapore (NRF Singapore)	Government agency or department
National Science Centre, Poland	Research funder
Natural Sciences and Engineering Research Council of Canada (NSERC)	Research funder
New England Journal of Medicine (NEJM)	Journal
Office of Global Affairs, Department of Health and Human Services, USA	Government agency or department
Oxford University Press	University press
Partners in Digital Health	Publisher
PeerJ	Journal
Penn State University	University
PLOS	Publisher

Organisation	Stakeholder type
PNAS – Proceedings of the National Academy of Sciences of the USA	Journal
Portuguese Society of Mental Health Nursing	Journal
Project HOPE	Non-governmental organisation
ProMED	Research network
Research Square	Preprint server
Rockefeller University Press (RUP)	University press
SAGE Publishing	Publisher
Science Europe	Sector organisation
Science Foundation Ireland	Research funder
Science Journals – American Association for the Advancement of Science	Publisher
ScienceOpen	Service provider
Sciencepaper Online	Journal
SciLifeLab – Science for Life Laboratory, Sweden	Research performing organisation
Society for Applied Microbiology	Learned society
South African Medical Research Council	Research funder
SPARC Europe	Library association
Springer Nature	Publisher
SSRN	Preprint server
STM	Publisher association
Strategic Initiative for Developing Capacity in Ethical Review (SIDCER)	Non-governmental organisation
Swedish Research Council	Research funder
Takeda	Pharmaceutical industry company
Taylor & Francis	Publisher
The British Medical Journal (BMJ)	Journal
The Department for International Development (DFID)	Government agency or department
The Global Health Network	Non-governmental organisation
The Institut Pasteur	Research performing organisation
The JAMA Network	Publisher
The Lancet	Journal
The Research Council of Norway	Research funder
The Royal Society	Learned society
UK Research and Innovation (UKRI)	Research funder
UNIMED – Mediterranean Universities Union	University association
University College London	University
University of Bristol	University
University of Southampton	University
Vivli	Not-for-profit, Service provider
Volkswagen Foundation (VolkswagenStiftung)	Research funder
Wellcome Trust	Research funder
WikiJournal User Group	Publisher
Wiley	Publisher

Organisation	Stakeholder type
Wolters Kluwer	Publisher
World Scientific Publishing	Publisher
XPRIZE Foundation	Not-for-profit
ZonMW – The Netherlands Organisation for Health Research and Development	Research funder

The Ingenuity Centre, University of Nottingham Innovation Park, Nottingham, NG7 2TU, UK

www.research-consulting.com