

# BMJ Open Timely access to trial data in the context of a pandemic: the time is now

Rebecca Li <sup>1,2</sup>, Julie Wood,<sup>1</sup> Amrutha Baskaran,<sup>1</sup> Stanley Neumann,<sup>1</sup> Elizabeth Graham,<sup>1</sup> Marcia Levenstein,<sup>1</sup> Ida Sim<sup>3</sup>

**To cite:** Li R, Wood J, Baskaran A, *et al*. Timely access to trial data in the context of a pandemic: the time is now. *BMJ Open* 2020;**10**:e039326. doi:10.1136/bmjopen-2020-039326

► Prepublication history for this paper is available online. To view these files, please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2020-039326>).

Received 10 April 2020  
Revised 30 September 2020  
Accepted 12 October 2020



© Author(s) (or their employer(s)) 2020. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

<sup>1</sup>Vivli, Cambridge, Massachusetts, USA

<sup>2</sup>Center for Bioethics, Harvard Medical School, Boston, Massachusetts, USA

<sup>3</sup>Division of General Internal Medicine, University of California San Francisco, San Francisco, California, USA

## Correspondence to

Dr Rebecca Li; [rl@vivli.org](mailto:rl@vivli.org)

## ABSTRACT

**Objective** Clinical trial data sharing has the potential to accelerate scientific progress, answer new lines of scientific inquiry, support reproducibility and prevent redundancy. Vivli, a non-profit organisation, operates a global platform for sharing of individual participant-level trial data and associated documents. Sharing of these data collected from each trial participant enables combining of these data to drive new scientific insights or assess reproducibility—not possible with the aggregate or summary data tables historically made available. We report on our initial experience including key metrics, lessons learned and how we see our role in the data sharing ecosystem. We also describe how Vivli is addressing the needs of the COVID-19 challenge through a new dedicated portal that provides a direct search function for COVID-19 studies, availability for fast-tracked request review and data sharing.

**Data summary** The Vivli platform was established in 2018 and has partnered with 28 diverse members from industry, academic institutions, government platforms and non-profit foundations. Currently, 5400 trials representing 3.6 million participants are shared on the platform. From July 2018 to September 2020, Vivli received 201 requests. To date, 106 of 201 requests received approval, 5 have been declined, 27 withdrew and 27 are in the revision stage.

**Conclusions** The pandemic has only magnified the necessity for data sharing. If most data are shared and in a manner that allows interoperability, then we have hope of moving towards a cohesive scientific understanding more quickly not only for COVID-19 but also for all diseases. Conversely, if only isolated pockets of data are shared then society loses the opportunity to close vital gaps in our understanding of this rapidly evolving epidemic. This current challenge serves to highlight the value of data sharing platforms—critical enablers that help researchers build on prior knowledge.

## OBJECTIVE

In recent years, journal editors,<sup>1</sup> funders<sup>2</sup> and others<sup>3</sup> have promoted policies to encourage reuse and sharing of individual participant-level trial data (IPD) the vast majority of which is never made available. As the world is currently experiencing a rare event that has caused nearly all of science to pivot towards one singular shared goal—finding a treatment or vaccine for COVID-19—the necessity for

sharing data has never been greater. Unlike chronic diseases, for COVID-19 trials that are conducted in a condensed time period if the majority of data generators share their data rapidly, they may be integrated into the cohesive model of scientific understanding. Conversely, if only limited or isolated pockets of data are shared, or there are significant delays in the sharing of key COVID-19 datasets, then society loses the opportunity to close vital gaps in our medical understanding of this rapidly evolving disease challenge. To promote access and lower barriers to entry Vivli, the global data sharing platform was established in 2018<sup>4,5</sup> and our partners have waived all fees to share, archive, anonymise and access studies for COVID-19 during this time of urgent, global, medical need.

There are currently a number of clinical trial data sharing platforms that serve specific constituencies<sup>6–9</sup>; however, many of these remain disconnected from each other thereby creating a fragmented ecosystem. Launched in July 2018, the Vivli platform facilitates data sharing of clinical trials on any condition from any country, funder or investigator with a broader vision to link data from existing sharing communities and platforms (eg, NHLBI-NIH BIOLINCC, Project Datasphere, Yale-YODA partnership with Janssen and others). Vivli's mission is to enable and support researchers who commit in their attestation to ICMJE's Data Sharing Statement for Clinical Trials to sharing their participant-level data, as well as industry members who are committed to sharing completed clinical trials results and others whose funders mandate sharing.

Vivli currently provides access to more than 5400 trials representing 3.6 million participants contributed from a diverse group of 28 data contributors (represented by 17 industry, 9 academic institution/non-profit foundations and 2 government platforms and listed here <https://vivli.org/members/ourmembers/>) and has continued to grow rapidly



since our launch in 2018. The Vivli platform is operated by a non-profit organisation that acts as a neutral convener with transparent governance to balance the diverse interests of multiple stakeholders. In the ecosystem of data sharing platforms, there is a variation of implementation approaches with some platforms and repositories utilising an ‘open access’ or ‘open data’ approach—meaning that data are typically available via download on signing of a legal use agreement. Other platforms such as Vivli offer a ‘managed access’ or ‘gatekeeper’ model whereby additional requirements are needed for data access including submission of a research proposal and either an independent review panel (such as the Wellcome Trust who serves as the Vivli secretariat to evaluate requests) as well as other review committees make a decision on each research proposal based on transparent criteria. Those that contribute the data are required to render the data anonymised prior to submission. To tip the culture of data sharing and encourage the broadest participation within the clinical research community (industry, academia and non-profit foundations), Vivli supports a range of data access approaches from fully downloadable data to data that is highly restricted and accessible only in a secure environment. Early data contributors played a major role in Vivli’s momentum and continue to be involved in Vivli’s ongoing governance through a standing Steering Committee. We discuss the potential of data sharing platforms, its impact on researchers and also reflect on the lessons learnt on the journey towards a sustainable data sharing platform. We report here on Vivli’s first 2 years of experience. Building on that experience, we recently launched a COVID-19 portal in which COVID-19 trials are being fast-tracked onto the Vivli platform with dedicated search and availability for expedited request review.

### Since its launch, Vivli has

- ▶ Demonstrated a non-profit business model through which costs are sustainably covered.
- ▶ Created a collaborative group of diverse data contributors who participate in the governance and prioritisation of data sharing features.
- ▶ Established a harmonised data request and provisioning process with common agreements and the option for an independent review of requests.
- ▶ Enabled datasets to be combined and accessed from multiple sources in ways not previously possible, by connecting with existing platforms (such as the YODA Project,<sup>6</sup> ImmPort<sup>8</sup> and NHLBI’s BioLINCC)<sup>10</sup> and enabling the combination from diverse data sources (academic, large pharma, NIH and small biotech).

### Data access process

With this disease-agnostic platform, the scientific community can discover, access and analyse participant-level data from completed trials through an online request process with user support for both data contributors and data requesters. A request for data must include a proposal that includes the scientific hypothesis underlying the data

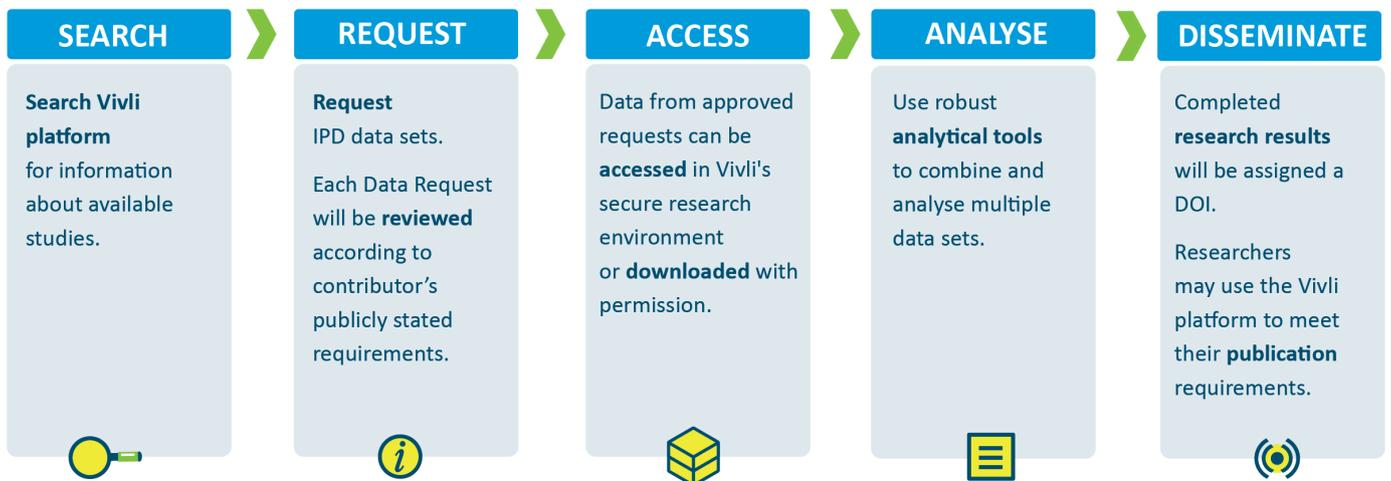
request as well as minimum requirements for statistical skills among the research team. A standard data use agreement provides access for 1 year and obligates researchers to disclose the results from their analyses publicly. Most requesters do not pay a fee for the first year after which a nominal fee is required to maintain continued access to the secure research environment. Data contributors choose whether their data are accessed via download or provisioned into a secure research environment. If access is provided in a secure research environment, the research team enters a user-friendly remote desktop area complete with a full suite of analytical tools (eg, SAS, STATA, R Studio and so on). Vivli data contributors declare transparent criteria for access to data, standard review criteria, any exceptions to these criteria and when their data would be available for sharing are available on the Vivli website. If a request is not approved, then the reason for non-approval is made public. [Figure 1](#) illustrates the Vivli request process.

### Data access metrics

At the time of this writing, there are 201 submitted requests (see [figure 2](#) depicting cumulative requests submitted over the last 2 years), with 80 requests (39.8%) that have had all data provided to the requester. Of the 201 total requests, 51 (25.4%) requested data from more than one data contributor whereas 150 (74.6%) requested data from one data contributor only. Requests range from 1 to 148 studies (median of 2 trials and a mean of 7 studies requested per proposal). Of the 28 Vivli data contributors, 19 have received at least one request for data. To date, 106 of the 201 requesting teams have received approval, 5 have been declined, 27 withdrew and 27 are in the revision stage. Eighty-four requests have completed the execution of the data use agreement. Data use agreements have been signed by requesters from 18 countries, a sign of the global interest in the data available on Vivli. The eighty fulfilled requests represent data from approximately 272 000 trial participants. Nine publications have resulted in the areas of oncology, dermatology and cardiology and a number are currently in progress.

Among the 5446 studies listed for sharing, 20.8% have been requested (1136 studies). Twenty-five of these studies are requested most frequently (five or more times). The breakdown of these 25 ‘frequently-requested studies’ is as follows: 12 rheumatoid arthritis studies; 4 ulcerative colitis/Crohn’s disease studies; 5 oncology studies; 1 COVID-19 study; 1 type 2 diabetes study and 1 cardiovascular study.

[Table 1](#) outlines the 18 subject areas of the submitted requests ranging from cardiovascular to vaccines, which indicates the breadth of interest in Vivli data. Methods and tool development requests as well as training proposals are combined under the ‘methods’ subject area. Most requests addressed either new questions surrounding treatment effectiveness or were designed to conduct participant-level meta-analysis rather than trying to confirm or reproduce prior analyses.



**Figure 1** Vivli Platform Request Process: illustrates the process for a researcher wishing to access IPD via Vivli. If a researcher wishes to submit a request, then they must complete a data request form. Data Requests may be reviewed by an Approving Entity. If approved, the Data Package consisting of the requested IPD and supporting material is provided and the requester signs a Data Use Agreement. The data are made available within a secure research environment or may.

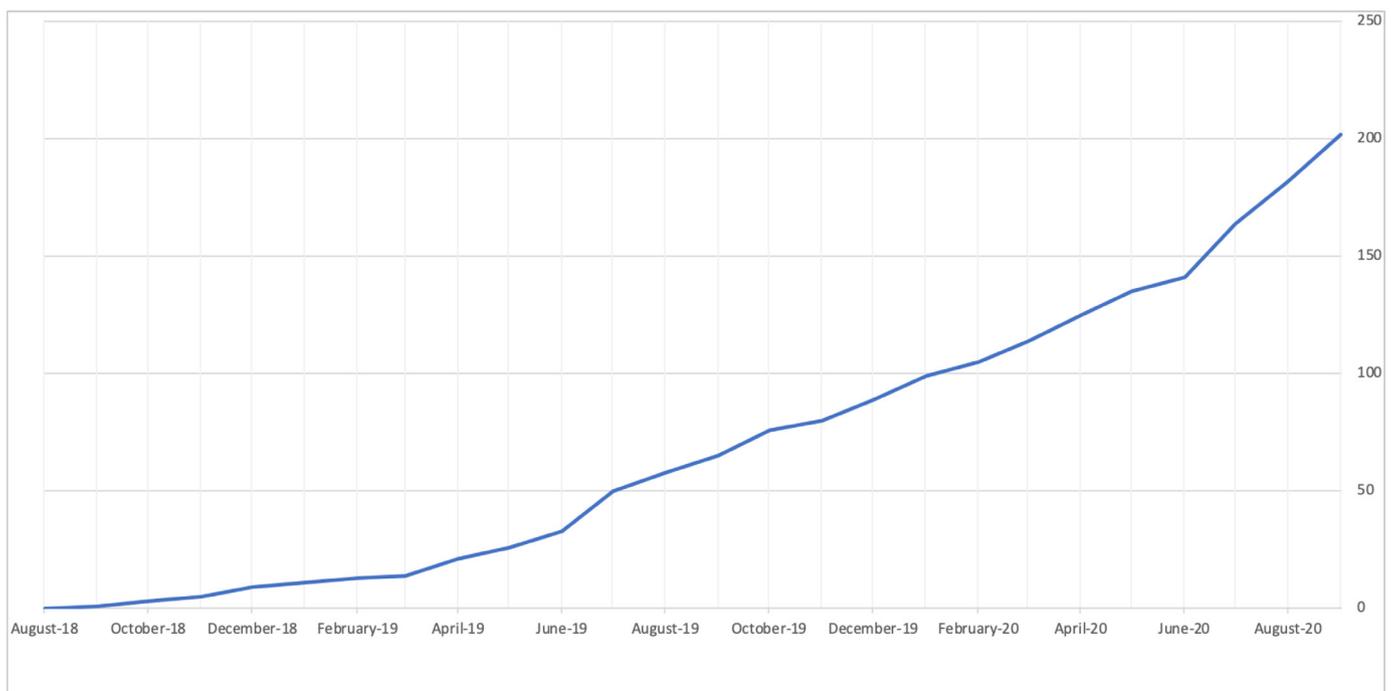
### Insights gained

The key challenges can be divided into those primarily affecting data contributors and those affecting requesters. Many data contributors face similar challenges including the sharing of imaging data and very large datasets, anonymisation of data (arduous and resource intense) and constraints surrounding the sharing of rare disease data. Data requesters have voiced challenges that the data formats they encounter are at times heterogeneous or non-standard; they wished that they had increased access to additional information during the request process (eg, protocols, data

dictionaries or clinical study reports) and the lengthy time from request to data access at times.

### Key insights we have gained include

- Understanding that as an independent non-profit organisation that manages a platform, we do not directly incentivise data sharing but we have a key role in enabling the implementation of incentives (eg, appropriate integration of digital object identifiers, governance of secondary datasets, training and provision of resources); it is the broader research ecosystem of funders, academia and publishers who



**Figure 2** Number of submitted requests from August 2018 to September 2020.

**Table 1** Proposals submitted by subject area

| Type               | Requests (n) |
|--------------------|--------------|
| Cardiovascular     | 25           |
| Dermatology        | 7            |
| Endocrinology      | 6            |
| Gastroenterology   | 7            |
| Gynaecology        | 1            |
| Haematology        | 2            |
| Immunology         | 2            |
| Infectious disease | 3            |
| Methods            | 11           |
| Nephrology         | 1            |
| Nephrology         | 22           |
| Oncology           | 48           |
| Orthopaedics       | 5            |
| Psychiatry         | 14           |
| Pulmonary          | 17           |
| Rheumatology       | 27           |
| Urology            | 1            |
| Vaccines           | 2            |
| Total              | 201          |

hold the crucial levers to reward investigators for sharing data for the public good. A future goal is to work with key funders synergistically to associate a researcher's role via their ORCID ID to a shared dataset thereby more directly linking data sharing with career progression.

- To facilitate data sharing, we realised that maximising flexibility for our data contributors in the area of data governance to the extent possible was important to lower the barriers for participation. For example contributors are able to maintain their own review processes (if a review is necessary to

access data). All data contributors publicly declare their criteria for sharing via a webpage on the Vivli platform. Harmonisation is required when necessary for the user experience. For instance, Vivli uses a single intake form for the scientific proposal and harmonised legal agreements have been essential to streamlining the overall process.

- Access to IPD requires more advocacy and we are just beginning to see an increase in demand for this type of data from a growing number of researchers. As more datasets accumulate in the Vivli repository enriching various disease areas, the ability to leverage 'big data insights' through combining multiple studies and pooling analyses to answer specific scientific questions steadily increases. Ultimately using IPD over summary-level data (or aggregate data) to conduct systematic reviews or meta-analyses allows for more powerful analyses such as new subgroup analyses and use of more diverse analytic methods but does require a greater resource commitment<sup>11</sup> Vivli is starting to see original publications based on IPD sharing emanate out of this work and a number are in the pipeline.

## CONCLUSION

Vivli depends on and fosters the collaboration of funders, sponsors, clinical trialists, secondary data researchers and, ultimately, trial participants in sustaining a functioning data sharing ecosystem (table 2). As commentators have noted, there continue to be challenges to clinical trial data sharing<sup>12</sup> but the combined experience of Vivli and other platforms<sup>4-10</sup> demonstrates that the first-order governance and technical challenges have been solved. The focus can now shift to achieving additional efficiencies in costs and processes, in facilitating best practices for data reuse and in tipping the academic clinical trial culture towards data sharing as a community norm. Vivli continues to partner with and bridge to other data sharing platforms as a collaborative

**Table 2** Vivli data sharing platform's role in the data sharing ecosystem

| Stakeholder                        | Stakeholder role  | Vivli's complementary role  |
|------------------------------------|---|---|
| Non-profit funder                  | Funds trials and supports grantees to make data available                         | Offers non-profit funders a mechanism for their grantees to share data beyond the grant period                          |
| Academic researcher                | Sharing and accessing data  | Provision of metrics and digital object identifiers to measure open science behaviours in future grant-making decisions |
| For-profit funder                  | Sponsor trials and makes data accessible to researchers                           | Offers for-profit sponsors a managed-access mechanism for sharing their trials  |
| Research participant               | Participates in trials  | Safeguards participant privacy while maximising the benefit of their contribution                                       |
| Clinical trialists and researchers | Conducts trials and/or conducts secondary research on clinical trial data         | Provides a mechanism for data access, long-term archiving and data reuse  |
| Journals and publishers            | Publishes findings from clinical trials and secondary analyses of clinical trials | Provides a controlled-access mechanism for journals and publishers that require clinical trial data to be shared        |

hub for major stakeholders to advance clinical trial data sharing and as an enabler of data sharing during a time of global public emergency. In conclusion, as COVID-19 is now a global pandemic with significant complexity and myriad unanswered questions, our goal is to encourage and support clinical researchers to join in this historic opportunity to share and facilitate access to clinical research data as rapidly as possible.

**Twitter** Marcia Levenstein @Mjl22g and Ida Sim @idasim

**Contributors** All authors conceived and gathered data for the manuscript and contributed to the tables and figures. RL summarised the data and wrote the manuscript. JW, AB, SN, EG, ML and IS provided comments on the manuscript and approved the final version of the manuscript for publication.

**Funding** The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

**Competing interests** Vivli receives funding support from the Arnold Ventures, Doris Duke Charitable Trust, the Helmsley Charitable Trust, pharmaceutical industry and academic institutional data contributors to Vivli.

**Patient consent for publication** Not required.

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

**Open access** This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

**ORCID iD**

Rebecca Li <http://orcid.org/0000-0002-2095-7550>

## REFERENCES

- 1 Taichman DB, Backus J, Baethge C, *et al*. Sharing clinical trial data: a proposal from the International Committee of medical Journal editors. *Ann Intern Med* 2016;164:505–6.
- 2 Kiley Ref *et al*. Data sharing from clinical trials—a research funder's perspective 2017:1990–2.
- 3 Institute of Medicine of the National Academies of Science. Sharing clinical trial data; maximizing benefits, minimizing risks. *National Academies Press* 2015.
- 4 Li R, Scott J, Rockhold F, *et al*. Moving data sharing forward: the launch of the Vivli platform. *NAM Perspectives* 2018;8.
- 5 Kaiser J. A new portal for patient data. *Science* 2018;361:212
- 6 Ross JS, Waldstreicher J, Bamford S, *et al*. Overview and experience of the yodA project with clinical trial data sharing after 5 years. *Sci Data* 2018;5:1–14.
- 7 Kochhar S, Knoppers B, Gamble C, *et al*. Clinical trial data sharing: here's the challenge. *BMJ Open* 2019;9:e032334.
- 8 Bhattacharya S, Andorf S, Gomes L, *et al*. ImmPort: disseminating data to the public for the future of immunology. *Immunol Res* 2014;58:234–9.
- 9 Humphreys GS, Tinto H, Barnes KI. Strength in numbers: the WWARN case study of Purpose-Driven data sharing. *Am J Trop Med Hyg* 2019;100:13–15.
- 10 Giffen CA, Wagner EL, Adams JT, *et al*. Providing researchers with online access to NHLBI biospecimen collections: the results of the first six years of the NHLBI BioLINCC program. *PLoS One* 2017;12:e0178141.
- 11 Tudur Smith C, Marcucci M, Nolan SJ, *et al*. Individual participant data meta-analyses compared with meta-analyses based on aggregate data. *Cochrane Database Syst Rev* 2016;9:MR000007.
- 12 Bauchner H, Fontanarosa PB. The challenges of sharing data in an era of Politicized science. *JAMA* 2019;322:2290–1.