



## REVIEW

# Open science approaches to COVID-19 [version 1; peer review: 2 approved]

Edwin G. Tse , Dana M. Klug , Matthew H. Todd 

Department of Pharmaceutical and Biological Chemistry, University College London School of Pharmacy, London, WC1N 1AX, UK

**V1** First published: 25 Aug 2020, 9:1043  
<https://doi.org/10.12688/f1000research.26084.1>  
 Latest published: 25 Aug 2020, 9:1043  
<https://doi.org/10.12688/f1000research.26084.1>

## Abstract

In only a matter of months, the coronavirus disease of 2019 (COVID-19) has spread around the world. The global impact of the disease has caused significant and repeated calls for quick action towards new medicines and vaccines. In response, researchers have adopted open science methods to begin to combat this disease *via* global collaborative efforts. We summarise here some of those initiatives, and have created an updateable list to which others may be added. Though open science has previously been shown as an accelerator of biomedical research, the COVID-19 crisis has made openness seem the logical choice. Will openness persist in the discovery of new medicines, after the crisis has receded?

## Keywords

Sars-CoV-2, COVID-19, open science, open data, open access, open source

## Open Peer Review

Reviewer Status  

Invited Reviewers

1

2

version 1



25 Aug 2020



report



report

1. **Matthew Hall** , National Institutes of Health, Rockville, USA  
**Kyle Brimacombe**, National Institutes of Health, Rockville, USA
2. **Alexander Tropsha** , University of North Carolina at Chapel Hill, Chapel Hill, USA

Any reports and responses or comments on the article can be found at the end of the article.

**Corresponding author:** Matthew H. Todd ([matthew.todd@ucl.ac.uk](mailto:matthew.todd@ucl.ac.uk))

**Author roles:** Tse EG: Conceptualization, Resources, Writing – Original Draft Preparation, Writing – Review & Editing; Klug DM: Conceptualization, Resources, Writing – Original Draft Preparation, Writing – Review & Editing; Todd MH: Conceptualization, Resources, Writing – Original Draft Preparation, Writing – Review & Editing

**Competing interests:** No competing interests were disclosed.

**Grant information:** This work was supported by the Artificial Intelligence and Augmented Intelligence for Automated Investigations for Scientific Discovery Network+, which has been funded by the Engineering and Physical Sciences Research Council (EPSRC) [EP/S000356/1].

*The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.*

**Copyright:** © 2020 Tse EG *et al.* This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

**How to cite this article:** Tse EG, Klug DM and Todd MH. **Open science approaches to COVID-19 [version 1; peer review: 2 approved]** F1000Research 2020, 9:1043 <https://doi.org/10.12688/f1000research.26084.1>

**First published:** 25 Aug 2020, 9:1043 <https://doi.org/10.12688/f1000research.26084.1>

## Introduction

In late 2019, reports began to emerge from Wuhan, China concerning cases of pneumonia of an unknown origin. Shortly thereafter, Chinese authorities identified this to be a novel type of coronavirus disease (now known as coronavirus disease 2019; COVID-19) caused by Severe Acute Respiratory Syndrome coronavirus 2 (SARS-CoV-2), with the outbreak being declared by the World Health Organization (WHO) as a public health emergency of international concern. Over the next two months, increasing numbers of COVID-19 cases were reported in countries outside China at an alarming rate, prompting the WHO to declare COVID-19 a global pandemic in March 2020<sup>1</sup>. To date, according to the WHO COVID-19 Situation Dashboard<sup>2</sup>, there have been around 18 million cases worldwide, in over 200 countries, resulting in around 690,000 deaths. As COVID-19 is caused by a novel coronavirus, there are no established methods for its treatment, and measures such as social distancing and self-isolation have become crucial to prevent further spread. The urgency to overcome this pandemic is furthered by the devastating effect on the world economy that has been seen as a result of the prolonged implementation of these measures<sup>3</sup>.

The enormous impact of this disease has resulted in significant activity towards new therapeutics, particularly for the development of a new vaccine. Much of this work is taking place in the private sector, alongside the usual requirement in that sector for secrecy. However, there has in parallel been a significant push for a more open approach because it is understood that openness leads to research acceleration. Broadly speaking, these initiatives can be grouped together as three types: 1) open access (the availability of research publications that are free to access and, often, re-use), 2) open data (the same, but with data) and 3) open source (in which broader community participation is allowed *via* liberal licence terms). There are many well-known and frequently-described advantages of openness (e.g. reduction of duplication of effort, faster communication of important outcomes) that nevertheless compete with a need for secrecy for many researchers, arising from the need for protected intellectual property or a perceived competitive advantage. These motivations for secrecy seem, in a time of crisis, to be lessened, and the increased prevalence of open initiatives relating to COVID-19 research has been striking.

Open science in biomedical research has gained increased traction over the past decade<sup>4,5</sup> from screening projects (e.g., CO-ADD<sup>6</sup>) and the sharing of physical samples (e.g., SGC probes<sup>7</sup>, MMV Boxes<sup>8</sup>) through to fully-fledged drug discovery (e.g., Open Source Malaria<sup>9</sup>, MycetOS<sup>10</sup>) and development (e.g., M4K Pharma<sup>11</sup>) campaigns. For COVID-19, data are being generated and shared (e.g. protein target structures, fragment hits), and initiatives have been created to identify and fast-track candidate compounds into clinical use. Even the lengthy process of drug approval is thought to be something that can be shortened: the urgency of the current situation and the use of open science has opened the possibility of reducing the timeline significantly to as little as 1.5 years<sup>12</sup> (though the

fastest vaccine to be developed, for the 2014–2016 Ebola virus epidemic, took 5 years)<sup>13,14</sup>.

This article collates the key open science resources and initiatives currently available for COVID-19 research (Figure 1). The three previously mentioned categories will be used to group the resources and a brief description of each will be given. This article forms the basis of a “living” collection of open science resources for COVID-19. As more resources become available, anyone may update [the repository](#) and discuss those additions.

## Open access

### Preprint servers

An obvious example of open access is the preprint server, such as bioRxiv<sup>15</sup> and medRxiv<sup>16</sup>. To date, there have been more than 7000 COVID-19 focussed articles submitted to these servers<sup>17</sup>. While this is a valuable resource, the articles are preliminary reports and are yet to be peer-reviewed: the main purpose of preprints is to allow researchers to quickly disseminate their results before official publication. It is noted that this surge in COVID-related studies has led to preprint servers like bioRxiv to implement stricter quality control of submitted articles, resulting in purely computational-based papers no longer being accepted<sup>18</sup>.

### Journals

Unlike preprint servers, open access journals publish peer-reviewed articles that are freely available for anyone to view without payment. Examples of open access journals featuring COVID-19 collections include the Public Library of Science journals (e.g. *PLoS One*, *PLoS Medicine*, etc.)<sup>19</sup>, *Nature Communications*<sup>20</sup> and Wellcome Open Research<sup>21</sup>. Other journals which are not fully open access, such as *JAMA* and *The BMJ*, have made their COVID collections free and open<sup>22,23</sup>.

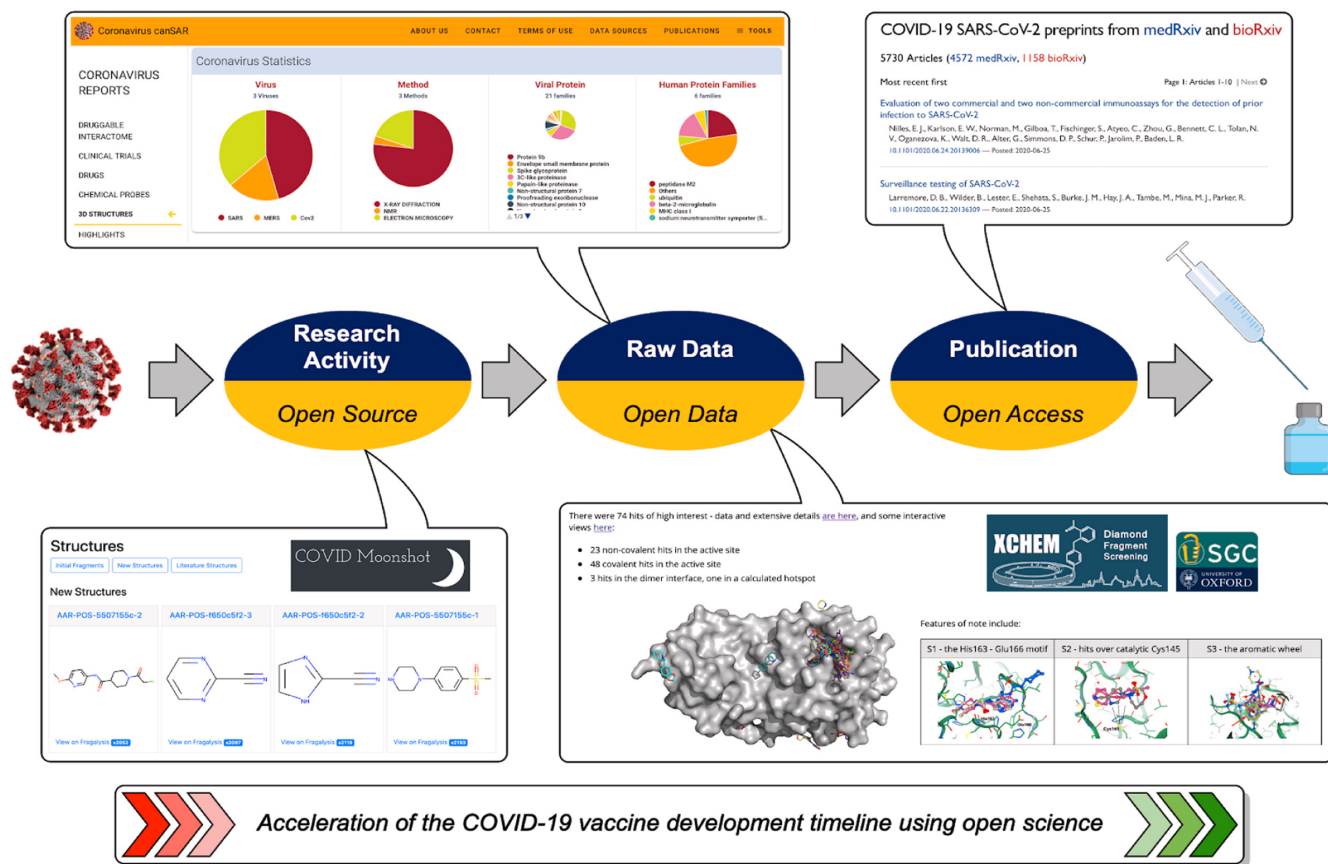
## Open data

### Diamond protease and fragment data

The Diamond Light Source is a synchrotron facility located in the UK that is utilised for a range of scientific areas including the investigation of protein structures and properties. During the coronavirus pandemic, Diamond scientists have used their facility to generate data on protein targets and fragments, all of which have been made publicly available<sup>24</sup>. Notably, by solving a high-resolution structure (1.39 Å) of the SARS-CoV-2 main protease<sup>25</sup>, it was possible to perform a screen of multiple fragment libraries to identify the most promising hits for fragment-based drug discovery. This screen resulted in 74 high interest hits and the full results of this screen have been made available to researchers and initiatives such as the PostEra COVID Moonshot project (*vide infra*). This main protease structure has also been used to screen for covalent probes from a library of electrophile fragments<sup>26</sup>.

### The Protein Data Bank

The Protein Data Bank (PDB) is an open database containing the 3D structural data of proteins and nucleic acids deposited by researchers from around the world. This database is an important resource for scientific research and many



**Figure 1. Application of open science for COVID-19 vaccine/treatment development.** Representative examples from the development timeline include the COVID Moonshot project, the ICR canSAR tool, the Diamond XChem fragment screen and preprint servers.

scientific journals now require authors to submit their structural data to the PDB. The PDB are maintaining a collection of a wide range of SARS-CoV-2 structures including the main protease and spike protein/receptors<sup>27</sup>.

### NCATS OpenData COVID-19

The National Center for Advancing Translational Sciences (NCATS) has focussed its efforts on drug repurposing for COVID-19 by generating datasets created from the screening of SARS-CoV-2-related assays against FDA-approved drugs and anti-infectious agents. Multiple compound collections are actively being screened in eight assays (with more in development) that focus on various stages of the SARS-Cov-2 life cycle in both human and viral targets. The results of the screen, as well as all assay protocols, have been made available online<sup>28</sup>.

### ICR Coronavirus canSAR

The Institute of Cancer Research (ICR) have developed a tool named canSAR which is a knowledgebase that collates multidisciplinary data and applies machine learning approaches to provide useful predictions for cancer drug discovery. The ICR have repurposed this tool for the current research efforts

against coronavirus, allowing people to freely search for information including the druggable interactome, ongoing and completed clinical trials, and lists of active compounds, probes and targets under investigation<sup>29</sup>.

### Data aggregation initiatives

A number of platforms have been created with the purpose of aggregating and curating openly available data that has been generated during COVID-19 research. The European Commission is working with a number of partners, including the EMBL-EBI, to create a platform that aggregates data ranging from sequencing and expression data to protein structures, drug targets and compounds<sup>30</sup>. The COVID-19 dataset is a large, machine-readable database intended to facilitate machine learning and data mining approaches to COVID-19 research<sup>31,32</sup>. The COVID-19 Molecular Structure and Therapeutics Hub maintains a repository of input files and analysis scripts for molecular simulation and dynamics studies related to COVID-19<sup>33</sup>.

### Open source

#### PostEra COVID Moonshot

PostEra AI is a for-profit startup company that specialises in integrating molecular design with chemical synthesis. As a

result of the coronavirus pandemic, and stemming from the data produced by the Diamond Light Source (*vide supra*), PostEra have collaborated with academic institutions and industry around the world, while adopting open science principles, to design new inhibitors of the SARS-CoV-2 main protease<sup>34</sup>. This global collaboration effort allows anyone to suggest new inhibitors based on the initial Diamond fragment hits. Following this, the most attractive compounds will be identified using machine learning algorithms, synthesised by a contract synthesis company and evaluated in inhibition assays (fluorescence and RapidFire mass spectrometry) against the SARS-CoV-2 main protease in labs around the world. Importantly, all stages of this process will be made publicly available. To date, there have been over 12000 unique compounds designed by the community. Over 1300 compounds have been ordered commercially, around 950 compounds have been synthesised, and over 900 compounds have been assayed against SARS-CoV-2.

### JEDI GrandChallenge

The Joint European Disruptive Initiative (JEDI) is a search for breakthrough technologies in the European Union. The GrandChallenge is a three stage campaign for the development of lead compounds against multiple SARS-CoV-2 targets<sup>35</sup>. Stage 1 is an open competition that focuses on *in silico* screening of compounds against high-resolution protein structures. Teams use simulation approaches (e.g. molecular dynamics, deep learning, docking, etc.) to score libraries of compounds against a chosen target. By comparing multiple different approaches errors can be averaged out and the best compounds for each protein target chosen for progression to the next stage. Stage 2 is an *in vitro* screening stage focussing on identifying the compounds from Stage 1 that provide 99% viral suppression. Teams must provide experimental evidence for this through either selective testing, high-throughput screening or smart combinatorial methods. Stage 3 is the *in vivo* screening stage aimed at finding novel drug combinations. This stage is run independently from the first two stages but lead compounds from those stages may be incorporated in this stage (provided they have been FDA-approved). Following the conclusion of each stage, the top-ranked team will be awarded a cash prize (€250,000 in both Stage 1 and 2, and up to €1,000,000 in Stage 3).

### Sample sharing: MMV COVID Box

The Medicines for Malaria Venture (MMV) is a not-for-profit organisation that brings together the public and private sectors for the discovery and development of new antimalarial medicines. MMV have previously created, and freely distributed, collections of promising candidate compounds in well plates to researchers around the world to enable a more efficient starting point for new drugs (see Malaria Box<sup>36</sup>, Pathogen Box<sup>37</sup> & Pandemic Box<sup>38</sup>). MMV have now created, and made available on request, the COVID Box, which contains 80 compounds of both marketed drugs and compounds in development that possess known or predicted activity against SARS-CoV-2<sup>39</sup>. A stipulation of this open research project is that the resulting data generated by researchers using the COVID Box must be shared in the public domain within 2 years of its generation.

### Sample sharing: COVID-19 Protein Portal

The COVID-19 Protein Portal is a UK-based initiative led by Wellcome and UKRI, that provides SARS-CoV-2-related protein reagents for UK scientists to use, free of charge<sup>40</sup>. These include viral proteins, human proteins and antibodies, all of which are searchable in their online database. All results generated from the use of these reagents will be made publicly available.

### Nextstrain pathogen evolution

Nextstrain is an open source, interactive data visualisation platform that provides “real-time tracking of pathogen evolution” *via* the analysis of sequencing data<sup>41</sup>. The tools used to achieve this are freely available to use and modify, and have already been used to track the evolution of a range of pathogens including the seasonal flu, Zika virus, and the West Nile virus. In response to the current pandemic, Nextstrain is maintaining a SARS-CoV-2 phylogenetic tree based on the analysis of contributed sequencing data.

### Folding@home

Folding@home (F@H) is a distributed computing project involving multiple research labs and citizen scientists from around the world that focuses on simulating protein dynamics. F@H provides software that enables users to donate unused computing power towards the computational analysis of protein folding<sup>42</sup>. Thus far, F@H’s COVID-19 projects have focused on simulating the interactions between the SARS-CoV-2 spike protein and the human ACE2 receptor to which it binds. All input files are available through GitHub, which is also where the research outputs will be made openly available<sup>43</sup>.

### Open Source COVID-19 Research Consortium

The Open Source COVID-19 (OSC19) Drug Discovery program is utilising computer science and biochemistry to enable the rapid screening of existing drug molecules for use against COVID-19<sup>44</sup>. Scientists from a range of fields are encouraged to participate, including synthetic chemists to make drug candidates, biochemists and virologists to run assays and donors and volunteers to aid in publicity and fundraising. As an open source project, all research results will be freely available with no intellectual property claims for any discoveries made.

### Other resources

In a similar manner to this article, a number of additional resources have been created to help spread word of the growing list of open science efforts for COVID-19 research. Examples include Initiatives such as Joinup EU<sup>45</sup> and SPARC<sup>46</sup>, which have created hubs of open source research projects and resources. Funding agencies like the UKRI and UKCDR have provided lists of funded COVID-19 research projects to help researchers identify and fill funding gaps<sup>47,48</sup>. To further advocate the use of open science, the Open COVID Pledge has aimed to encourage researchers and businesses to make their COVID-related intellectual property freely available by providing the Open COVID Licence<sup>49</sup>. Finally, the Virus Outbreak Data Network (VODAN) is working to ensure that data related to the COVID outbreak is findable, accessible, interoperable, and reusable<sup>50</sup>.

## Conclusion

A significant amount of effort has been made to progress COVID-19 research since the beginning of the pandemic. With so many scientific minds working on this problem together, it is clear that conducting scientific research in an open manner can accelerate the research process. The resources and initiatives highlighted in this article demonstrate the benefits of open science approaches and its potential to accelerate research timelines. This article itself will form the basis of an open science “living” resource hosted on a [public repository](#). It is understandable that a global public health crisis causes us to

adopt innovations in how we work in the search for an effective solution. After the COVID-19 crisis has faded there will remain many other crises that we face in the search for an effective therapy to alleviate suffering, whether the affected population is a billion people or a single individual in search of a cure. It is hoped that open science will be seen as a “new normal” approach in those crises too.

## Data availability

### Underlying data

No data are associated with this article.

## References

1. WHO Director-General's Open Remarks at the Media Briefing on COVID-19 – 11 March 2020. 2020; Accessed 9 June, 2020.  
[Reference Source](#)
2. WHO Coronavirus Disease (COVID-19) Dashboard. 2020; Accessed 9 June, 2020.  
[Reference Source](#)
3. Gopinath G: **The Great Lockdown: Worst Economic Downturn Since the Great Depression**. 2020; Accessed 24 June, 2020.  
[Reference Source](#)
4. McKiernan EC, Bourne PE, Brown CT, *et al.*: **How Open Science Helps Researchers Succeed**. *eLife*. 2016; 5: e16800.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
5. Shaw DL: **Is Open Science the Future of Drug Development?** *Yale J Biol Med*. 2017; 90(1): 147–151.  
[PubMed Abstract](#) | [Free Full Text](#)
6. Zuegg J, Hansford KA, Elliott AG, *et al.*: **How to Stimulate and Facilitate Early Stage Antibiotic Discovery**. *ACS Infect Dis*. 2020; 6(6): 1302–1304.  
[PubMed Abstract](#) | [Publisher Full Text](#)
7. Müller S, Ackloo S, Arrowsmith CH, *et al.*: **Donated Chemical Probes for Open Science**. *eLife*. 2018; 7: e34311.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
8. Van Voorhis WC, Adams JH, Adelfio R, *et al.*: **Open Source Drug Discovery with the Malaria Box Compound Collection for Neglected Diseases and Beyond**. *PLoS Pathog*. 2016; 12(7): e1005763.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
9. Williamson AE, Ylloja PM, Robertson MN, *et al.*: **Open Source Drug Discovery: Highly Potent Antimalarial Compounds Derived from the Tres Cantos Arylpyrroles**. *ACS Cent Sci*. 2016; 2(10): 687–701.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
10. Lim W, Melse Y, Konings M, *et al.*: **Addressing the Most Neglected Diseases Through an Open Research Model: The Discovery of Fenarimols as Novel Drug Candidates for Eumycetoma**. *PLoS Negl Trop Dis*. 2018; 12(4): e0006437.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
11. Morgan MR, Roberts OG, Edwards AM: **Ideation and Implementation of an Open Science Drug Discovery Business Model – M4K Pharma [version 1; peer review: 2 approved, 1 approved with reservations]**. *Wellcome Open Res*. 2018; 3: 154.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
12. Lurie N, Saville M, Hatchett R, *et al.*: **Developing Covid-19 Vaccines at Pandemic Speed**. *N Engl J Med*. 2020; 382(21): 1969–1973.  
[PubMed Abstract](#) | [Publisher Full Text](#)
13. WHO Prequalifies Ebola Vaccine, Paving the Way for its use in High-Risk Countries. 2019; Accessed 9 June, 2020.  
[Reference Source](#)
14. Herder M, Graham JE, Gold R: **From Discovery to Delivery: Public Sector Development of the rVSV-ZEBOV Ebola Vaccine**. *J Law Biosci*. 2020; Isz019.  
[Publisher Full Text](#)
15. **bioRxiv - The Preprint Server for Biology**. 2020; Accessed 8 June, 2020.  
[Reference Source](#)
16. **medRxiv - The Preprint Server for Health Sciences**. 2020; Accessed 8 June, 2020.  
[Reference Source](#)
17. **COVID-19 SARS-CoV-2 Preprints from medRxiv and bioRxiv**. 2020; Accessed 8 June, 2020.  
[Reference Source](#)
18. Kwon D: **How Swamped Preprint Servers are Blocking Bad Coronavirus Research**. *Nature*. 2020; 581(7807): 130–131.  
[PubMed Abstract](#) | [Publisher Full Text](#)
19. **PLoS COVID-19 Updates**. 2020; Accessed 23 June, 2020.  
[Reference Source](#)
20. **SARS-CoV-2 - Latest Research and News**. 2020; Accessed 23 June, 2020.  
[Reference Source](#)
21. **Articles from collection Coronavirus (COVID-19)**. 2020; Accessed 23 June, 2020.  
[Reference Source](#)
22. **Coronavirus (COVID19)**. 2020; Accessed 29 June, 2020.  
[Reference Source](#)
23. **Coronavirus (covid-19): Latest News and Resources**. 2020; Accessed 29 June, 2020.  
[Reference Source](#)
24. **Main Protease Structure and XChem Fragment Screen**. 2020; Accessed 8 June, 2020.  
[Reference Source](#)
25. Owen CD, Lukacik P, Strain-Damerell CM, *et al.*: **SARS-CoV-2 main protease with unliganded active site (2019-nCoV, coronavirus disease 2019, COVID-19)**. *wwwPDB*. 2020.  
[Publisher Full Text](#)
26. Resnick E, Bradley A, Gan J, *et al.*: **Rapid Covalent-Probe Discovery by Electrophile-Fragment Screening**. *J Am Chem Soc*. 2019; 141(22): 8951–8968.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
27. **COVID-19/SARS-CoV-2 Resources**. 2020; Accessed 30 June, 2020.  
[Reference Source](#)
28. **OpenData | COVID-19**. 2020; Accessed 8 June, 2020.  
[Reference Source](#)
29. **canSAR Coronavirus Research Tool**. 2020; Accessed 16 June, 2020.  
[Reference Source](#)
30. **The COVID-19 Data Portal**. 2020; Accessed 23 June, 2020.  
[Reference Source](#)
31. **COVID-19 Open Research Dataset Challenge (CORD-19)**. 2020; Accessed 24 June, 2020.  
[Reference Source](#)
32. **CORD-19 COVID-19 Open Research Dataset**. 2020; Accessed 24 June, 2020.  
[Reference Source](#)
33. **COVID-19 Molecular Structure and Therapeutics Hub**. 2020; Accessed 24 June, 2020.  
[Reference Source](#)
34. **COVID Moonshot**. 2020; Accessed 8 June, 2020.  
[Reference Source](#)
35. **JEDI Grand Challenge**. 2020; Accessed 8 June, 2020.  
[Reference Source](#)
36. **The Malaria Box**. 2020; Accessed 16 June, 2020.  
[Reference Source](#)
37. **The Pathogen Box**. 2020; Accessed 16 June, 2020.  
[Reference Source](#)
38. **The Pandemic Response Box**. 2020; Accessed 16 June, 2020.  
[Reference Source](#)
39. **The COVID Box**. 2020; Accessed 16 June, 2020.  
[Reference Source](#)

40. **COVID-19 Protein Portal**. 2020; Accessed 23 July, 2020.  
[Reference Source](#)
41. **Real-Time Tracking of Pathogen Evolution**. 2020; Accessed 24 June, 2020.  
[Reference Source](#)
42. **COVID-19**. 2020; Accessed 24 June, 2020.  
[Reference Source](#)
43. **Folding@home COVID-19 Efforts**. 2020; Accessed 24 June, 2020.  
[Reference Source](#)
44. **Open Source COVID-19 Research Consortium**. 2020; Accessed 16 July, 2020.  
[Reference Source](#)
45. **Digital Response to COVID-19 Open Source Solutions**. 2020; Accessed 24 June, 2020.  
[Reference Source](#)
46. **The Coronavirus and Open Science: Our Reads and Open Use Cases**. 2020; Accessed 24 June, 2020.  
[Reference Source](#)
47. **COVID-19 Research and Innovation Supported by UKRI**. 2020; Accessed 24 June, 2020.  
[Reference Source](#)
48. **COVID-19 Research Project Tracker by UKCDR & GloPID-R**. 2020; Accessed 24 June, 2020.  
[Reference Source](#)
49. **Open COVID Pledge**. 2020; Accessed 24 June, 2020.  
[Reference Source](#)
50. **Virus Outbreak Data Network (VODAN)**. 2020; Accessed 24 June, 2020.  
[Reference Source](#)

# Open Peer Review

Current Peer Review Status:  

---

## Version 1

Reviewer Report 26 October 2020

<https://doi.org/10.5256/f1000research.28785.r70312>

© 2020 Tropsha A. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



**Alexander Tropsha** 

Laboratory for Molecular Modeling; UNC Eshelman School of Pharmacy, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

This is a very timely, excellent manuscript that both highlights and supports open drug discovery research especially for COVID-19 and possible future pandemics. The authors provide a nice overview of different types of open research activities related to the research response to COVID-19.

There has been indeed an unprecedented level of sharing data and resources since the beginning of the pandemic and it is difficult to cover all of these developments in a short review. I would nevertheless suggest a few additions. Thus, I recommend adding a statement on the availability of computational resources besides Folding@home such as COVID-19 HPC consortium (<https://www.xsede.org/covid19-hpc-consortium>).

I would also suggest a broader acknowledgment of multiple journals and publishers that agreed to provide open access to all papers on the topic of COVID-19. A possibly complete list is available from the Wellcome Trust (<https://wellcome.org/press-release/publishers-make-coronavirus-covid-19-content-freely-available-and-reusable>).

Finally, I would recommend that the authors expand on the important issue of the quality of research publications. The authors mention that even preprint servers stopped accepting manuscripts with countless purely computationally repurposed drugs against the disease with no experimental confirmation. It would be prudent to cite a recent publication in Nature Biotechnology on the need for automated peer review of the publications tsunami (<https://www.nature.com/articles/s41587-020-0686-x>)<sup>1</sup> and perhaps, emphasize the importance of rigor, experimental support, and reproducibility of research findings reported in both preprint servers and peer-reviewed publications.

## References

1. Levin JM, Oprea TI, Davidovich S, Clozel T, et al.: Artificial intelligence, drug repurposing and

peer review. *Nat Biotechnol.* 2020; **38** (10): 1127-1131 [PubMed Abstract](#) | [Publisher Full Text](#)

**Is the topic of the review discussed comprehensively in the context of the current literature?**

Yes

**Are all factual statements correct and adequately supported by citations?**

Yes

**Is the review written in accessible language?**

Yes

**Are the conclusions drawn appropriate in the context of the current research literature?**

Yes

**Competing Interests:** No competing interests were disclosed.

**Reviewer Expertise:** Cheminformatics and Computer Aided Drug Discovery

**I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.**

Reviewer Report 14 September 2020

<https://doi.org/10.5256/f1000research.28785.r70313>

© 2020 Hall M et al. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The author(s) is/are employees of the US Government and therefore domestic copyright protection in USA does not apply to this work. The work may be protected under the copyright laws of other jurisdictions when used in those jurisdictions.



**Matthew Hall** 

National Center for Advancing Translational Sciences, National Institutes of Health, Rockville, MD, USA

**Kyle Brimacombe**

National Center for Advancing Translational Sciences, National Institutes of Health, Rockville, MD, USA

### Comments to the Author

The authors have worked to summarize the open science efforts made in the wake of the COVID-19 pandemic. They have collected COVID-19-related open research resources and categorized them as open access, open data, or open source. In summarizing these approaches, the authors seek to highlight the benefits of open science and the promise of accelerated therapeutic development in this unprecedented COVID era.



The review as a whole is succinct, well-written and easy-to-understand, and presents a clear summary of the accumulated open science resources. The curation work the authors have undertaken is commendable, and given the number of COVID-related initiatives becoming public, it could prove valuable to researchers looking to find an appropriate resource.

It's clear the review was meant to summarize the state of open science initiatives for COVID-19 research without offering commentary or analysis on the approaches taken, but this reviewer can't help but wonder what lessons the scientific community will learn from this first big foray (for many) into open science. E.g. who were the main users of these open science resources? Are there (or should there be) minimum requirements to be considered an "open science" resource? Is it enough to say that data/results will be made public in the future?

Regarding the list being compiled by the authors, this reviewer understands why Github is a logical home for a living, open source resource, but it may inadvertently limit the audience reached (or require a bit of exploration for users new to the platform). Github will likely be familiar to those in the informatics or developer communities, but to those in basic/bench research, the sight of a Github page may signal that the resource is not meant for them. The authors might consider mirroring the resource somewhere in a format more familiar to bench researchers or perhaps work to direct users to the wiki rather than the code tab, where the sight of a classic README.md file/structure might scare away potential readers who otherwise recognize that as the sign of a wrong turn taken.

To improve the utility of their lists, the authors might also consider categorizing the resources by subject matter (e.g. structural data vs small molecule screening/activity vs prediction platforms) in addition to classifying by type of open science resource. It's likely that users may be in search of a particular type of COVID resource/dataset matched to their area of expertise, more so than a particular class of open science platform (open data vs open source). Those in search of the former will need to read through a handful of separate resource write-ups before being able to decide which, if any, are relevant to them.

The value of this type of resource is highly dependent on how frequently it is updated and curated, so the authors should be commended on their undertaking of this work.

#### **Minor edit to text:**

The NCATS OpenData Portal (cited as #28) also has a preprint which goes into greater detail about the mission and design of the site: <https://www.biorxiv.org/content/10.1101/2020.06.04.135046v1>  
1

#### **References**

1. Brimacombe KR, Zhao T, Eastman RT, Hu X, et al.: An OpenData portal to share COVID-19 drug repurposing data in real time. *bioRxiv*. 2020. [PubMed Abstract](#) | [Publisher Full Text](#)

#### **Is the topic of the review discussed comprehensively in the context of the current literature?**

Yes

#### **Are all factual statements correct and adequately supported by citations?**

Yes

**Is the review written in accessible language?**

Yes

**Are the conclusions drawn appropriate in the context of the current research literature?**

Yes

**Competing Interests:** No competing interests were disclosed.

**Reviewer Expertise:** Open data sharing, Translational science, Assay development, Experimental therapeutics, Drug discovery

**We confirm that we have read this submission and believe that we have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.**

---

The benefits of publishing with F1000Research:

- Your article is published within days, with no editorial bias
- You can publish traditional articles, null/negative results, case reports, data notes and more
- The peer review process is transparent and collaborative
- Your article is indexed in PubMed after passing peer review
- Dedicated customer support at every stage

For pre-submission enquiries, contact [research@f1000.com](mailto:research@f1000.com)

**F1000Research**