



OPINION ARTICLE

**REVISED** Open laboratory notebooks: good for science, good for society, good for scientists [version 2; peer review: 2 approved, 1 approved with reservations]

Matthieu Schapira <sup>1,2</sup>, The Open Lab Notebook Consortium, Rachel J. Harding <sup>1</sup>

<sup>1</sup>Structural Genomics Consortium, University of Toronto, Toronto, ON, M5G 1L7, Canada

<sup>2</sup>Department of Pharmacology and Toxicology, University of Toronto, Toronto, ON, M5G 1L7, Canada

**v2** First published: 22 Jan 2019, 8:87 (<https://doi.org/10.12688/f1000research.17710.1>)  
 Latest published: 02 Apr 2019, 8:87 (<https://doi.org/10.12688/f1000research.17710.2>)

**Abstract**

The fundamental goal of the growing open science movement is to increase the efficiency of the global scientific community and accelerate progress and discoveries for the common good. Central to this principle is the rapid disclosure of research outputs in open-access peer-reviewed journals and on pre-print servers. The next bold step in this direction is open laboratory notebooks, where research scientists share their research — including detailed protocols, negative and positive results — online and in near-real-time to synergize with their peers. Here, we highlight the benefits of open lab notebooks to science, society and scientists, and discuss the challenges that this nascent movement is facing. We also present the implementation and progress of our own initiative at [openlabnotebooks.org](http://openlabnotebooks.org), with more than 20 active contributors after one year of operation.

**Keywords**

open lab notebooks, open science, peer-review, preprints, publishing, science communication



This article is included in the **Science Policy Research gateway**.

**Open Peer Review**

Reviewer Status ? ✓ ✓

	Invited Reviewers		
	1	2	3
<b>REVISED</b>		✓	✓
<b>version 2</b> published 02 Apr 2019		report	report
		↑	
<b>version 1</b> published 22 Jan 2019	? report	? report	

- Matthew H. Todd** , University College London (UCL), London, UK  
**Edwin Tse** , University of Sydney, Sydney, Australia  
**Marat Korsik**, University of Sydney, Sydney, Australia  
**Mathamsanqa Bhebhe**, University of Sydney, Sydney, Australia
- Lara M. Mangravite** , Sage Bionetworks, Seattle, USA
- Jean-Baptiste Poline** , McGill University, Montreal, Canada

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

**Corresponding authors:** Matthieu Schapira ([matthieu.schapira@utoronto.ca](mailto:matthieu.schapira@utoronto.ca)), Rachel J. Harding ([rachel.harding@utoronto.ca](mailto:rachel.harding@utoronto.ca))

**Author roles:** Schapira M: Supervision, Writing – Original Draft Preparation; Harding RJ: Conceptualization, Investigation, Writing – Review & Editing

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

**Grant information:** The Structural Genomics Consortium (SGC) is a registered charity (number 1097737) that receives funds from AbbVie, Bayer Pharma AG, Boehringer Ingelheim, Canada Foundation for Innovation, Eshelman Institute for Innovation, Genome Canada through Ontario Genomics Institute [OGI-055], Innovative Medicines Initiative (EU/EFPIA) [ULTRA-DD grant no. 115766], Janssen, Merck KGaA, Darmstadt, Germany, MSD, Novartis Pharma AG, Ontario Ministry of Research, Innovation and Science (MRIS), Pfizer, São Paulo Research Foundation-FAPESP, Takeda, and The Wellcome Trust [106169/ZZ14/Z]. RJH is a recipient of the HDSA Berman/Topper HD Career Development Fellowship. PR is funded by the Wellcome Trust Leicester ISSF award reference 204801/Z/16/Z and the Leicester Institute of Chemical and Structural Biology (LISCB).

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

**Copyright:** © 2019 Schapira M *et al.* This is an open access article distributed under the terms of the [Creative Commons Attribution Licence](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

**How to cite this article:** Schapira M, The Open Lab Notebook Consortium and Harding RJ. **Open laboratory notebooks: good for science, good for society, good for scientists [version 2; peer review: 2 approved, 1 approved with reservations]** F1000Research 2019, 8:87 (<https://doi.org/10.12688/f1000research.17710.2>)

**First published:** 22 Jan 2019, 8:87 (<https://doi.org/10.12688/f1000research.17710.1>)

**REVISED Amendments from Version 1**

The revised version includes discussions on data ownership and the permanence of open lab notebooks. In particular, we have now implemented a robust archival system in GitHub and <https://archive.org> for data entered at <https://openlabnotebooks.org>. As suggested, we also provide a brief history of the open lab notebook community. Finally, we discuss the risk that dubious experiments find their way on open platforms, get amplified over the internet and mislead scientists, patient groups or other communities.

**See referee reports**

## Introduction

The function of the scientific peer-reviewed system is to provide greater confidence that published research is scientifically sound. This system is widely accepted as the best available, although imperfect (as peer reviewers may miss technical flaws or be biased<sup>1</sup>), to guide the global scientific community towards progress. Peer-reviewed publishing is also used by research scientists, funders and institutions as a mechanism to claim ownership of their discoveries. As a result, the community widely believes that findings should be kept secret until they are published in a peer-reviewed journal. This tradition of secrecy, which protects the scientist as opposed to the science, has been transmitted from mentor to trainee for centuries (Galileo kept his discoveries to himself until they were published). In the life sciences, this belief can reach near-mystical levels<sup>2</sup>, and can be compounded by constraints associated with patent protection procedures or the absence of clear mechanism to make one's data publicly available. The peer-review and publication process grew in an era where communication was largely in paper format. Today, in the age of instant communication, one would imagine there should be more efficient ways to operate.

## Open lab notebooks: good for science and society

We believe that open laboratory notebooks, where research scientists record their work online and in near-real time, are an efficient way to disseminate data before it is published in peer-reviewed journals, and has several advantages over the traditional "release after publication" system<sup>3</sup>. First, making the data accessible within weeks rather than keeping it hidden for years means that others will be able to build upon the research, and avoid spending time and resources on redundant experiments<sup>4</sup>. Second, open lab notebooks should include detailed protocols that can be reproduced, which is often not the case in peer-reviewed publications<sup>5,6</sup>. Third, negative data, which are almost never disclosed in the current publishing system but are provided in open lab notebooks, can sometime provide important insight<sup>7,8</sup>. Fourth, open lab notebooks offer a space for anyone to comment on experimental records. This allows experts to provide insight, but also to flag technically unsound experiments, thereby reducing the potential for flawed science to appear in peer-reviewed journals and in pre-print media. Open lab notebooks can therefore help save time, resources, and knowledge. If adopted by many, they should lead to a more synergistic way to do science and to more efficient use of public funds.

## Good for scientists

Many believe that the chances of getting scooped before one publishes their work in a peer-reviewed journal increase when openly sharing their work online<sup>9</sup>. We argue that open lab notebooks have compensating advantages that are good for scientists. To succeed in academia, one must get funding, assert primacy over discoveries, be known in a field of research and be able to present work and ideas clearly and convincingly. Open lab notebooks can help in all aspects.

First, funding agencies are seeing the open science movement as a long lasting and far-reaching shift for the best, and are increasingly supportive of efforts to embrace open science principles. For instance, the symposium set to launch [openlabnotebooks.org](https://openlabnotebooks.org) was entirely sponsored by the Wellcome Trust and the Canadian Institute of Health Research, and senior representatives from the Gates Foundation and the Chan-Zuckerberg Initiative were also in attendance (<https://www.thesgc.org/openlab-notebooks-2018>). The NIH's National Institute on Aging dedicated an entire session to open science at their 2018 Alzheimer's research summit (<https://www.nia.nih.gov/research/nih-ad-summit-2018-program-agenda>), as did the 2018 Enroll-HD congress of the CHDI Huntington's Disease Foundation (<https://www.enroll-hd.org/enroll-hd-congress-2018/>). The Wellcome Trust has recently launched the [Wellcome Open Research publishing platform](#) and [Open Research Fund](#). Our personal observations seem to indicate that grant applications highlighting the use of open lab notebooks are being viewed positively. For example, Huntington's disease (HD) research funders such as the CHDI Foundation, the Huntington Society of Canada and the Huntington Society of America, have all generously funded studies of HD biochemistry at the SGC Toronto.

Second, results in open lab notebook are date-stamped, thus claiming temporal priority of the data. Indeed, public repositories such as Zenodo<sup>10</sup> add a date-stamp to depositions, and assign a citable DOI to open lab notebook records (detailed below): once a record has been published, it can no longer be modified, but revised versions can be appended if necessary.

Third, early career scientists can use their open notebooks to connect with their peers and with experts in the field, start new collaborations and build their own network. Fourth, the use of open lab notebooks provides opportunity to present work clearly and concisely to both experts and non-experts. This is an important skill to master in order to write convincing grant applications. Fifth, junior scientists will also find their open lab notebook a good medium to showcase their technical skills and scientific insight, and may find it useful to add a link in their resume when applying for their next position. Finally, many will find a personal satisfaction in embracing open science and FAIR data principles<sup>11</sup>.

## Implementation of an open lab notebook platform

Open lab notebooks have been pioneered and championed by a number of practitioners but remain a niche activity in the scientific community. Jean-Claude Bradley first coined the term "open notebook science" in 2006 and his definition of this method of scholarly communication have laid the foundations

for our own efforts<sup>12</sup>. In addition to the notebooks of individual researchers following Bradley's template, open notebook examples now include collective efforts from the Open Lab Notebook Network (<http://onsnetwork.org>) and Open Source Malaria (<http://opensourcemalaria.org>). However, the open lab notebook community remains small, the practice is not consistently defined or implemented and the impact of these efforts in the field have not been systematically evaluated.

Following our prediction that open lab notebooks should be good for science and good for scientists, and after a 2-year pilot study where Rachel Harding, a post-doctoral fellow at the Structural

Genomics Consortium (SGC) shared her work on Huntington's disease at [labscribbles.com](http://labscribbles.com)<sup>13</sup>, we launched [openlabnotebooks.org](http://openlabnotebooks.org) in January 2018, where 12 scientists from the SGC started reporting their work live, online<sup>14,15</sup>. Each post is composed of two documents. (1) A detailed and rigorous experimental record, including all data and protocols, which experts can evaluate, comment on or build upon (Figure 1); (2) a blog, aimed at the non-specialist that explains in simple terms the motivation and rationale for the experiment, summarizes results – positive and negative – and outlines next steps (Figure 2). The blogs, posted at [openlabnotebooks.org](http://openlabnotebooks.org), are managed by a webserver downloaded from [wordpress.org](http://wordpress.org) (the open-source online system LabTrove

## Viability screen of ACVR1/ALK2 inhibitory compounds on Diffuse Intrinsic Pontine Glioma cells

 Brown, Elizabeth;  Farnie, Gillian;  Bullock, Alex

6 compounds developed by M4K Pharma to inhibit the ACVR1/ALK2 tyrosine kinase were tested for their ability to reduce viability of both mutant and wild-type ACVR1 DIPG cell lines (HSJD-DIPG-007 and HSJD-DIPG-011 respectively). Compounds were tested on cells grown as 2D sheets and 3D neurospheres, as well as with and without radiation treatment. M4K2009 was highly effective in ACVR1 mutant DIPG cells, but this was not the case in ACVR1 wild-type cells.

**Funding Acknowledgment:** The SGC is a registered charity (number 1097737) that receives funds from AbbVie, Bayer Pharma AG, Boehringer Ingelheim, Canada Foundation for Innovation, Eshelman Institute for Innovation, Genome Canada through Ontario Genomics Institute [OGI-055], Innovative Medicines Initiative (EU/EFPIA) [ULTRA-DD grant no. 115766], Janssen, Merck KGaA, Darmstadt, Germany, MSD, Novartis Pharma AG, Ontario Ministry of Research, Innovation and Science (MRIS), Pfizer, São Paulo Research Foundation-FAPESP, Takeda, and Wellcome.

Preview



Page: 2 of 7 — Automatic Zoom

**Day 7**  
Scan cells with image cytometer to record final confluence/sphere diameter and propidium iodide intensity. Harvest cells with CellTiter-Glo® 3D (for all cells and all conditions) as a measure of culture viability according to [manufacturer's instructions](#).

**Data analysis**  
All values were normalised to the average of the vehicle controls on the plate (or the plate divided into sections using the vehicle controls, if visible variation across the plate).

**Results**  
In the mutant ACVR1 cell line HSJD-DIPG-007 the most effective probe is M4K2009 (with the caveat that this dataset only includes one biological repeat), with all M4K2000 series compounds being more effective than M4K3000 series compounds (Figure 1). Comparison of additional endpoint read outs such as confluency or cell death supports this ranking (Figure 2). Consideration of the effectiveness of each compound in different growth conditions i.e. 2D vs 3D and with or without radiation treatment suggests that M4K2043 and M4K2096 may be less effective in addition to radiation, however M4K2009 remains highly efficient at killing HSJD-DIPG-007 cells (Figure 3). Results in the ACVR1 WT HSJD-DIPG-011 cell line were more variable but indicate that M4K2009 is less effective, with M4K2096 in this case being the most effective probe (Figure 4). These contrasting screen results could indicate that the effectiveness of M4K ACVR1 inhibitory compounds could be dependent on the mutations present. For the HSJD-DIPG-011 cell line, in two out of three biological repeat viability and confluence results are concurrent (Figure 5). Comparison of growth conditions in the HSJD-DIPG-011 cell line is difficult because of the variation

Files (553.7 kB)

Name	Size	Preview	Download
DIPG_cell_M4K_compound_screen_181206.pdf	553.7 kB		

md5:d789819fa3d9dcf0c2938bf154271273

74 views  
31 downloads  
See more details...

7  
Tweeted by 10  
See more details

Indexed in

OpenAIRE

**Publication date:**  
December 6, 2018

**DOI:**  
DOI 10.5281/zenodo.2000673

**Keyword(s):**  
#ipg diffuse intrinsic pontine glioma m4k pharma compound screen viability assay acvr1 alk2

**Communities:**  
Open Notebooks  
SGC Open Notebooks

**License (for files):**  
Creative Commons Attribution 4.0 International

**Versions**

Version 1 10.5281/zenodo.2000673 Dec 5, 2018

**Cite all versions?** You can cite all versions by using the DOI 10.5281/zenodo.2000672. This DOI represents all versions, and will always resolve to the latest one. [Read more](#).

**Share**



**Cite as**  
Brown, Elizabeth, Farnie, Gillian, & Bullock, Alex (2018). Viability screen of ACVR1/ALK2 inhibitory compounds on Diffuse Intrinsic Pontine Glioma cells. <http://doi.org/10.5281/zenodo.2000673>

Start typing a citation style...

**Figure 1. Detailed experimental records including protocols, positive and negative data are posted on Zenodo.** A citable DOI is automatically generated (right-middle panel), and the number of visits and downloads provided (top right, numbers as of December 2018).

## Screening ACVR1 inhibitors on mutant and non-mutant ACVR1 DIPG cells – effectiveness may vary

7th December 2018 Elizabeth Brown [Leave a comment](#)

Hi there! The last month of my life was taken over by making sure my PhD first year report was beautifully polished, but I have returned with results of a small compound screen:

These are all compounds that [Jong Fu](#) has already tested with his assays so we know they effectively inhibit ACVR1, but I thought it would be a good idea to try them out on my DIPG cells and see how well they can kill them. I included a few different variables in this screen to test what could make a difference to the outcome:

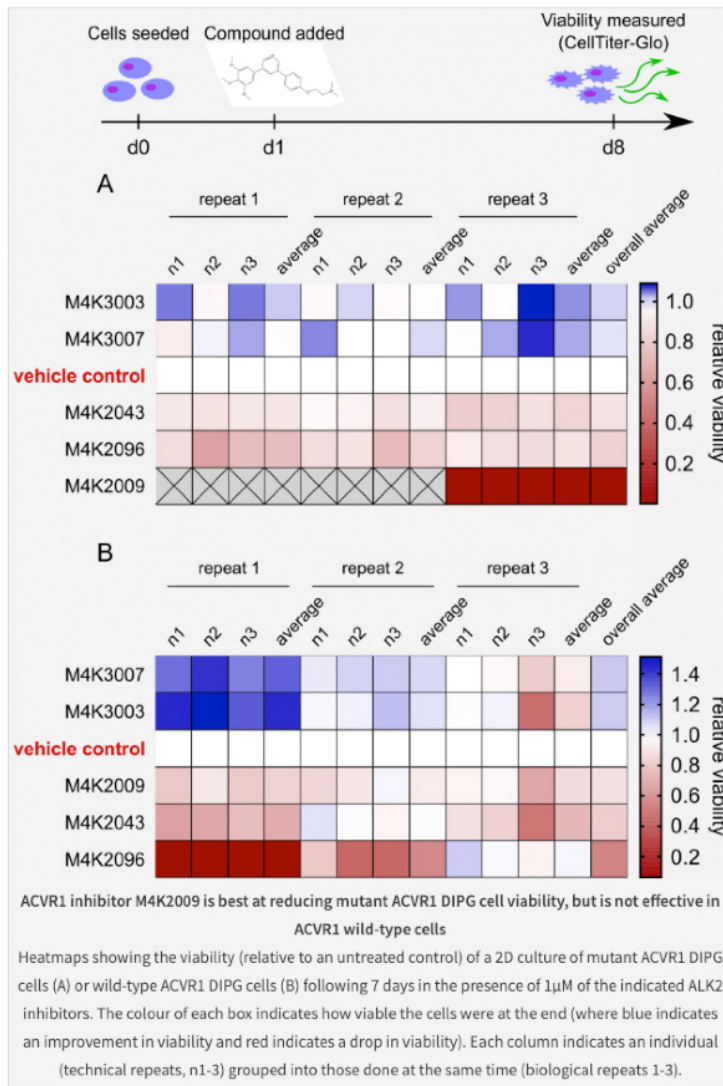
- I tested cells grown in 2D and 3D because this has been noted to change the outcome of compound screens in other systems
- I tested the compounds alone, or in addition to radiation in order to test whether the compounds were effective in addition to current DIPG treatments.

I tested how effective each compound is by growing cells for 7 days in the presence of the compound. At the end I measured how alive or healthy the cells were (i.e. their viability) by using a pre-developed kit that measures the amount of ATP inside the cell by using an enzyme that can produce light when ATP is present. ATP is measured because this is sometimes called the 'energy currency of the cell' so living cells have to have ATP inside them, whereas dead cells won't and won't be counted.

From that heatmap you can see that M4K2009 is the best at killing ACVR1 mutant cells (the top heatmap), but it isn't as effective in cells where the ACVR1 gene is not mutated ("wild-type" cells). In those cells instead M4K2096 is the most effective compound (but still not to the same extent as M4K2009).

But this is only the data for cells grown in 2 dimensions without a radiation treatment. If you want to see the full dataset, you can read my [Zenodo post!](#)

ACVR1/ALK2, Brain Tumours, Diffuse intrinsic pontine glioma, Liz Brown, Understudied Kinases



**Figure 2.** A blog explains in language accessible to non-scientists (such as patient groups) the rationale and take-home message of the experiment.

would be a valid alternative<sup>16</sup>), archived weekly to GitHub (repository <https://github.com/thesgc/static-openlabnotebooks>), quarterly to archive.org ([https://wayback.archive-it.org/6473/\\*/https://opennotebook.thesgc.org/](https://wayback.archive-it.org/6473/*/https://opennotebook.thesgc.org/)), and link to the experimental records, which are deposited at Zenodo ([zenodo.org](https://zenodo.org)), but can also be made available from other public repositories, such as GitHub ([github.com](https://github.com)) or Figshare ([figshare.com](https://figshare.com)). While the experimental details posted at Zenodo are important scientifically, the blog written in layman's term can be used to engage with scientists that may have a complementary set of expertise for future collaborations as well as other stakeholders in the research process, including patient groups, a dimension that most in academia are missing.

The Zenodo repository enables sharing research outputs from across all fields of research, creation and curation of complete digital repositories, flexible licensing with controlled degree of openness and safe storage of the data for the future in the same

cloud infrastructure as CERN's own LHC research data. Open laboratory notebooks need to guarantee that the data will remain accessible, in order to avoid the fate suffered by the pioneer open notebook of Jean-Claude Bradley, which is still accessible while its associated raw data wiki is not. Zenodo is strongly committed to preserving the data it archives. CERN has existed since 1954 and has an experimental program defined for the next 20+ years. Each file has two replicas located on different disk servers. In the highly unlikely event that Zenodo closes operations, they guarantee migration of all content to other suitable repositories, and since all uploads have DOIs, citations and links to Zenodo resources (including data) will not be affected.

The ultimate goal of this open lab notebook initiative is not only to increase the impact of our work but also, along with precursors in the field such as Open Source Malaria (<http://opensourcemalaria.org/>) and other isolated open lab

notebook efforts, to inspire others to follow, and contribute to the creation of a new open science movement in the life sciences. While it is too early to judge the success of this initiative, the number of contributing scientists and institutions is steadily increasing (Figure 3). While only one scientist was contributing in November 2017, 23 scientists from six institutions (University of Toronto, University of Oxford, University of North Carolina, University of Leicester, the Karolinska Institute in Sweden and University of Montpellier in France) are recording their work at [openlabnotebooks.org](http://openlabnotebooks.org) as of December 2018.

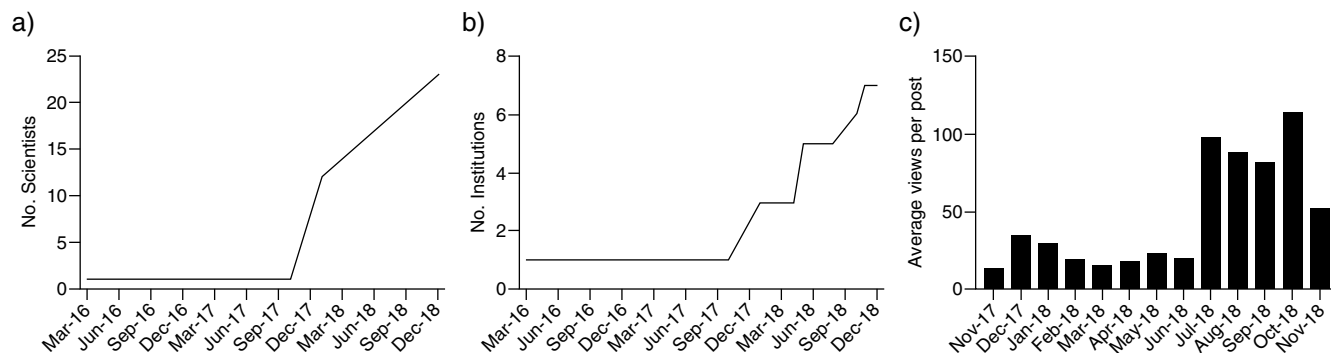
As importantly, impact is also increasing, judging by the average number of views per experimental record calculated from statistical data available at [Zenodo.org](http://Zenodo.org) (Figure 3). Some reports raised a considerable interest. For instance, the crystal structure of USP5 in complex with small molecule fragments has 821 unique views and 324 unique downloads as of December 2018<sup>17</sup>. If the initiative is successful, we anticipate that within three to five years, usage metrics are comparable at [openlabnotebooks.org](http://openlabnotebooks.org) and [bioRxiv](http://bioRxiv), the preprint server for biology.

Data posted at [openlabnotebooks.org](http://openlabnotebooks.org) are raising interest in academic groups, but also in the industry. For instance, a notebook contributor was directly contacted by a big pharmaceutical company to further discuss the results that he had shared online, and a big biotech company asked permission to another contributor to include their data in a presentation at a public scientific meeting. Some of the research reported at [openlabnotebooks.org](http://openlabnotebooks.org) is of direct relevance to patient groups. For instance, four scientists record their results on testing chemical inhibitors of the kinase ALK2, a potential therapeutic target for the treatment of the pediatric brain tumor diffuse intrinsic pontine glioma (DIPG), and the heterotopic ossification disorder fibrodysplasia ossificans progressive (FOP)<sup>18,19</sup>. The compounds, developed by the open science biotech company M4KPharma, are still in pre-clinical phase of development but should ultimately lead to clinical trials for these incurable diseases<sup>20</sup>. Scientists working on projects with a clear path to the clinic are eager to share their enthusiasm and commitment with patient groups (sometimes using social media to announce their latest open notebook post) who, in turn, follow their work.

### The challenges of open lab notebooks

Three antagonizing points that inhibit scientists from starting their own open lab notebook are the fear of being scooped, the inability to report collaborative work when collaborators want to keep data secret, and the concern that an open notebook will take time away from an already overburdened schedule<sup>21</sup>. The language barrier for non-native English speakers, and the availability of open lab notebook solutions can also be challenging. It is indeed likely that maintaining an open lab notebook increases the chances of being scooped, but it is too early at this point to know whether this effect is minor or significant. Paradoxically, and given the territorial nature of the current frameworks for funding and managing scientific research, entries in one's open lab notebook may mark one's area very effectively, especially in a conceivable future when funding trusts and councils start looking into them. We would argue that most, if not all, scientists get scooped during their career, and that open lab notebooks serve as a safety net for early career scientists who have a citable record of their work if they ever get scooped. Obtaining permission from collaborators to report collaborative work in open lab notebooks can be challenging. We believe that the best way to avoid such a situation is to clearly state at the outset of a collaboration the intention to adopt open science principles<sup>22</sup>. Scientists are more likely to agree if presented with the idea well in advance. The time invested in practicing clear, concise and engaging scientific writing is not lost on one's career. After some practice, maintaining an open lab notebook should not take more time than using a regular lab notebook.

Open notebooks being published before peer-review, there is a risk that dubious experiments, erroneous analysis or misinterpretations find their way on open platforms, get amplified over the internet and mislead colleague scientists, patient groups or other communities. Once they become indexed by popular search engines, open lab notebooks could become a source of pollution of the scientific (and non-scientific) literature. This risk, which is not limited to open notebooks but extends to the increasing number of Journals that adopt a post-publication peer-review mechanism, is real, serious, and should be monitored. We believe that the best way to mitigate this risk is for open notebooks to provide a platform for open comments. In



**Figure 3. Information on openlabnotebooks.org.** The Number of (a) scientists and (b) institutions actively contributing to [openlabnotebooks.org](http://openlabnotebooks.org). (c) The average number of unique visits for each experimental record.

principle, this could be an even stronger quality control than the current peer-review system in place in most scientific journals, as the number of “open reviewers” for any given report is limitless. At the moment, we find that very few comments are posted at [openlabnotebooks.org](http://openlabnotebooks.org), a platform that is only a year old, but we see that comments are mainstream, and sometimes turn into healthy discussions at Open Source Malaria, a pioneer in the field<sup>23</sup>.

### Future directions and conclusion

Open lab notebooks represent a major departure from current practices in science (especially biomedical sciences) and hold a mix of promises and risks. As the community producing these lab notebooks is increasing, there is an opportunity to move beyond ideology and anecdotal data to evidence-based policy design. In the spirit of openness, we call on colleagues from both the life science and the social sciences communities to conduct systematic evaluation of the benefits and downsides of open lab notebooks. It will be important to compare several parameters on a yearly basis. These may include the frequency of research being scooped among scientists disclosing their work in open lab notebooks versus a less open reference group; the frequency of new collaborations; the frequency of comments and ideas received by the authors of open notebooks; and instances where open lab notebooks were essential for compliance with funder or institutional requirements. More difficult to assess will be issues such as recognition, career progression, speeding up research, and impact on reproducibility, but they could all be addressed with appropriate questionnaires and data analytics.

Our goal is to see the number of open lab notebooks increase exponentially over the coming years. Future implementation of novel features, such as the ability to search for experiments containing compounds with specific chemical templates, is expected to extend the reach of the platform to medicinal and computational chemists. Indexing of open lab notebooks by popular search engines such as Google Scholar (which already indexes pre-prints and other non-peer-reviewed documents) would increase the visibility and impact of open notebooks. Importantly, open lab notebook data deposited at [Zenodo.org](http://Zenodo.org) is already searchable with [Google's Dataset search engine](#). To further encourage scientists to break free from the tradition of secrecy that has been passed on for generations, a cultural change needs to be supported at institutional and governmental levels. Funding bodies are starting to define and enforce open science publication practices<sup>24</sup>. Similarly, universities could take a more proactive role, for instance by including adherence to open-access principles as an evaluation criteria

for career advancement<sup>25</sup>. Indeed, while strong incentives described above already exist for junior scientists to start their own open lab notebook, the benefit to their PIs who already have established a professional network and don't need to showcase their skills is not always as clear. As long as scientists are not convinced that open science is good for them, Science 2.0 will have to wait.

### Data availability

No data are associated with this article.

### Grant information

The Structural Genomics Consortium (SGC) is a registered charity (number 1097737) that receives funds from AbbVie, Bayer Pharma AG, Boehringer Ingelheim, Canada Foundation for Innovation, Eshelman Institute for Innovation, Genome Canada through Ontario Genomics Institute [OGI-055], Innovative Medicines Initiative (EU/EFPIA) [ULTRA-DD grant no. 115766], Janssen, Merck KGaA, Darmstadt, Germany, MSD, Novartis Pharma AG, Ontario Ministry of Research, Innovation and Science (MRIS), Pfizer, São Paulo Research Foundation-FAPESP, Takeda, and The Wellcome Trust [106169/ZZ14/Z]. RJH is a recipient of the HDSA Berman/Topper HD Career Development Fellowship. PR is funded by the Wellcome Trust Leicester ISSF award reference 204801/Z/16/Z and the Leicester Institute of Chemical and Structural Biology (LISCB).

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

### Acknowledgments

We thank the following principal investigators whose group members are contributing to [openlabnotebooks.org](http://openlabnotebooks.org): Cheryl Arrowsmith, Dalia Barsyte, Paul Brennan, Alex Bullock, David Drewry, Susanne Gräslund, Brian Marsden, Dave Morris, Panagiotis Prinos, Frank Von Delft, Tim Willson, and Wyatt Yue. Nicholas Worby at University of Toronto Libraries set-up the quarterly archiving of [openlabnotebooks.org](http://openlabnotebooks.org) to <https://archive.org/>.

As of December 2018, the Open Lab Notebook Consortium is made up of Roslin Adamson, Jose Brandao-Neto, Elizabeth J. Brown, Antoine Claessens, David Damerell, David Dilworth, Thomas Durcan, Benjamin J. Eduful, Aled M. Edwards, Opher Gileadi, Jolene Caifeng Ho, Leonidas Koukouffis, Tobias Krojer, Genna M. Luciani, Sabrina Mackinnon, Mandeep Mann, Carolyn Marks, Sean O'Byrne, Alfredo Picado, Pietro Roversi, Louisa Temme, Eleanor Williams, Jong Fu Wong, Wen Yih Aw.

### References

- Smith R: **Peer review: a flawed process at the heart of science and journals.** *J R Soc Med.* 2006; **99**(4): 178–82.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Resnik DB: **Openness versus Secrecy in Scientific Research Abstract.** *Episteme (Edinb).* 2006; **2**(3): 135–147.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Woelfle M, Olliaro P, Todd MH: **Open science is a research accelerator.** *Nat Chem.* 2011; **3**(10): 745–8.  
[PubMed Abstract](#) | [Publisher Full Text](#)
- Powell K: **Does it take too long to publish research?** *Nature.* 2016; **530**(7589): 148–51.  
[PubMed Abstract](#) | [Publisher Full Text](#)

5. Wallach JD, Boyack KW, Ioannidis JPA: **Reproducible research practices, transparency, and open access data in the biomedical literature, 2015–2017.** *PLoS Biol.* 2018; **16**(11): e2006930.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
6. **Reality check on reproducibility.** *Nature.* 2016; **533**(7604): 437.  
[PubMed Abstract](#) | [Publisher Full Text](#)
7. Mlinari A, Horvat M, Šupak Smolčić V: **Dealing with the positive publication bias: Why you should really publish your negative results.** *Biochem Med (Zagreb).* 2017; **27**(3): 030201.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
8. Carroll HA, Toumpakari Z, Johnson L, *et al.*: **The perceived feasibility of methods to reduce publication bias.** *PLoS One.* 2017; **12**(10): e0186472.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
9. Grubb AM, Easterbrook SM: **On the lack of consensus over the meaning of openness: an empirical study.** *PLoS One.* 2011; **6**(8): e23420.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
10. Nielsen LH: **Sharing your data and software on Zenodo.** 2017.
11. Wilkinson MD, Dumontier M, Aalbersberg IJ, *et al.*: **The FAIR Guiding Principles for scientific data management and stewardship.** *Sci Data.* 2016; **3**: 160018.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
12. Drexel CoAS E-Learning: **Open Notebook Science.**  
[Reference Source](#)
13. Harding RJ: **Open notebook science can maximize impact for rare disease projects.** *PLoS Biol.* 2019; **17**(1): e3000120.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
14. **Open notebooks galore: The Structural Genomics Consortium.** *eLife.* [Accessed: 24-Dec-2018] 2018.  
[Reference Source](#)
15. Schapira M: **Open Lab Notebooks to increase impact and accelerate discovery.** *Research Data at Springer Nature.* [Accessed: 24-Dec-2018] 2018.  
[Reference Source](#)
16. Badiola KA, Bird C, Brocklesby WS, *et al.*: **Experiences with a researcher-centric ELN.** *Chem Sci.* 2015; **6**(3): 1614–1629.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
17. Mann M, Harding R, Ravichandran M, *et al.*: **Co-crystal structures of USP5 ZF-UBD and weak binding compounds.** *Zenodo.* 2018.  
[Publisher Full Text](#)
18. van Dinther M, Visser N, de Gorter DJ, *et al.*: **ALK2 R206H mutation linked to fibrodysplasia ossificans progressiva confers constitutive activity to the BMP type I receptor and sensitizes mesenchymal cells to BMP-induced osteoblast differentiation and bone formation.** *J Bone Miner Res.* 2010; **25**(6): 1208–1215.  
[PubMed Abstract](#) | [Publisher Full Text](#)
19. Taylor KR, Vinci M, Bullock AN, *et al.*: **ACVR1 Mutations in DIPG: lessons learned from FOP.** *Cancer Res.* 2014; **74**(17): 4565–4570.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
20. Morgan MR, Roberts OG, Edwards AM: **Ideation and implementation of an open science drug discovery business model – M4K Pharma.** *Wellcome Open Res.* 2018; **3**: 154.  
[Publisher Full Text](#)
21. Robertson MN, Ylloja PM, Williamson AE, *et al.*: **Open source drug discovery - a limited tutorial.** *Parasitology.* 2014; **141**(1): 148–157.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
22. Masum H, Rao A, Good BM, *et al.*: **Ten simple rules for cultivating open science and collaborative R&D.** *PLoS Comput Biol.* 2013; **9**(9): e1003244.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
23. 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](#)
24. Else H: **Radical open-access plan could spell end to journal subscriptions.** *Nature.* 2018; **561**(7721): 17–18.  
[PubMed Abstract](#) | [Publisher Full Text](#)
25. Alperin JP, Fischman GE, McKiernan EC, *et al.*: **How significant are the public dimensions of faculty work in review, promotion, and tenure documents?** 2018.  
[Publisher Full Text](#)



# Open Peer Review

Current Peer Review Status:



---

Version 2

Reviewer Report 12 August 2019

<https://doi.org/10.5256/f1000research.20551.r50782>

© 2019 Poline J. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution Licence](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



**Jean-Baptiste Poline** 

McGill University, McGill University, Montreal, QC, Canada

The article by Matthieu Schapira and Rachel Harding "Open laboratory notebooks: good for science, good for society, good for scientists" is a clear presentation of the benefits of open notebook for the scientific community. The central argument is that one of the main benefits is an increased efficiency (both in time and cost) of the scientific process through the faster re-utilization of positive or negative results. The authors discuss in detail the challenges of the initiative and the results of the current openlabnotebook.org implementation and results.

I liked in particular the straightforward discussion on the current incentive system from the authors. The perceived or real danger of being scooped is a clear obstacle to the efficient use of these. The arguments of why opennotebooks are good for scientists are clear and compelling although based mostly on anecdotal data.

The implementation of the infrastructure is solid, reusing standard and long lasting components (zenodo, github). The search capacity across notebooks could have been detailed more. An component around the standardization of terms, use of controlled vocabulary (or even ontologies) could be added in the future. For instance, an extension of the figshare data documentation system would be helpful to label notebook appropriately.

The uptake in both the building and the use of openlabnotebooks is promising, although absolute number of adaption are still limited since this is still at early stages.

The challenges for wider adoption are clear - (namely, fear of being scooped, collaborators with non open practices, and fear that the time invested would not be worth (in terms of recognition), risk of non solid information being released. The solutions to mitigate these challenges are not as well developed, but as this is a new technology and practice, only the future will say if these are indeed going to prevent the openlabnotebooks developments. I think that the close involvement of both national and private funding agencies could rapidly solve for these adoption challenges in a top-down manner.

Last, I have one question: what is the process to join the openlabnotebooks consortium?

**Is the topic of the opinion article discussed accurately in the context of the current literature?**

Yes

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

Yes

**Are arguments sufficiently supported by evidence from the published literature?**

Yes

**Are the conclusions drawn balanced and justified on the basis of the presented arguments?**

Yes

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

**Reviewer Expertise:** Brain Imaging data analysis, reproducible science, Open science, Neuro-informatics, Imaging genetics, neuro-data-science, publishing and culture of science

**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 18 April 2019

<https://doi.org/10.5256/f1000research.20551.r46587>

© 2019 Mangravite L. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution Licence](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



**Lara M. Mangravite** 

Sage Bionetworks, Seattle, WA, USA

The authors did a reasonable job of addressing my questions. I have no further concerns.

**Is the topic of the opinion article discussed accurately in the context of the current literature?**

Yes

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

Yes

**Are arguments sufficiently supported by evidence from the published literature?**

Yes

**Are the conclusions drawn balanced and justified on the basis of the presented arguments?**

Yes

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

**Reviewer Expertise:** open science, bioinformatics, pharmaceutical science

**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.**

---

**Version 1**

Reviewer Report 04 March 2019

<https://doi.org/10.5256/f1000research.19363.r43420>

© 2019 Mangravite L. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution Licence](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

**Lara M. Mangravite**

Sage Bionetworks, Seattle, WA, USA

The opinion piece on Open Laboratory Notebooks written by Drs Schapira, Harding, and the Open Lab Notebook Consortium provides a thoughtful commentary on an interesting and relevant topic. The Open Labnotebook provides an interesting and potentially powerful approach to support transition of scientific communication beyond traditional peer-reviewed scholarly publications. As the authors note, the availability of digital media provides a much wider array of approaches to support meaningful research dissemination – and to promote scientific reuse and collaboration.

The authors speak primarily from their own experience – and, because the use of open lab notebooks is quite new, there is often not a lot of data to back up their comments. As such, it was hard to understand what ideas were well supported and what ideas were in need of evaluation. Along those lines, the authors do not put their work into context of the broader field – I do not have a clear sense of who else is using an open labnotebook approach, whether their experiences are similar or different, and what portions of the scientific community were finding open labnotebooks useful. This context would help in synthesizing the ideas in the manuscript.

Secondly, it would be enormously helpful to understand how the authors anticipate that open labnotebooks would operate in the context of the current scientific communication framework. Do you expect labnotebooks to replace peer-review journals or to complement them? How might a researcher convey the overarching hypothesis that they are looking to address through the set of experiments run across a series of notebook entries? It seems that neither the labnotebook approach nor the existing publication approach are sufficient to fully convey scientific output by themselves. I would very much enjoy understanding the authors' perspective on this.

Along these same lines, I am curious as to how notebook readers will incorporate this form of communication into their workflow. It will have to be useful – and easy – in order to gain the desired level of adoption. I appreciated the inclusion of the notebook use statistics – it would be helpful to understand how users found these posts, what kind of users were attracted, and how they responded to this medium as a form of scientific discourse.

The role of the gatekeeper was also not addressed. Open LabNotebooks are reliant on a series of technical platforms. These platforms will have an impact on the overall success based on their sustainability models, their archiving approaches, and their need for financial viability. Is there a cost to using the system – and, if not, how is it funded? Are the content fully owned by the authors and do they remain under full control of the authors? What are the responsibilities of the platforms and how might their choices impact the approach?

Finally, the authors have a tremendous opportunity to ground some of this theoretical work in their own practical application. It would be useful to understand how your scientists' experiences compared to those were anticipated. For example, one frequent concern heard by researchers is one of time – they are worried that use of the system will require an additional burden in an already over-committed schedule. How much time do the current users spend on their posts? Did this reduce over time? This same approach can be used to address the anticipate benefits.

**Is the topic of the opinion article discussed accurately in the context of the current literature?**

Partly

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

Partly

**Are arguments sufficiently supported by evidence from the published literature?**

Partly

**Are the conclusions drawn balanced and justified on the basis of the presented arguments?**

Yes

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

**Reviewer Expertise:** open science, bioinformatics, pharmaceutical science

**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, however I have significant reservations, as outlined above.**

Author Response 26 Mar 2019

**Matthieu Schapira**, SGC, Toronto, Canada

*The authors speak primarily from their own experience – and, because the use of open lab notebooks is quite new, there is often not a lot of data to back up their comments. As such, it was hard to understand what ideas were well supported and what ideas were in need of evaluation. Along those lines, the authors do not put their work into context of the broader field – I do not have a clear sense of who else is using an open labnotebook approach, whether their experiences are similar or different, and what portions of the scientific community were finding open labnotebooks useful. This context would help in synthesizing the ideas in the manuscript.*

Open lab notebooks remain a niche of scholarly communication and the number of active examples is limited. An additional section has been added in “Implementation of an open lab notebook platform” describing the brief history and status of the open lab notebook community:

“Open lab notebooks have been pioneered and championed by a number of practitioners but remain a niche activity in the scientific community. Jean-Claude Bradley first coined the term “open notebook science” in 2006 and his definition of this method of scholarly communication have laid the foundations for our own efforts [12]. In addition to the notebooks of individual researchers following Bradley’s template, open notebook examples now include collective efforts from the Open Lab Notebook Network (<http://onsnetwork.org>) and Open Source Malaria (<http://opensourcemalaria.org>). However, the open lab notebook community remains small, the practice is not consistently defined or implemented and the impact of these efforts in the field have not been systematically evaluated.”

*Secondly, it would be enormously helpful to understand how the authors anticipate that open labnotebooks would operate in the context of the current scientific communication framework. Do you expect labnotebooks to replace peer-review journals or to complement them? How might a researcher convey the overarching hypothesis that they are looking to address through the set of experiments run across a series of notebook entries? It seems that neither the labnotebook approach nor the existing publication approach are sufficient to fully convey scientific output by themselves. I would very much enjoy understanding the authors’ perspective on this.*

We see open lab notebooks as a mechanism for scientists to share their work before they publish it in peer-reviewed journals. As such, they are complementary to peer-reviewed publication. This is specified in the article as follows:

“We believe that open laboratory notebooks, where research scientists record their work online and in near-real time, are an efficient way to disseminate data before it is published in peer-reviewed journals, and ...”

*Along these same lines, I am curious as to how notebook readers will incorporate this form of communication into their workflow. It will have to be useful – and easy – in order to gain the desired level of adoption. I appreciated the inclusion of the notebook use statistics – it would be helpful to understand how users found these posts, what kind of users were attracted, and how they responded to this medium as a form of scientific discourse.*

We agree that much is to be gained from knowing the demographics and motivations of open lab notebook readers. The readers input in the well populated comment section of the platform Open Source Malaria strongly suggest that most readers are scientists in the field, sharing common goals and interests.

*The role of the gatekeeper was also not addressed. Open LabNotebooks are reliant on a series of technical platforms. These platforms will have an impact on the overall success based on their sustainability models, their archiving approaches, and their need for financial viability. Is there a cost to using the system – and, if not, how is it funded? Are the content fully owned by the authors and do they remain under full control of the authors? What are the responsibilities of the platforms and how might their choices impact the approach?*

The server hosting [openlabnotebooks.org](http://openlabnotebooks.org) is maintained by the Structural Genomics Consortium at University of Oxford. It relies on Wordpress, which is an open source platform, and requires limited maintenance. Consequently, the service is free for both contributors and readers. Zenodo is hosted by the CERN and is also free. The issue of permanence, sustainability, archiving and

ownership are important ones and are addressed in our response to Reviewer 1 (sections #6 and #17) and discussed in the revised version of the manuscript.

*Finally, the authors have a tremendous opportunity to ground some of this theoretical work in their own practical application. It would be useful to understand how your scientists' experiences compared to those were anticipated. For example, one frequent concern heard by researchers is one of time – they are worried that use of the system will require an additional burden in an already over-committed schedule. How much time do the current users spend on their posts? Did this reduce over time? This same approach can be used to address the anticipate benefits.*

When initiating this project, prospective open notebookers were surveyed regarding their concerns, questions and attitudes about running their own open notebook projects which highlighted many of the concerns you raise, in particular, the time burden. We are planning to conduct a follow-up survey once the number of scientists joining our platform increases.

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

Reviewer Report 06 February 2019

<https://doi.org/10.5256/f1000research.19363.r43417>

© 2019 Todd M et al. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution Licence](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



**Matthew H. Todd**

School of Chemistry, The University of Sydney, Sydney, NSW, Australia

**Edwin Tse**

University of Sydney, Sydney, Australia

**Marat Korsik**

University of Sydney, Sydney, Australia

**Mathamsanqa Bhebhe**

University of Sydney, Sydney, Australia

This opinion piece is on a timely, important topic and is clearly and engagingly written. Anecdotally, we find that many of our colleagues in science are unaware that open lab notebooks exist. This article will help.

The authors identify several important advantages and challenges associated with the near-immediate deposition of results into the public domain, online. They use examples from their own research to highlight the possibilities.

The refereeing team behind this review are seasoned users of open lab notebooks, and so are in a good position to judge the piece. We judge it to have cleared peer review from our perspective, once the following comments and suggestions have been acted upon. There are a number, which should be read

not as criticism but as testament to our shared enthusiasm for this subject and its importance in the future of research.

1) **Secrecy.** In the introduction, reasons are suggested for why scientists may keep results secret. We would suggest that there are two important reasons that are not explicitly mentioned: i) that the scientist may want to patent something, and ii) that the scientist cannot be bothered to work out how to release research using atypical means. The first point is alluded to where mention is made of ownership, and the second point is alluded to by the mention of "paper" but we would argue these two factors are significant enough that they should be made explicit.

2) **Careers.** We'd be interested in whether there is a justification for the statement "Many believe that openly sharing work online will limit career opportunities." If there is none, then perhaps rephrase this more as a possibility?

3) **Grants.** The statement "Grant applications that highlight the use of open lab notebooks are being viewed positively" may be true (one hopes it is), but the evidence presented doesn't support that statement (the grants may have been funded because the science was so good, regardless of the dissemination plan), so again, this probably needs to be made more aspirational.

4) **Errors.** The authors need to address what happens if an experiment is recorded containing a mistake, and a mistake that might propagate *via* an ungrounded conclusion. Is there a danger in leading colleagues down the wrong path? Is there a danger to the reputation of young scientists? Would incorrect conclusions, if indexed by search engines, lead to literature pollution that might be hard to correct?

5) **Machines.** One must make the assumption that few people will ever *read* open lab notebooks, just as few people read regular lab notebooks. More needs to be made of this, since a) it is essential, rather than just desirable, that the notebook is searched and indexed by e.g. Google, and b) the entries need, ideally, to be machine readable. Could the authors comment on this - i.e. how Googleable the contents are (beyond the Dataset Search they mention) and what can be done to ensure that the entries can be understood by machines?

6) **Permanence.** There are many web links in the article. This is an academic publication, which is intended to last forever. It's likely those URLs will not last forever. Is there a way that the pages pointed to can be archived somehow? An important, relevant example: this manuscript does not refer to one of the pioneers of open lab notebooks - indeed he coined the phrase "open notebook science" - Jean-Claude Bradley. Bradley's blog can still be **accessed** but the wiki that used to house all the raw data has **gone**. In publishing our most recent paper that made use of open lab notebooks<sup>1</sup>, we took pains to archive the lab notebooks on a repository to mitigate potential loss by external providers, and to back up web pages (as PDFs) to similar places. Can the authors address the two issues here: i) should we link to web pages in academic articles without backing them up? 2) How are the lab notebooks backed up and archived?

7) **Time Stamps.** The authors mention "results in open lab notebook are date-stamped." It would be interesting to know whether the authors have looked into whether the date stamps are "convincing" from a legal standpoint. Would they, for example be sufficient to claim priority in a legal challenge? Is there an issue with the fact that ELN pages can be edited, over an extended period of time, or is it sufficient to ensure that the page has a robust revision history?

8) **Scooping.** A scooper could claim that they were simply unaware of the open scientist's work. This seems a reasonable defense. What needs to happen for ignorance of an open lab notebook not to be a

defense against ignorance of prior art?

9) **Citations.** We are cited twice (thank you), but there are two other relevant papers the authors might be interested in citing should these be deemed by them to be appropriate: i) an extensive discussion on the use cases of an open source ELN called Labtrove<sup>2</sup> pioneered by Jeremy Frey's team at Southampton, and ii) a large medchem project by Open Source Malaria which was conducted entirely using open lab notebooks<sup>1</sup> and which may be instructive in terms of how to publish a paper based on open notebook work (also described informally [here](#)).

10) **Examples of Inputs.** Are there any examples the authors can point to where the open ELN has helped the research, e.g. where suggestions have been made, and acted upon, to help the science?

11) **Examples of ELNs.** Are there other examples of open lab notebooks being used in biomedical research? Is this, in fact, still a highly niche activity?

12) **Language.** The language barrier is mentioned, but we're not sure this is relevant to this paper. Language is already a barrier in the current system, though English is essentially a *lingua franca* across science.

13) **License.** What is the license that covers the authors' i) blog posts and ii) notebooks? To what extent is this an important choice? For example, can others take and re-use the content without restriction, or only for non-commercial purposes?

14) **Raw Data.** The authors mention that the ELN entries include "all data". This is an important feature of open ELNs that distinguishes the practice from a great deal of open science in which highlights might be discussed, or blogged about, but without the attendant raw data. We investigated briefly to see how much of the raw data could be accessed in the authors' cases. We have noticed that a fair number of the lab notebook entries recorded by the team in Zenodo contains PDF/Word-style summaries of data, rather than the data themselves. This limits re-use of the data. Could the authors comment on the central importance of the availability of all raw data?

Examples:

- [This](#) blog post mentions performing synthesis of analogues based on lit procedures but there's no link to an ELN.
- Same with [this](#) page.
- And [this](#).
- [This](#) page has a link (in the 2nd paragraph) to a (structure?) data file but no backlink from the data back to the blog page, risking orphaning of content?
- [This](#) blog page does have a link to the ELN. There is more experimental detail, but there is no raw characterization data. Also the [backlink](#) to the blog page is broken.
- Many biology entries contain pictures of purifications and for the most part, the ELN pages only have a docx or pdf there. Should there be any raw data available for these kinds of pages, or are docx and PDF enough?

15) **Identifiers.** It is mentioned that is is desirable to have "the ability to search for experiments containing compounds with specific chemical templates". We agree. Some ELNs already allow this (e.g. Chemotion<sup>3</sup> and C6H6<sup>4</sup>). We have found a stop-gap is the manual inclusion of chemical strings such as SMILES, InChI. Can the authors comment on whether similar things can be included in order to allow machines to understand the biology contained within the entries, e.g. UniProt numbers?



16) **Metrics re Access.** Is there a way of distilling out access by non-team people, or even access by the author of the page who may have accessed multiple times during the editing of a page?

17) **Ownership.** When a researcher moves on, does that researcher have any responsibility towards the ELN? e.g. if a mistake is found, who needs to correct it? Is it the responsibility of the PI to act as future curator?

## References

1. Williamson AE, Ylioja PM, Robertson MN, Antonova-Koch Y, Avery V, Baell JB, Batchu H, Batra S, Burrows JN, Bhattacharyya S, Calderon F, Charman SA, Clark J, Crespo B, Dean M, Debbert SL, Delves M, Dennis AS, Deroose F, Duffy S, Fletcher S, Giaever G, Hallyburton I, Gamo FJ, Gebbia M, Guy RK, Hungerford Z, Kirk K, Lafuente-Monasterio MJ, Lee A, Meister S, Nislow C, Overington JP, Papadatos G, Patiny L, Pham J, Ralph SA, Ruecker A, Ryan E, Southan C, Srivastava K, Swain C, Tarnowski MJ, Thomson P, Turner P, Wallace IM, Wells TN, White K, White L, Willis P, Winzeler EA, Wittlin S, Todd MH: 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](#)
2. Badiola KA, Bird C, Brocklesby WS, Casson J, Chapman RT, Coles SJ, Cronshaw JR, Fisher A, Frey JG, Gloria D, Gossel MC, Hibbert DB, Knight N, Mapp LK, Marazzi L, Matthews B, Milsted A, Minns RS, Mueller KT, Murphy K, Parkinson T, Quinnell R, Robinson JS, Robertson MN, Robins M, Springate E, Tizzard G, Todd MH, Williamson AE, Willoughby C, Yang E, Ylioja PM: Experiences with a researcher-centric ELN. *Chem Sci.* 2015; **6** (3): 1614-1629 [PubMed Abstract](#) | [Publisher Full Text](#)
3. Tremouilhac P, Nguyen A, Huang YC, Kotov S, Lütjohann DS, Hübsch F, Jung N, Bräse S: Chemotion ELN: an Open Source electronic lab notebook for chemists in academia. *J Cheminform.* 2017; **9** (1): 54 [PubMed Abstract](#) | [Publisher Full Text](#)
4. Patiny L, Zasso M, Kostro D, Bernal A, Castillo AM, Bolaños A, Asencio MA, Pellet N, Todd M, Schloerer N, Kuhn S, Holmes E, Javor S, Wist J: The C6H6 NMR repository: An integral solution to control the flow of your data from the magnet to the public. *Magn Reson Chem.* 2018; **56** (6): 520-528 [PubMed Abstract](#) | [Publisher Full Text](#)

**Is the topic of the opinion article discussed accurately in the context of the current literature?**

Partly

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

Partly

**Are arguments sufficiently supported by evidence from the published literature?**

Partly

**Are the conclusions drawn balanced and justified on the basis of the presented arguments?**

Partly

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

**Reviewer Expertise:** Open source drug discovery, organic and medicinal chemistry

**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, however we have significant reservations, as outlined above.**

Author Response 26 Mar 2019

**Matthieu Schapira**, SGC, Toronto, Canada

1) **Secrecy.** *In the introduction, reasons are suggested for why scientists may keep results secret. We would suggest that there are two important reasons that are not explicitly mentioned: i) that the scientist may want to patent something, and ii) that the scientist cannot be bothered to work out how to release research using atypical means. The first point is alluded to where mention is made of ownership, and the second point is alluded to by the mention of "paper" but we would argue these two factors are significant enough that they should be made explicit.*

Points well taken. The following statement was added to the Introduction "...and can be compounded by constraints associated with patent protection procedures or the absence of clear mechanism to make one's data publicly available."

2) **Careers.** *We'd be interested in whether there is a justification for the statement "Many believe that openly sharing work online will limit career opportunities." If there is none, then perhaps rephrase this more as a possibility?*

This was not clear. The sentence was replaced as follows:

"Many believe that the chances of getting scooped before one publishes their work in a peer-reviewed journal increase when openly sharing their work online [9]"

3) **Grants.** *The statement "Grant applications that highlight the use of open lab notebooks are being viewed positively" may be true (one hopes it is), but the evidence presented doesn't support that statement (the grants may have been funded because the science was so good, regardless of the dissemination plan), so again, this probably needs to be made more aspirational.*

This was revised as follows:

"Our personal observations seem to indicate that grant applications highlighting the use of open lab notebooks are being viewed positively."

4) **Errors.** *The authors need to address what happens if an experiment is recorded containing a mistake, and a mistake that might propagate via an ungrounded conclusion. Is there a danger in leading colleagues down the wrong path? Is there a danger to the reputation of young scientists? Would incorrect conclusions, if indexed by search engines, lead to literature pollution that might be hard to correct?*

We agree that this important point was missing. The following was added to the section "The challenges of open lab notebooks":

Open notebooks being published before peer-review, there is a risk that dubious experiments, erroneous analysis or misinterpretations find their way on open platforms, get amplified over the internet and mislead colleague scientists, patient groups or other communities. Once they become indexed by popular search engines, open lab notebooks could become a source of pollution of the scientific (and non-scientific) literature. This risk, which is not limited to open notebooks but extends to the increasing number of Journals that adopt a post-publication peer-review

mechanism, is real and serious. We believe that the best way to mitigate this risk is for open notebooks to provide a platform for open comments. In principle, this could be an even stronger quality control than the current peer-review system in place in most scientific journals, as the number of “open reviewers” for any given report is limitless. At the moment, we find that very few comments are posted at [openlabnotebooks.org](http://openlabnotebooks.org), a platform that is only a year old, but we see that comments are mainstream, and sometimes turn into healthy discussions at Open Source Malaria, a pioneer in the field.

**5) Machines.** *One must make the assumption that few people will ever read open lab notebooks, just as few people read regular lab notebooks. More needs to be made of this, since a) it is essential, rather than just desirable, that the notebook is searched and indexed by e.g. Google, and b) the entries need, ideally, to be machine readable. Could the authors comment on this - i.e. how Googleable the contents are (beyond the Dataset Search they mention) and what can be done to ensure that the entries can be understood by machines?*

We believe that in the life sciences, most would find the constraint of formatting their notebook in a standardized language an unacceptable burden. Open lab notebooks should be indexable by Google or other search engines as free text records. Chemists would benefit from the integration of cheminformatics functionality, such as fingerprint-based substructure or similarity searches. These tools are openly provided by RDKit, and we expect that they will soon be integrated into open lab notebook platforms.

**6) Permanence.** *There are many web links in the article. This is an academic publication, which is intended to last forever. It's likely those URLs will not last forever. Is there a way that the pages pointed to can be archived somehow? An important, relevant example: this manuscript does not refer to one of the pioneers of open lab notebooks - indeed he coined the phrase "open notebook science" - Jean-Claude Bradley. Bradley's blog can still be [accessed](#) but the wiki that used to house all the raw data has [gone](#). In publishing our most recent paper that made use of open lab notebooks<sup>1</sup>, we took pains to archive the lab notebooks on a repository to mitigate potential loss by external providers, and to back up web pages (as PDFs) to similar places. Can the authors address the two issues here: i) should we link to web pages in academic articles without backing them up? 2) How are the lab notebooks backed up and archived?*

We appreciate this concern. We have multiple layers of backup for the [openlabnotebooks.org](http://openlabnotebooks.org) web-site including snapshots, replication, and off-site tape storage. We also export all blog posts automatically at the end of each week to the following GitHub repository <https://github.com/thesgc/static-openlabnotebooks>. The Information Technology Services at University of Toronto is now also maintaining an archive of the website at [https://wayback.archive-it.org/6473/\\*/https://opennotebook.thesgc.org/](https://wayback.archive-it.org/6473/*/https://opennotebook.thesgc.org/), which is updated quarterly. The data posted on Zenodo is also backed-up, as specified at <https://about.zenodo.org/infrastructure/>. Archiving each website that one cites is an impressive commitment to the permanence of one's work. We assume that consent should first be obtained to archive someone else's research output. We are not planning to follow this path at the moment. The following statement was added to the article:

“The blogs, posted at [openlabnotebooks.org](http://openlabnotebooks.org), are managed by a webserver downloaded from [wordpress.org](http://wordpress.org), archived weekly to GitHub (repository <https://github.com/thesgc/static-openlabnotebooks>), quarterly to [archive.org](http://archive.org)

([https://wayback.archive-it.org/6473/\\*/https://opennotebook.thesgc.org/](https://wayback.archive-it.org/6473/*/https://opennotebook.thesgc.org/)), and link..."

We also added the following section:

"Open laboratory notebooks need to guarantee that the data will remain accessible, in order to avoid the fate suffered by the pioneer open notebook of Jean-Claude Bradley, which is still accessible while its associated raw data wiki is not. Zenodo is strongly committed to preserving the data it archives. CERN has existed since 1954 and has an experimental program defined for the next 20+ years. Each file copy has two replicas located on different disk servers. In the highly unlikely event that Zenodo closes operations, they guarantee migration of all content to other suitable repositories, and since all uploads have DOIs, citations and links to Zenodo resources (including data) will not be affected."

*7) **Time Stamps.** The authors mention "results in open lab notebook are date-stamped." It would be interesting to know whether the authors have looked into whether the date stamps are "convincing" from a legal standpoint. Would they, for example be sufficient to claim priority in a legal challenge? Is there an issue with the fact that ELN pages can be edited, over an extended period of time, or is it sufficient to ensure that the page has a robust revision history?*

We believe that whether a particular electronic date stamp is sufficient to prove the date of disclosure would depend on how well the system was configured and how good the records were logged. Website businesses do depend on date stamping to establish when a particular change to their terms & conditions was made, for example. Our understanding is that the system implemented at Zenodo clearly indicates when a record was posted and when it was edited, which should be sufficient evidence for open notebooks to count as prior art.

To clarify the date-stamp/versioning mechanism in place at Zenodo, we added the statement "once a record has been published, it can no longer be modified, but revised versions can be appended if necessary"

*8) **Scooping.** A scooper could claim that they were simply unaware of the open scientist's work. This seems a reasonable defense. What needs to happen for ignorance of an open lab notebook not to be a defense against ignorance of prior art?*

We believe that scientists learn to live with the potential and reality of being scooped. Plagiarism is another problem, but we believe that it is more an academic conduct issue than a legal issue. If materials & methods and results published in a date-stamped open notebook are later authored by another scientist in a peer-reviewed journal, this should in principle raise reputational problem for that person.

*9) **Citations.** We are cited twice (thank you), but there are two other relevant papers the authors might be interested in citing should these be deemed by them to be appropriate: i) an extensive discussion on the use cases of an open source ELN called Labtrove <sup>2</sup> pioneered by Jeremy Frey's team at Southampton, and ii) a large medchem project by Open Source Malaria which was conducted entirely using open lab notebooks <sup>1</sup> and which may be instructive in terms of how to publish a paper based on open notebook work (also described informally [here](#)).*

We added a citation to the LabTrove paper in the section “Implementation of an open lab notebook platform”, and to the Open Source Malaria paper just above the “Future directions and conclusion” section.

*10) **Examples of Inputs.** Are there any examples the authors can point to where the open ELN has helped the research, e.g. where suggestions have been made, and acted upon, to help the science?*

Yes. Examples include researchers from a pharmaceutical company that contacted a scientist to further discuss his results, a big biotech company that kindly asked permission to include data from an open notebook at a conference talk, and academics interested in specific chemical templates included in an open report.

Rachel Harding’s open notebook has had input from a number of readers including Professor Ray Truant (McMaster University) and Professor Jeff Carroll (Western Washington University), both of whom have initiated discussions about the research questions Dr. Harding is pursuing and ultimately lead to collaborative experimental design and analysis of data.

*11) **Examples of ELNs.** Are there other examples of open lab notebooks being used in biomedical research? Is this, in fact, still a highly niche activity?*

Open lab notebooks are far from mainstream and the number of active examples remains small. However, we enjoy following contributors of the Open Notebook Science Network <http://onsnetwork.org/> amongst others.

*12) **Language.** The language barrier is mentioned, but we're not sure this is relevant to this paper. Language is already a barrier in the current system, though English is essentially a lingua franca across science.*

We agree, but we have examples of scientists who do not feel confident enough in English to use our platform.

*13) **Licence.** What is the license that covers the authors' i) blog posts and ii) notebooks? To what extent is this an important choice? For example, can others take and re-use the content without restriction, or only for non-commercial purposes?*

We currently license all of our open notebook blog posts and Zenodo deposits using a Creative Commons Attribution license. This allows our work to be shared and reused as widely as possible whilst asking for attribution of the original work only. More complex licenses may put off readers of the notebooks from reusing the content, which is in opposition of the open lab notebook ethos.

*14) **Raw Data.** The authors mention that the ELN entries include "all data". This is an important feature of open ELNs that distinguishes the practice from a great deal of open science in which highlights might be discussed, or blogged about, but without the attendant raw data. We investigated briefly to see how much of the raw data could be accessed in the authors' cases. We have noticed that a fair number of the lab notebook entries recorded by the team in Zenodo*

contains PDF/Word-style summaries of data, rather than the data themselves. This limits re-use of the data. Could the authors comment on the central importance of the availability of all raw data?

Examples:

- [This](#) blog post mentions performing synthesis of analogues based on lit procedures but there's no link to an ELN.
- Same with [this](#) page.
- And [this](#).
- [This](#) page has a link (in the 2nd paragraph) to a (structure?) data file but no backlink from the data back to the blog page, risking orphaning of content?
- [This](#) blog page does have a link to the ELN. There is more experimental detail, but there is no raw characterization data. Also the [backlink](#) to the blog page is broken.
- Many biology entries contain pictures of purifications and for the most part, the ELN pages only have a docx or pdf there. Should there be any raw data available for these kinds of pages, or are docx and PDF enough?

We agree that all raw data should be included in open notebook reports, and we encourage scientists using openlabnotebooks.org to adhere to this principle. We are also aware that enforcing virtuous habits can put-off new users, and are opting for a more permissive approach, where we prefer to educate in best practices rather than imposing them.

15) **Identifiers.** It is mentioned that is desirable to have "the ability to search for experiments containing compounds with specific chemical templates". We agree. Some ELNs already allow this (e.g. Chemotion <sup>3</sup> and C6H6 <sup>4</sup>). We have found a stop-gap is the manual inclusion of chemical strings such as SMILES, InChI. Can the authors comment on whether similar things can be included in order to allow machines to understand the biology contained within the entries, e.g. UniProt numbers?

Algorithm could easily be integrated in an notebook platform to interpret UniProt numbers. For machine to understand the biology associated with a UniProt number, if there is such a thing, is beyond our expertise.

16) **Metrics re Access.** Is there a way of distilling out access by non-team people, or even access by the author of the page who may have accessed multiple times during the editing of a page?

The access metrics are directly generated by Zenodo, provides the number of visits, downloads, unique visits and unique downloads. Even if the author revisits or downloads multiple times his or her record, it would count as a single "unique event".

17) **Ownership.** When a researcher moves on, does that researcher have any responsibility towards the ELN? e.g. if a mistake is found, who needs to correct it? Is it the responsibility of the PI to act as future curator?

All records posted at openlabnotebooks.org and Zenodo remain indefinitely accessible. Our position is that the labnotebook is owned by the scientist who wrote it.

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

## Comments on this article

Version 1

Reader Comment 22 Jan 2019

**Muhammad Shahzad Aslam**, Xiamen University Malaysia, Malaysia

Informative paper and we'll written

***Competing Interests:*** Nill

---

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**