Reproducibility through Open Protocols: Opportunities for Libraries

ACRL Choice webinar
April 18, 2024

Anneliese Taylor, UCSF
Lenny Teytelman, protocols.io
What is reproducibility?

*Reproducibility* is obtaining consistent results using the same input data; computational steps, methods, and code; and conditions of analysis.

*Replicability* is obtaining consistent results across studies aimed at answering the same scientific question, each of which has obtained its own data.

From *Reproducibility and Replicability in Science*, National Academies Press, 2019
Reproducibility issues

70%

of Scientists said they had failed to reproduce experiments*

*Baker, M. 1,500 scientists lift the lid on reproducibility. Nature 533, 452–454 (2016). https://doi.org/10.1038/533452a
What are protocols?

- Detailed instructions for how research is conducted
- Methods vs. protocols
  - Method: the approach used to examine a research question
- The terms are often used interchangeably
Library support for protocol resources

Methods and Protocols Resources

Free and UCSF-subscribed methods & protocol resources are listed below. See how to get access to subscription resources through the UCSF Library.

1. Bio-protocol — a free, peer-reviewed life sciences protocol journal founded by Stanford scientists
2. Cold Spring Harbor Protocols (subscription) — an interdisciplinary, peer-reviewed biomedical research
   techniques journal
   techniques and practical overviews in over a dozen topical areas
4. JMIR Research Protocols — an open access, peer-reviewed journal of research ideas, grant proposals, and
   study and trial protocols of health-related research and technology innovations
5. JoVE (Journal of Visualized Experiments) (subscription) — peer-reviewed, experimental techniques published
   as videos with accompanying text
6. Methods in Enzymology (subscription) — long running peer-reviewed publication of research methods in
   biochemistry and biotechnology
7. Protocol Exchange — open and free protocol sharing site with 1000 contributions, many with videos.
   Managed by Nature Protocols
8. protocols.io - an open repository with thousands of publicly shared protocols with interactive features, as
   well as a platform for running and customizing protocols. UC's premium account allows unlimited numbers
   of private protocols and groups. LabArchives electronic lab notebook users can also synchronize protocols
   between protocols.io and a LabArchives notebook.
9. SAGE Research Methods (subscription) - more than 1000 online books, reference works, journal articles, and
   instructional videos by academics from across the social sciences, including a large collection of qualitative
   methods books
10. Science of Synthesis (subscription) — critical review of the synthetic methodology developed from the early
    1800s to date for the entire field of organic and organometallic chemistry
11. Springer Nature Experiments (UC subscription) — over 75,000 protocols and methods from across the life
    Primers Methods. Filter results by technique, antibody, organism, and more

Resources for designing protocols

1. Clinical Trial Protocol Development from UCSF's Clinical Research Resource HUB offers several protocol templates and related
   resources. Cancer Center-specific templates and resources are also available.
2. ClinicalTrials.gov Protocol Registration & Results System (PRS) or another appropriate registry is required for publication of clinical trial
   results in many journals. See definitions of protocol data elements, templates, and tutorials.
3. EQUATOR Network provides reporting guidelines for all study types
4. How to write an easily reproducible research protocol from American Journal Experts
5. How to make your protocol more reproducible, discoverable, and user-friendly from protocols.io
6. Systematic review protocol resources from Cornell University Library

https://tinyurl.com/methodsprotocols
Protocols & Their Role in Open Science
“Closed” science workflow
“Open science is the practice of science in such a way that others can collaborate and contribute, where research data, lab notes, and other research processes are freely available, under terms that enable reuse, redistribution, and reproduction of the research and its underlying data and methods.”

Foster Open Science
Open Science Workflow
UCSF Library & Open Science
The University of California system

- UC libraries
  - Governance & advisory structure
  - Content licensing
  - Discovery & preservation

- California Digital Library (CDL)
  - Part of UC Office of the President
  - UC Curation Center (UC3)
A bit about UC San Francisco (UCSF)

▷ One of the 10 University of California (UC) campuses
▷ Exclusively health sciences
▷ 3200 graduate & professional students
▷ 1200 postdocs; 5K academic employees
▷ Top public recipient of NIH funding: 1484 awards, $790M
▷ 8000+ research articles annually
UCSF Open Science Group

- Established by the Library and a postdoctoral scholar
- Members include students, faculty, researchers & staff
  - Including Associate Dean for Graduate Programs, Liz Silva
- Other Library personnel:
  - Data Services Lead, Ariel Deardorff
  - Data Science Lead, Karla Lindquist

Ibraheem Ali, Anneliese Taylor, Ariel Deardorff, Karla Lindquist, and Liz Silva
Protocol sharing event - 2018

https://calendars.library.ucsf.edu/event/4218010
Reproducibility workshops - 2019 & 2021

- Library, Graduate Division, and Open Science Group
- NIH’s Responsible Conduct for Research (RCR) training credit
- Before and after assessment

Shared materials
- Chapter in The Scholarly Communications Cookbook
  escholarship.org/uc/item/6gr2d33k
- Course materials publicly accessible: 2019 and 2021
Reproducibility Workshop Series: Reproducible Methods & Protocols

One of the barriers to reproducibility of research is the lack of detailed methods in the research environment and in published articles. This workshop will examine resources and tools to simplify and improve documentation throughout the research lifecycle and to ensure that methods and protocols are findable and reproducible. Please register below.

Learning Objectives
By the end of this workshop, learners will be able to:

- Summarize tools and best practices for documenting research methods
- Identify resources for finding existing methods and protocols
- Define the benefits of publishing open research protocols

Instructors & Presenters: Ibrahim Ali and Stephen Floor

Ibrahim received his PhD in Biomedical Sciences at UCSF and did a brief post-doc focusing on epigenetics and transcription regulation at the Gladstone Institutes. In his current role as Sciences Data Librarian at UCLA, he teaches data visualization strategies and best-practices for preprints and methods publishing to pre-clinical and clinical researchers.

Stephen Floor, PhD, is Assistant Professor in the Department of Cell & Tissue Biology. The Floor Lab studies the mechanisms by which RNA impacts human gene expression in health and disease. RNA chaperones are a major focus of the lab, which remodel RNA structures and RNA-protein interactions. We use deep sequencing, molecular biology of purified proteins, and imaging approaches to define the in vitro and cellular mechanisms of RNA chaperones. Many DEAD-box RNA chaperones are genetically altered in cancers and other diseases. We study the molecular basis of such alterations using cancer cell lines and genetically engineered stem cell models, aiming to better understand human biology and nominate new therapeutic targets or treatment regimens.
Assessment of workshop series
Credit-based reproducibility course - Fall 2022

Classes:

1. Reproducibility overview
2. Data & documentation
3. Methods, protocols & code
4. Data sharing & repositories
5. Publishing & open access
6. Authorship & peer review
7. Tracking and getting credit for your work

Shared materials at courses.ucsf.edu/course/view.php?id=9717
Collaborative open science efforts

Bay Area Open Science Group
A virtual community for Stanford, UCSF, & Berkeley

bayareaopensciencegroup.github.io

UC Love Data Week
February 12 - 16, 2024

UC Love Data Week is a week-long offering of presentations and workshops focused on data access, management, security, sharing, and preservation. Whether you're working on qualitative or quantitative data, we've got events for you! All members of the University of California community are welcome to attend.

Explore this year's presentations and workshops by using the calendar or workshop info views below. For full descriptions and to register for a specific workshop, click on the workshop title. Make sure to register with your UC campus email.

Tip for a successful UC Love Data Week Experience: Check your junk email folder for accidentally misdirected workshop confirmations and Zoom information!

uc-love-data-week.github.io
Support for open protocols at UC

The University of California Curation Center (UC3) helps researchers and the UC Libraries manage, preserve, and provide access to their important digital assets.

- **DMTTool**: Helps researchers create and manage data management plans.
- **DRYAD**: Members of DRYAD are institutional members of DRYAD, a self-service data publication platform for researchers to describe, upload, and share their research data.
- **EZID**: EZID (easy-identifier) makes it easy to create and manage long-term, globally unique identifiers for your data and sources, ensuring their future discoverability.

**UC-wide pilot of protocols.io**

By John Ghodacki / May 31, 2019

**UNIVERSITY OF CALIFORNIA**

**protocols.io**

Merritt is a repository service that lets you manage, archive, and share valuable digital content.
Example protocol on protocols.io

https://doi.org/10.17504/protocols.io.4pagvie
Protocol metrics on protocol.io
### UC San Francisco

#### Quarterly protocols.io Summary

<table>
<thead>
<tr>
<th>Year</th>
<th>Q1</th>
<th>Q2</th>
<th>Q3</th>
<th>Q4</th>
<th>Growth</th>
<th>Site Growth</th>
</tr>
</thead>
<tbody>
<tr>
<td>Users</td>
<td>655</td>
<td>672</td>
<td>694</td>
<td>709</td>
<td>11%</td>
<td>18%</td>
</tr>
<tr>
<td>Private protocols</td>
<td>676</td>
<td>687</td>
<td>703</td>
<td>712</td>
<td>9%</td>
<td>28%</td>
</tr>
<tr>
<td>Public protocols</td>
<td>143</td>
<td>151</td>
<td>154</td>
<td>157</td>
<td>16%</td>
<td>23%</td>
</tr>
<tr>
<td>Users</td>
<td>96</td>
<td>151</td>
<td>218</td>
<td>267</td>
<td>251%</td>
<td>72%</td>
</tr>
<tr>
<td>Private protocols</td>
<td>101</td>
<td>132</td>
<td>240</td>
<td>299</td>
<td>240%</td>
<td>73%</td>
</tr>
<tr>
<td>Public protocols</td>
<td>39</td>
<td>49</td>
<td>52</td>
<td>58</td>
<td>49%</td>
<td>59%</td>
</tr>
</tbody>
</table>
Additional resources

- Open Science 101 (UCSF Library)
- 10 Ways to Support Open Research (UCSF Library)
- Open Science Team Agreement (Bay Area Open Science Group)
- Reproducibility of Research (Univ of Utah, Eccles Health Sciences Library)

Pre-registered research studies
- AsPredicted (Wharton Credibility Lab, U Penn)
- OSF Registries (Center for Open Science)
protocols.io perspective
Current directions in academic libraries

1. Continue the migration from print to electronic and realign service operations
2. Review location of lesser-used collections
3. Continue to repurpose library as primary learning space
4. Reposition library expertise and resources to be more closely embedded in research and teaching enterprise outside library
5. Extend focus of collection development from external purchase to local curation

(2015 NFAIS Conference, slideshare.net/cmkeithw/nfais-web-version)
Impact of the UCSF protocols.io site license

Public and private protocols per user

- UCSF
- all protocols.io

<table>
<thead>
<tr>
<th>Protocols per user</th>
<th>private protocols</th>
<th>public protocols</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>UCSF</td>
<td>all protocols.io</td>
</tr>
<tr>
<td>0.00</td>
<td>0.25</td>
<td>0.25</td>
</tr>
<tr>
<td>0.25</td>
<td>0.50</td>
<td>0.25</td>
</tr>
<tr>
<td>0.50</td>
<td>0.75</td>
<td>0.25</td>
</tr>
<tr>
<td>0.75</td>
<td>1.00</td>
<td>0.25</td>
</tr>
<tr>
<td>1.00</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Impact of the UC protocols.io site license

2019-2022: UC researchers are 2.64% of the total registered users, but they have 5.31% of the total private protocols. Just as importantly, public UC-affiliated protocols have increased from 2%-3% of the total public protocols in 2019 to 8% in 2022.

Across all of protocols.io: Premium users are more than twice as likely to share protocols publicly, and over time, share five times more protocols than free users.
Reproducibility courses (OA, of course!)

https://www.cos.io/services/training
Reproducibility courses (OA, of course!)

https://www.repro4everyone.org/
Reproducibility courses (OA, of course!)

The power of sharing detailed methods: credit, preservation, and reproducibility

Open and Reproducible Practices for Experimental Research 2023

Gabriel Gasque, PhD
Head of Outreach
gabriel@protocols.io

September 7th, 2023

Webinar

The Art of Writing Good Methods

Emma Ganley, PhD
emma@protocols.io

https://www.protocols.io/webinars
Questions?

Contact us:
anneliese.taylor@ucsf.edu
lenny@protocols.io
Supplemental Slides
2021 Reproducibility series

- Library & Graduate Division
- Fewer guest presenters
- Before and after assessment
- Class materials shared publicly courses.ucsf.edu