# Rigor and Reproducibility: an overview

New Faculty Orientation

**August 17, 2018**

Jane Strasser, Ph.D.
Associate Vice President for Research
Research Integrity Officer
Jane.strasser@uc.edu
558-5034

http://researchcompliance.uc.edu/Home.aspx

University of CINCINNATI

A survey by Nature revealed that **52% of researchers believe that there is a 'significant reproducibility crisis'** and 38% believe there is a 'slight crisis'.

**HOW DO WE FIX IT?**

Nearly 90% (over 1,000 respondents) indicated:
More robust experimental design
Better statistics
Better mentorship

**so why haven't we done it?**

Money, time, reward system….

Nature. 2016. 533:452-454.

University of CINCINNATI

Scientists, policy makers, and journalists should use precisely defined terms and definitions when discussing research rigor and transparency to promote uniform understanding.

**a. Replicability:** the ability to repeat a prior result using the same source materials and methodologies. This term should only be used when referring to repeating the results of a specific experiment rather than an entire study

- technical replicates- repeated measures of the same sample
- sample (biological, chemical, environmental, etc.) replicates- from parallel measurements of distinct samples to capture random variation which may be a source of noise

**b. Reproducibility:** the ability to achieve similar or nearly identical results using comparable materials and methodologies. This term may be used when specific findings from a study are obtained by an independent group of researchers

**c. Generalizability:** the ability to apply a specific result or finding more broadly across settings, systems, or other conditions

**d. Translatability:** the ability to apply research discoveries from experimental models to human health applications

**e. Rigor:** the use of unbiased and stringent methodologies to analyze, interpret, and report experimental findings

**f. Transparency:** the reporting of experimental materials and methods in a manner that provides enough information for others to independently assess and/or reproduce experimental findings

https://www.faseb.org/Portals/2/PDFs/opa/2016/FASEB_Enhancing%20Research%20Reproducibility.pdf

University of CINCINNATI

## What are other issues that affect the quality / rigor of science?

Rigor, the strict application of the scientific method to ensure robust and unbiased experimental design, methodology, analysis, interpretation, and reporting of results.

1. Insufficient, inappropriate controls
2. Lack of investigator blinding, sample randomization
3. Improper statistical analysis
4. Bad reagents/input (chemicals, antibodies, poor survey collection, etc.)

Scientific Utopia II. Restructuring Incentives and Practices to Promote Truth Over Publishability doi: 10.1177/1745691612459058

University of CINCINNATI

## What are other issues that affect the quality / rigor of science?

Approaches

(a) leveraging chance by running many low-powered studies, rather than a few high-powered ones;
(b) uncritically dismissing "failed" studies as pilot tests or because of methodological flaws but uncritically accepting "successful" studies as methodologically sound;
(c) selectively reporting studies with positive or "clean" results and not studies with negative results;
(d) stopping data collection as soon as a reliable effect is obtained;
(e) continuing data collection until a reliable effect is obtained;
(f) including multiple independent or dependent variables and reporting the subset that "worked"
(g) maintaining flexibility in design and analytic models, including the attempt of a variety of data exclusion or transformation methods, and reporting a subset
(h) reporting a discovery as if it had been the result of a confirmatory test
(i) once a reliable effect is obtained, not doing a direct replication

**Scientific Utopia II. Restructuring Incentives and Practices to Promote Truth Over Publishability**
doi: 10.1177/1745691612459058

University of CINCINNATI

# Rigorous Experimental Design

Scientific rigor is the strict application of the scientific method to ensure **robust and unbiased experimental design, methodology, analysis, interpretation and reporting** of results. This **includes full transparency in reporting** experimental details so that others may reproduce and extend the findings. NIH expects applicants to describe how they will achieve robust and unbiased results when describing the experimental design and proposed methods. Robust results are obtained using methods designed to avoid bias and can be reproduced under well-controlled and reported experimental conditions.

- Use of Standards
- Sample size estimation (power analysis, justification)
- Randomization
- Blinding
- Appropriate replicates
- Controlling for inter-operator variability
- Statistical methods planned
- Inclusion and exclusion criteria
- Subject retention and attrition
- Plan to handle missing data
- Other

https://www.nih.gov/research-training/rigor-reproducibility

University of CINCINNATI

**Consideration of Sex and Other Relevant Biological Variables** **Cost also is not a consideration** in determining whether both sexes are to be included in experiments.
NIH expects that sex as a biological **variable will be factored into research designs, analyses, and reporting** in vertebrate animal and human studies. Appropriate analysis and transparent reporting of data by sex may enhance the rigor, transparency, and applicability of preclinical biomedical research. Strong justification from the scientific literature, preliminary data or other relevant considerations must be provided for applications proposing to study only one sex. Please refer to NOT-OD-15-102 for further consideration of NIH expectations about sex as a biological variable.
Similarly, investigators should consider other biological variables, as appropriate, in the design and analyses of their proposed studies. Research plans and findings should clearly indicate which biological variables are tested or controlled. Clear justification should be provided for exclusion of variables that may be relevant but are not considered in the research plan. For example, studies using young adult animals should clearly describe their study population and not generalize findings to juvenile or aged animals.

**Authentication of Key Biological and/or Chemical Resources**

https://www.nih.gov/research-training/rigor-reproducibility

University of CINCINNATI

**Table 1. A Primer of Best Practices to Enhance Rigor and Reproducibility**

| Topic | Best Practice | Benefits |
|---|---|---|
| Experimental Design | Describe experiment planning in manuscript Methods section, including:<br>• Power calculations (endpoint sensitivity, variability, effect size, desired level of confidence, definition and rationale for n).<br>• Inclusion/exclusion of data sets, description of pilot, and final data sets included in analyses.<br>• Random assignment to treatment groups, description of exceptions.<br>• Procedures to achieve blinding, exceptions to blinding, and resulting interpretive caveats.<br>• Details of reagents and assays sufficient to facilitate independent replication.<br>• Positive and negative controls. | Capture thinking in incomplete information landscape.<br>Iterative hypothesis refinement.<br>Deep understanding of assessments in advance of execution.<br>Reduce testing to foregone conclusion.<br>Optimize resource allocation and use.<br>Create roadmap to assembling publication. |
| Analysis and Statistics | Describe statistical analysis plan in manuscript Methods section, including:<br>• Methods to test for significance.<br>• Interim analyses, futility assessments.<br>• Data inclusion/exclusion, attrition.<br>• Statistical treatment of technical and biological replicates.<br>• Test-retest approaches.<br>• Statement of central tendency, variance, statistical test, and p value for significant and nonsignificant differences.<br>• Descriptive statistics for groups as well as pooled values. | Enhance awareness of and reduce sources of potential unconscious bias.<br>Minimize type 1 error. |

Steward and Balice-Gordon. Neuron (2014) 84:572-581. Rigor or Mortis: Best practices for preclinical research in neuroscience.

University of CINCINNATI

**Table 1. A Primer of Best Practices to Enhance Rigor and Reproducibility**

| Topic | Best Practice | Benefits |
|---|---|---|
| Data Management | Develop lab standards for indexing and maintaining information, including:<br>• Recording of key experimental design and execution parameters.<br>• Archiving raw data and at least one backup with appropriate frequency.<br>• Curation of process from raw data to summary figure to conclusion. | Ensure all information supporting a conclusion can be located during and after study completion. |
| Resource Sharing | Include lists of resources in manuscripts that will be made available and point of contact for requests.<br>Indicate time limit for resource availability, if any.<br>Include budget line item to support resource sharing in funding applications.<br>Deposit animal lines at commercial vendor within 3 months of publication.<br>Provide raw data upon request. | Simplify sharing of reagents, protocols, raw data to facilitate replication, interpretation of data.<br>Help distinguish lack of conceptual validation versus lack of replication.<br>Enable meta-analyses and data basing. |
| Publication and Reporting | Provide comprehensive review checklist for methodology, reagents, and resource sharing.<br>Two-stage review: if manuscript meets general journal criteria (novelty, impact, general interest), initiate second stage of review for technical merit including details relating to rigor. | Raise awareness of key metrics for determining rigor.<br>Facilitate replication of key findings. |

Steward and Balice-Gordon. Neuron (2014) 84:572-581. Rigor or Mortis: Best practices for preclinical research in neuroscience.

University of CINCINNATI

## Things to keep in mind:

- Is the work properly powered to get significant results (how was sample size determined and why)?
- Is there a rigorous (and appropriate) statistical analysis?
- Are the appropriate controls included?
- Have you minimized the likelihood of bias (e.g. blinding)?
- Do you have the correct (and validated) reagents, model systems, and tools?
- What are you inclusion/exclusion criteria (should be stated in publications)?
- Are the published methods clear and sufficiently detailed to facilitate replication?
- Are all data and reagents available upon request (or provided with publication?)?

https://www.nih.gov/research-training/rigor-reproducibility

University of CINCINNATI

## References:


**NIH Rigor and Transparency** (extensive resource)
https://www.nih.gov/research-training/rigor-reproducibility

**Enhancing Research Reproducibility: Recommendations from the Federation of American Societies for Experimental Biology** Effective January 14, 2016
https://www.faseb.org/Portals/2/PDFs/opa/2016/FASEB_Enhancing%20Research%20Reproducibility.pdf

**Scientific Utopia II. Restructuring Incentives and Practices to Promote Truth Over Publishability** Brian A. Nosek, Jeffrey R. Spies and Matt Motyl
doi: 10.1177/1745691612459058
Perspectives on Psychological Science 2012 vol. 7 no. 6 615-631

**Rigor or Mortis: Best practices for preclinical research in neuroscience**
Oswald Steward and Rita Balice-Gordon.
http://dx.doi.org/10.1016/j.neuron.2014.10.042
Neuron. 2014. 84:572-581


University of CINCINNATI

## Data Challenges:

- Data are varied and may be of poor quality
- E-Data may have format, version, and access control issues
- Security and availability
- Completeness and robustness
- Source/original data access, transparency, audit trail
- Ownership/retention
- Resources to do the above

University of CINCINNATI

## Data Planning Checklist:

Managing your data before you begin your research and throughout its life cycle is essential to ensure its current usability and long-run preservation and access.
To do so, begin with a planning process.

1. What type of data will be produced? Will it be reproducible? What would happen if it got lost or became unusable later?
2. How much data will it be, and at what growth rate? How often will it change?
3. Who will use it now, and later?
4. Who controls it (PI, student, lab, UC, funder)?
5. How long should it be retained? e.g. 3-5 years, 10-20 years, permanently
6. Are there tools or software needed to create/process/visualize the data?
7. Any special privacy or security requirements? e.g., personal data, high-security data
8. Any sharing requirements? e.g., funder data sharing policy
9. Any other funder requirements? e.g., data management plan in proposal
10. Is there good project and data documentation?
11. What directory and file naming convention will be used?
12. What project and data identifiers will be assigned?
13. What file formats? Are they long-lived?
14. Storage and backup strategy?
15. When will I publish it and where?
16. Is there an ontology or other community standard for data sharing/integration?
17. Who in the research group will be responsible for data management?

http://libraries.mit.edu/guides/subjects/data-management/checklist.html
http://guides.libraries.uc.edu/c.php?g=222496&p=1472563

University of CINCINNATI

## Documentation should include:

1. Who generated the record
2. What they did
3. When they did it
4. Why the did it
5. What the overall goal/project was
6. How they did it (protocol/methodology)
7. What materials were used
8. The results
9. The analysis
10. The interpretation
11. The next step(s)

Remember to check for data entry errors
Can you audit (are changes in the database saved so you can identify if/when an error occurred?/do you have version controls?)?

University of CINCINNATI

Modified from "Guidelines for Scientific Record Keeping in the Intramural Research Program at NIH" Michael Gottesman, MD. 2008

## The research record includes:

- Lab notes, spreadsheets, databases
- Equipment/access logs, etc.
- Posters
- Seminars
- Funding proposals
- Progress reports
- Manuscripts

ARCHIVE: final raw data set, documented program that prepared the data set, documented program that conducted the analysis, output from the program (the analysis)

University of CINCINNATI

## Why do we care about scientific records?

1. Good records are essential for data analysis, publication, collaboration, peer review, and other research activities

2. Record keeping/record retention is required by funding agencies, by the state of Ohio, and by UC. UC Rule 10-43-18 mandates that all scientific records be maintained for at least 5 years

3. Necessary to support intellectual property claims

- Failure to retain records is considered evidence of misconduct

**Without the data it isn't science, it's science fiction**

University of CINCINNATI

Modified from "Guidelines for Scientific Record Keeping in the Intramural Research Program at NIH" Michael Gottesman, MD. 2008

Documentation should be:

1. Reasonably permanent
   - Paper (organized!)
   - Electronic (with back up)

2. Appropriately secured

3. Meet the FAIR standard
   - Findable
   - Accessible
   - Interoperable
   - Reusable

Good mentoring including consistent review of raw data reduces the likelihood of misconduct

University of CINCINNATI

**BOX 9-1**
**Best Practices Checklist for Researchers**

Research Integrity

- Maintain high standards in own work.
- Understand policies.
- Raise questions and problems promptly and professionally.
- Strive to be a generous and collegial colleague.

Data Handling

- Develop data management and sharing plan at the outset of a project.
- Incorporate appropriate data management expertise in the project team.
- Understand and follow data collection, management, and sharing standards, policies, and regulations of the discipline, institution, funder, journal, and relevant government agencies.

Authorship and Communication

- Ensure that general and disciplinary standards are followed for research publications.
- Acknowledge the roles and contributions of authors.
- Be transparent when communicating with all audiences.

Mentoring and Supervision

- Model and instruct on research best practices.
- Regularly check work of subordinates and ensure adherence to best practices.
- Clarify expectations.

Peer Review

- Provide complete and timely review.
- Maintain confidentiality.
- Disclose conflicts, and eliminate or manage them as appropriate.

Research Compliance

- Protect human subjects and laboratory animals.
- Follow environmental and other safety regulations.
- Do not engage in misuse.
- Disclose and manage conflicts of interest.

University of CINCINNATI

Fostering Integrity. National Academies of Sciences, Engineering, and Medicine. 2017. www.nap.edu

**BOX 9-3**

**Best Practices Checklist for Journals**

Practicing Transparency

- Adopt up-to-date policies and instructions.
- Publish retractions/corrections and reasons in articles, in tables of contents, and as metadata in a timely fashion.
- Provide a link to data and code that support articles, and facilitate long-term access.
- Require full descriptions of methods in methods sections or electronic supplements.
- Provide for postpublication review and commentary.
- Be transparent in negotiating with authors and in adjudicating disputes.
- Establish a conflict of interest policy covering editorial staff.
- Provide open access consistent with business viability.

Adopt Policies that Ensure Openness Regarding:

- Data, code, and records of any image alterations.
- Author funding and conflicts of interest.
- Peer reviewer conflicts of interest.

Author Contributions

- Describe author roles.

Training and Education

- Facilitate training for editors, reviewers, and authors.

Collaboration

- Participate in science, engineering, technology, and medical publishing efforts to develop tools and approaches to foster integrity.

University of CINCINNATI

Fostering Integrity. National Academies of Sciences, Engineering, and Medicine. 2017. www.nap.edu

**Box 9-4: Best Practices Checklist for Research Sponsors and Users of Research**

Aligning Policies with Research Integrity

- Maintain clear policies on research misconduct, and implement them consistently.
- Increase awareness of how policies and practices affect research integrity and quality, and act on that knowledge.
- Work to harmonize policies and practices across agencies, sectors, and national borders.

Public Access to Data and Code

- Develop data and code access policies for extramural grants appropriate to the research being funded, and make fulfillment of these policies a condition of future funding.
- Cover the costs borne by researchers and institutions to make data and code available.
- Practice transparency of data and code for intramural programs.
- Promote responsible sharing of data in areas such as clinical trials.
- Practice impartiality and transparency in utilizing research for the development of policy and regulations.

University of CINCINNATI

Fostering Integrity. National Academies of Sciences, Engineering, and Medicine. 2017. www.nap.edu

**Research Misconduct** applies to the research record and is specific to:

- Fabrication
- Falsification
- Plagiarism
- Serious deviations from discipline norms (sabotage, malicious allegations of misconduct)

Research Misconduct does NOT include honest error, nor does it cover authorship disputes.

Misconduct can be reckless, knowing, and/or intentional.

University of CINCINNATI

If you believe research misconduct may have occurred:

**PROTECT THE EVIDENCE** -Bring the RIO in ASAP

1. Research Integrity Officer (RIO) does initial assessment (specific and credible)

2. RIO notifies accused (this includes notice of article 9 for AAUP) and sequesters all data

3. RIO & College Dean identify panel of un-conflicted experts to perform inquiry (initial review to determine if meets the definition). If merits moving forward RIO must notify federal funders

4. RIO & College Dean identify panel of un-conflicted experts to perform full investigation (findings reported to Dean, Provost, VP for Research, President, and funding agencies). RIO works with journals (and authors if appropriate) to correct research record.

http://researchcompliance.uc.edu/Home/ResearchIntegrity.aspx

University of CINCINNATI

## Additional considerations:

- We protect whistleblowers
- We coordinate with the Dean of the Graduate School when allegations involve a graduate student
- We treat allegations and investigations confidentially and protect reputations

➢ Failure to oversee can result in findings of reckless misconduct

➢ Failure to retain records is considered evidence of misconduct

University of CINCINNATI

UNIVERSITY OF Cincinnati

UC Libraries



# Data Management Planning

Search Submit Query

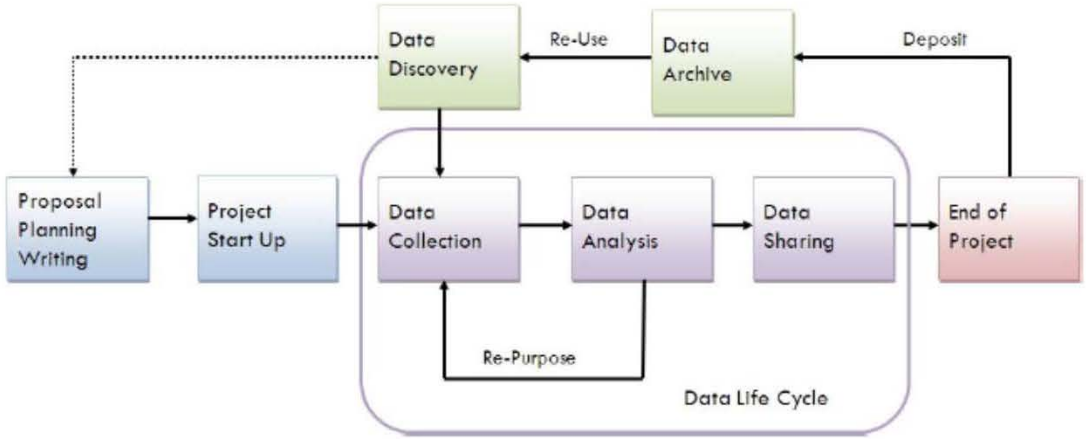
This guide on data management planning and data discovery focuses on STEM fields.



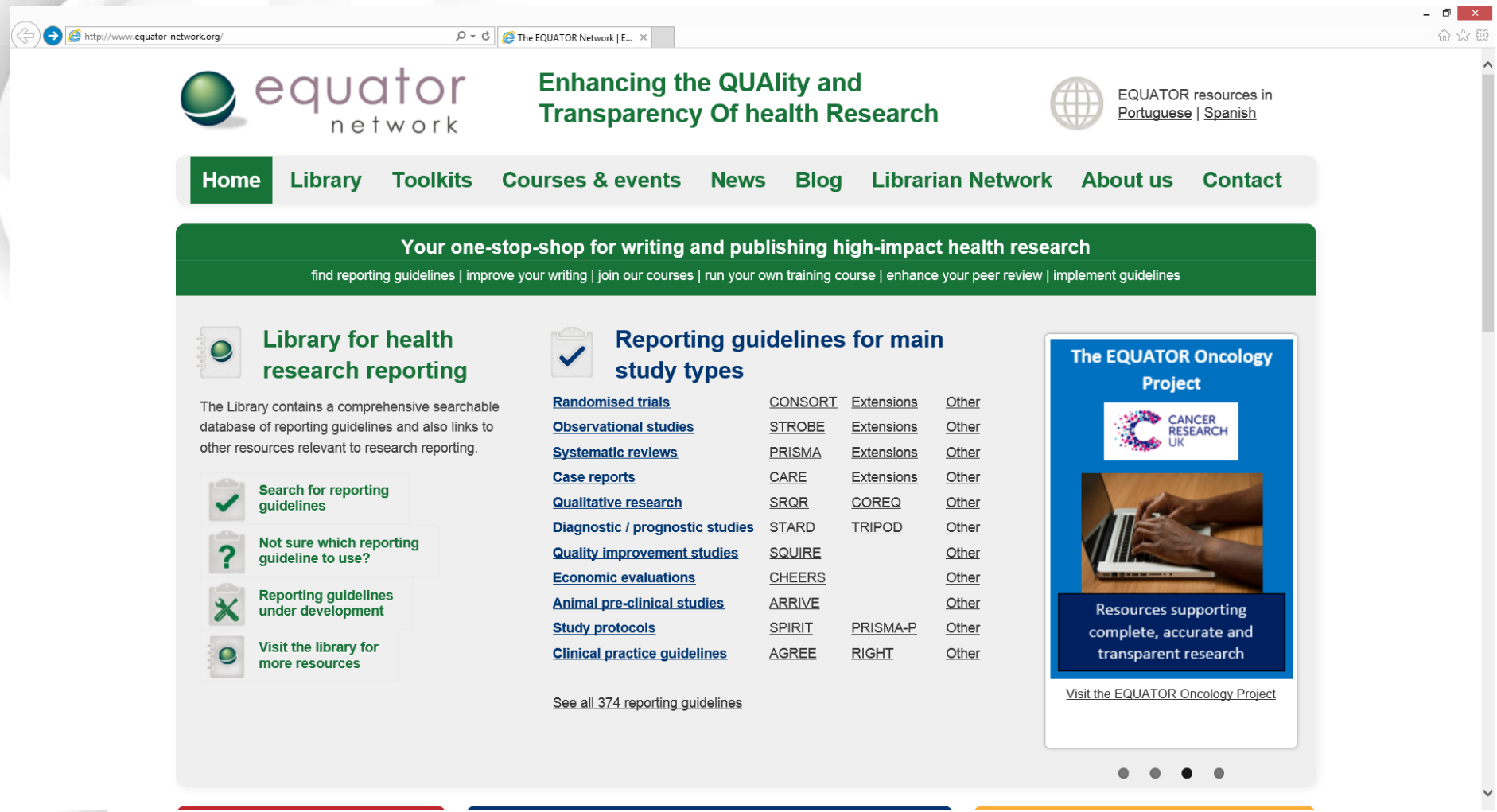
## Welcome!

Welcome to UCL's resource guide for Data Management Planning!

## Data Management Planning Process

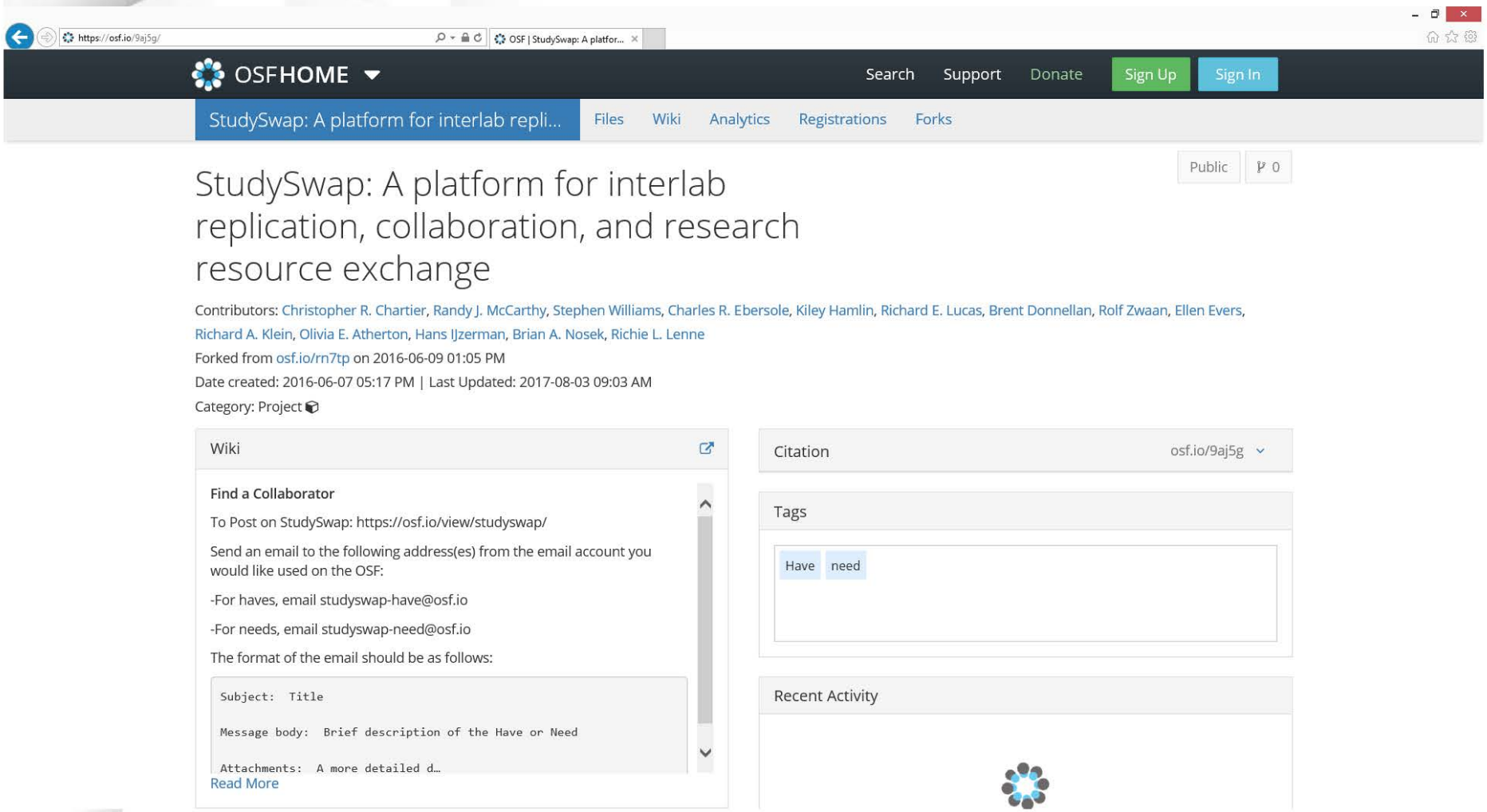



Source: http://dmconsult.library.virginia.edu/files/2013/03/DataLifeCycle1.jpg




http://www.equator-network.org/

University of CINCINNATI




University of CINCINNATI

https://osf.io/9aj5g/

# Where to go for help

**Office of Research** - if we can't help you, we can point you to someone who can http://research.uc.edu/home/officeofresearch/administrativeoffices.aspx

A few of the many other useful links:

**Evaluation Services Center** (http://www.uc.edu/evaluationservices.html )

**Center for Clinical and Translational Sciences and Training** (https://cctst.uc.edu/ )

**Libraries Research Data Services** (http://libraries.uc.edu/digital-scholarship/data-services.html )

**Office of Institutional Research**
(http://www.uc.edu/provost/about-us/peopleandoffices/institutional_research.html )

University of CINCINNATI