National Institute of Diabetes and Digestive and Kidney Diseases

Research Project Management 101

Emily Leary, PhD
Chief, Biostatistics Program Office
Office of the Director
NIDDK

National Institute of Diabetes and Digestive and Kidney Diseases

National Institute of Diabetes and Digestive and Kidney Diseases

Upcoming Sessions

August 20, 2026 from 3:00-4:30pm Eastern

Optimize Research Data Collection and Sharing

Speakers: James M Welch MGC CGC (NIDDK), Ken Wilkins PhD (NIDDK), Sungyoung Auh PhD (NIDDK)

https://bit.ly/3TXScn6

August 27, 2026 from 2:00-3:30pm Eastern

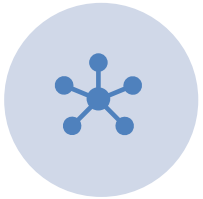
The Data Lifecycle: Leveraging Best Practices and Institutional Requirements for High-Quality Data and Reporting

Speakers: Courtney Duncan MSW LICSW (NIDDK), Emma J Stinson MPH (NIDDK), Marika Heinicke PharmD BCGP (NIDDK), Patti Young (NIDDK), Veronica Sansing-Foster PhD (NIDDK)

https://bit.ly/45WKOuP

Why is this important?

Reduces delays and rework



Improves collaboration, moving the science forward

Strengthens rigor & reproducibility

Basically, this is a small time and energy investment which pays off significantly over the project lifecycle.

And yes, this is always a good use of time!

What is research project management?

Project management is the application of processes, methods, skills, knowledge and experience to achieve specific project objectives according to the project acceptance criteria within agreed parameters.

Research project management is the process of planning, organizing, and overseeing scientific or academic studies. It includes the systemic processes for managing people, timelines, information tracking, and ensuring ethical compliance.

Research vs Traditional

ASPECT	RESEARCH PROJECT MANAGEMENT	TRADITIONAL PROJECT MANAGEMENT
◎ **Focus**	Generate knowledge & scientific understanding	Deliver a defined product, service, or outcome
↻ **Approach**	Iterative, exploratory, and adaptive	Linear, predictive, and plan-driven
? **Uncertainty**	High and expected	Minimized through planning and requirements
↗ **Scope**	Evolves with evidence and new discoveries	Defined upfront and controlled for stability
★ **Success Metrics**	Scientific validity, impact, publications, learnings	On-time, on-budget, quality, customer satisfaction
✓ **Regulatory/Compliance**	Continuous throughout the research lifecycle	Often at key phases or project close
▦ **Milestones & Deadlines**	Flexible milestones based on learning	Fixed milestones and deadlines
! **Risk Management**	Dynamic; scientific, technical, and ethical risks	Planned; schedule, cost, and resource risks

RESEARCH CREATES SIGNIFICANT CHALLENGES BEYOND TRADITIONAL PROJECT MANAGEMENT

Key Components

Iterative process

Continuously refine and adapt

No beginning or end

Ongoing and continuous

Interdependent components

Work together as a system

Goals/Scope of Work

- Outcomes / Deliverables
- SMART goals
- Milestones/timeline
- Resources
- People

SMART Objectives

S **Specific**: Is the objective sufficiently detailed to be meaningful? Can everyone understand goal?

M **Measurable**: Is the objective quantifiable? How do I know when I'm done?

A **Achievable**: Is this something that can reasonably be done?

R **Relevant**: Is this useful and realistic to ensure overall goals?

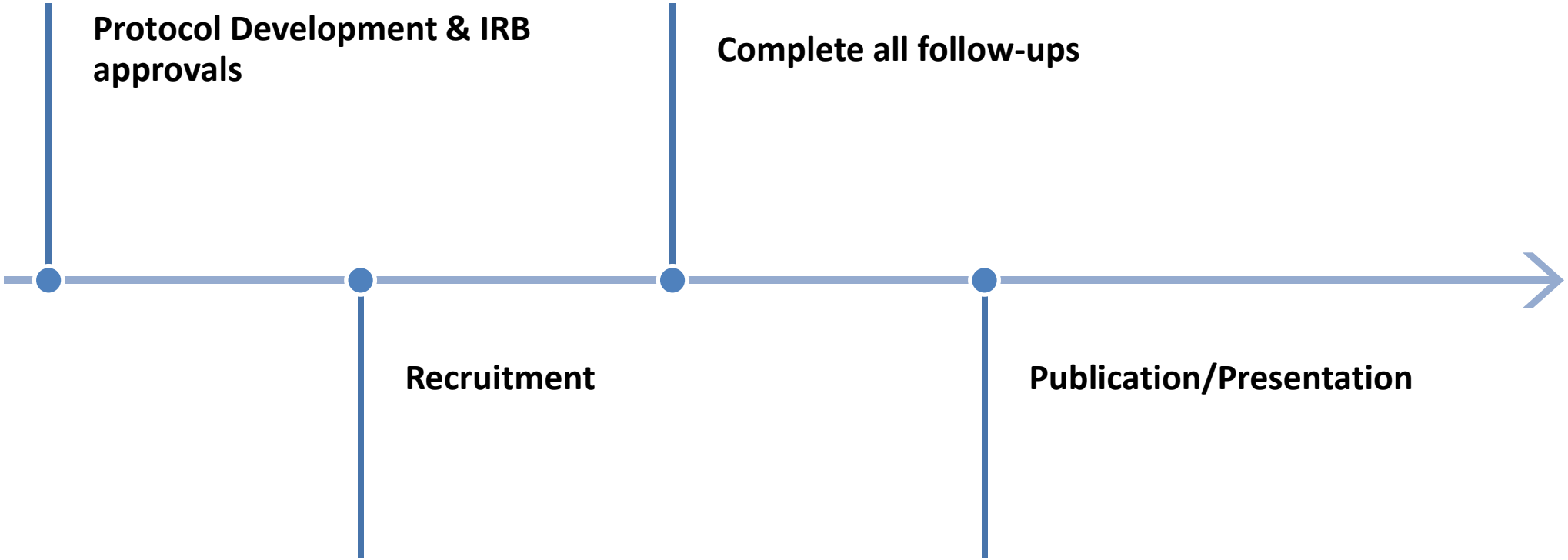
T **Timebound**: When does this need to be accomplished?

I'm going to lose weight!

I'm going to go to the gym for 60 min, 3x per week for the next 3 months!

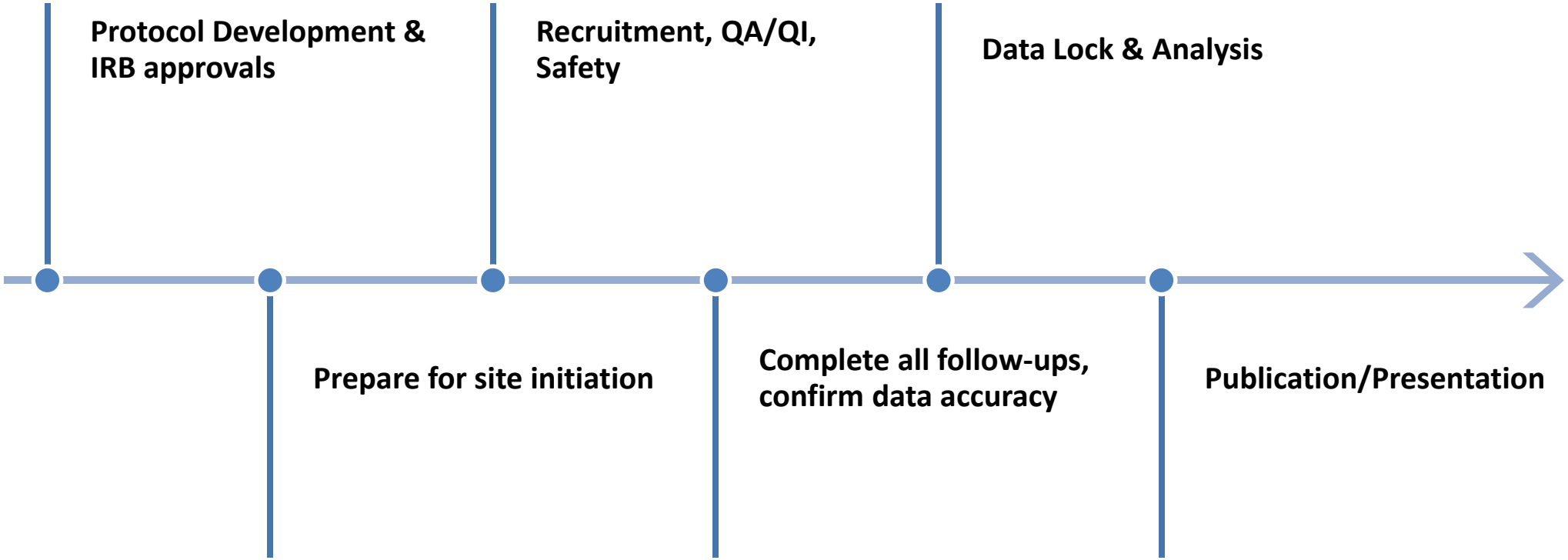
Outcomes/Milestones

What is the overall goal or objectives that all activities support?



Outcomes/Milestones

What is the overall goal or objectives that all activities support?



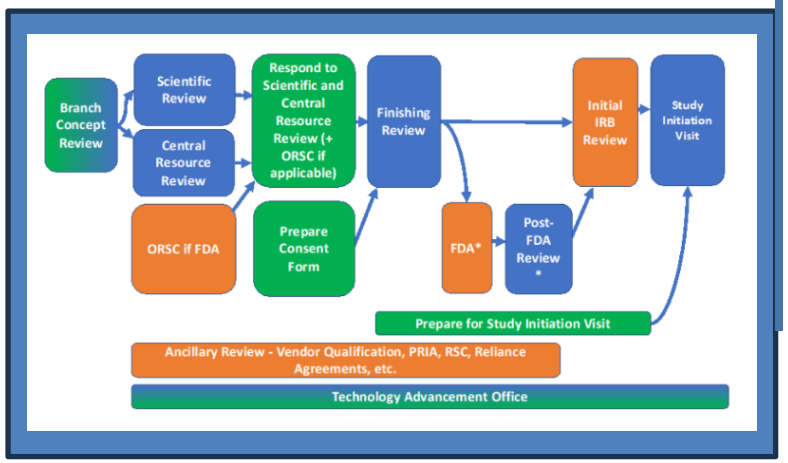
Outcomes/Milestones

What is the overall goal or objectives that all activities support?

Protocol Development & IRB approvals

Recruitment, QA/QI, Safety

Data Lock & Analysis



Prepare for site initiation

Complete all follow-ups, confirm data accuracy

Publication/Presentation

Clarifying Roles and Responsibilities

RACI matrix is a responsibility matrix that makes ownership and communication explicit

Use the four roles to quickly identify who acts, owns, advises, and stays updated.

RACI matrix – resources/experience

R	Responsible	A	Accountable	C	Consulted	I	Informed

Research Project		Collaborators					Resources/Experts					
	PI	Lead Researcher	Associate Investigators	Other Topic experts	Trainees/mentees	Radiology/imaging	Pharmacy	Clinic/in-patient	Lab/cores	Regulatory	Other tests/testing/equipme
Protocol Development											
Develop the research aims/goals	A	R	A	C	I	C	C	C	C	I	C
Write the protocol	A	R	C	C	I	C	C	C	C	C	I
Internal reviews - Branch, Resource, Scientific, etc	A/R	R	I	I	I	I	I	I	I	I	I
IRB submission	A/R	R	I	C/I	I	I	I	I	I	I	I
Site preparation											

Link to free RACI templates here

Threats to Goals: Scope creep

In scientific research, scope creep occurs when new questions, analyses, experiments, or project goals are added after the project has begun without formally evaluating their impact on time, budget, personnel, or the primary scientific objectives.

1 **Define in-scope & out-of-scope work**

2 **Prevent scope creep**

3 **Review and assess changes as needed**

4 **Protocol considerations; secondary protocols**

Threats to Goals: Resource limitations

1 Capacity & Enrollment

Clinical / animal capacity & throughput

Enrollment caps or logistical limits

2 Staffing & Problem Support

Coverage across components and shifts

Troubleshooting and surge support

3 Ethics, Oversight & Approvals

Ethical issues and a plan to revisit them

Oversight, approvals, and compliance

4 Resources, Tools & Supply

Access to assays, equipment, drugs, and tools

Shortages and critical supplies

5 Logistics & Access

Site, specimen, storage, and transport needs

Dependencies, handoffs, and access constraints

6 Budget & Computing

Budget flexibility and contingency

Compute, storage, software, security, and support

People

- Roles and timing
- Structure / organization
- Training
- Continuity & resilience

Who, what, where?

1 Roles & Timing

Who needs to do what — and when?

Define owners, backups, handoffs, and coverage across phases.

2 Team Structure

How will people be organized?

Clarify reporting, decision rights, interfaces, and escalation paths.

3 Training & Authorization

What must each person have before starting?

Map role-specific training, approvals, credentials, and required documentation.

4 Continuity & Resilience

How will the project absorb personnel changes?

Cross-train critical roles, document workflows, SOPs, and maintain succession / backup plans.

RACI matrix – Expertise/People

R	Responsible	A	Accountable	C	Consulted	I	Informed

Research Project	Collaborators					Resources/Experts					
	PI	Lead Researcher	Associate Investigators	Other Topic experts	Trainees/mentees	Radiology/imaging	Pharmacy	Clinic/in-patient	Lab/cores	Regulatory	Other tests/testing/equipme
Protocol Development											
Develop the research aims/goals	A	R	A	C	I	C	C	C	C	I	C
Write the protocol	A	R	C	C	I	C	C	C	C	C	I
Internal reviews - Branch, Resource, Scientific, etc	A/R	R	I	I	I	I	I	I	I	I	I
IRB submission	A/R	R	I	C/I	I	I	I	I	I	I	I
Site preparation											

Tracking People and Info

- All personnel involved with a study should be IRB listed personnel
  - NIH's Policy 102 Investigator Conflict of Interest policy
  - Biostatisticians, even when only access to de-identified data
  - Can be listed as Associate Investigators or Statistician (not engaged)

- Some data, equipment or tools require additional training or agreements
  - IRB trainings/approvals/reviews
  - Protocols
  - Consent forms for project
  - Data specific training/DUA up to date
  - Ethical oversight/ COI clearance
  - Authentication of Key resources – mouse model, cell lines, assay etc.

Plan for continuity

A training-rich environment creates turnover by design

Short-term trainees bring energy and expertise—but project knowledge must remain with the study, not with any one person.

Knowledge continuity

Use shared documentation, clear ownership, and structured handoffs.

Trainee transitions

Plan overlap, backups, and cross-training before rotations end.

No loss of assets

Store data, files, code, protocols, and equipment records in durable project locations.

Communication

- When
- How
- Who
- Why

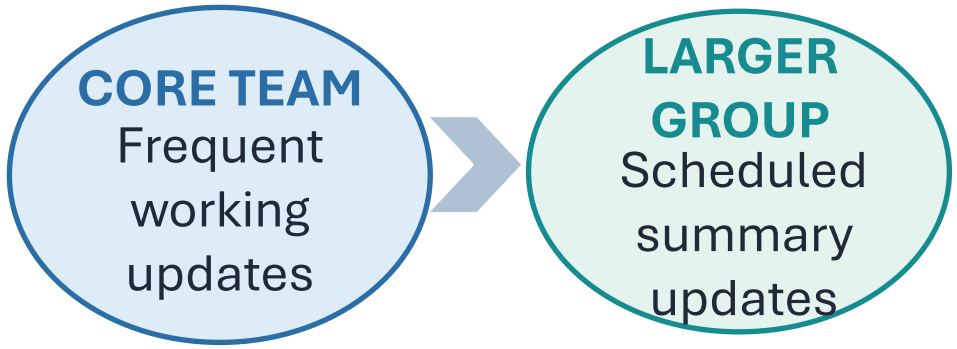
Communication

1 WHO needs to know?
Define the audience by role and decision need

2 WHEN do they need it?
Match timing to urgency and cadence

3 HOW should it be shared?
Use the right channel for the message

COMMUNICATION CADENCE



Set type + frequency

MAKE COMMUNICATION ACTIONABLE

1. **Meeting schedule** — Create a predictable rhythm
2. **Decision log** — Record what was decided + why
3. **Action tracking** — Owner • due date • status
4. **Scheduled updates** — Summarize progress and changes

- **CALENDAR** — Meetings + milestones
- **SHARED TRACKER** — Actions + decisions
- **EMAIL / CC** — Targeted awareness
- **TEAM SPACE** — Files + ongoing context

RACI matrix – Expertise/People

R	Responsible	A	Accountable	C	Consulted	I	Informed

Research Project		Collaborators					Resources/Experts					
	PI	Lead Researcher	Associate Investigators	Other Topic experts	Trainees/mentees		Radiology/imaging	Pharmacy	Clinic/in-patient	Lab/cores	Regulatory	Other tests/testing/equipme
Protocol Development												
Develop the research aims/goals	A	R	A	C	I		C	C	C	C	I	C
Write the protocol	A	R	C	C	I		C	C	C	C	C	I
Internal reviews - Branch, Resource, Scientific, etc	A/R	R	I	I	I		I	I	I	I	I	I
IRB submission	A/R	R	I	C/I	I		I	I	I	I	I	I
Site preparation												

Link to free RACI templates here

Email example: Who, When, How (What)

Info requested by August 14th, 2026 | SLP Book Club & Follow-up materials

Dear all,

Thank you to those who were able to join us for a fantastic July 30th kick-off book club meeting for *Atomic Habits*. Cheryse and Julie have graciously shared their notes/slides on the book (attached picture and PowerPoint) and the bonus chapters and other resources mentioned are available here.

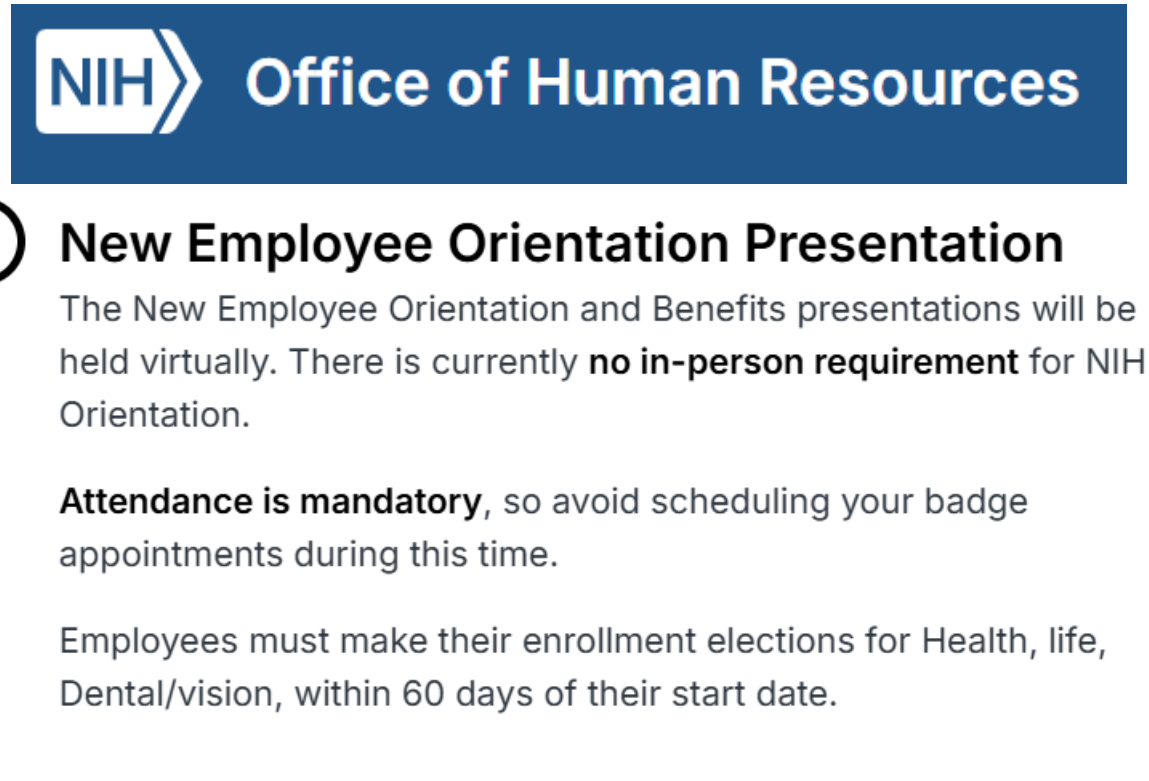
Earlier this week, Phil sent out his compiled reference list from each of the speakers from our SLP Cohort (attached). I've created a very short survey (hyperlinked here) to solicit votes for the next book, targeting end of September 2026. I'd appreciate your survey responses by 5pm Eastern August 14th to allow enough time for planning and reading.

Thanks again for your insights and engagement – I always learn so much!

Emily

Onboarding as communication

- The planned, organized process of integrating a new staff member into an organization
- Research onboarding
  - Training requirements
  - Expectations/Policies
    - Schedule/working hours
    - Safety
    - Lab meetings
    - Research meetings
    - Journal clubs/conferences



Organization and Documentation

- SOPs, notes
- Access, permissions
- File naming, templates
- Meta data
- Version control

Standard Operating Procedures (SOPs)

Written document with standardized instructions for a specific process that must be accomplished in the same way each time

- Should have sufficient detail
- Should be accessible for personnel
- May include templates or other guidance
- A research study can have MANY SOPs

PLoS Negl Trop Dis. 2016 Nov 3;10(11):e0005053. doi: 10.1371/journal.pntd.0005053

The Art of Writing and Implementing Standard Operating Procedures (SOPs) for Laboratories in Low-Resource Settings: Review of Guidelines and Best Practices

Barbara Barbé [1,*], Kristien Verdonck [1], Deby Mukendi [2], Veerle Lejon [1,3], Jean-Roger Lilo Kalo [2], Emilie Alirol [4], Philippe Gillet [1], Ninon Horié [4], Raffaella Ravinetto [1,5], Emmanuel Bottieau [1], Cedric Yansouni [6], Andrea S Winkler [7,8], Harry van Loen [1], Marleen Boelaert [1], Pascal Lutumba [2,9], Jan Jacobs [1,10]

EDITORIAL PLoS Comput Biol. 2020 Sep 3;16(9):e1008095. doi: 10.1371/journal.pcbi.1008095

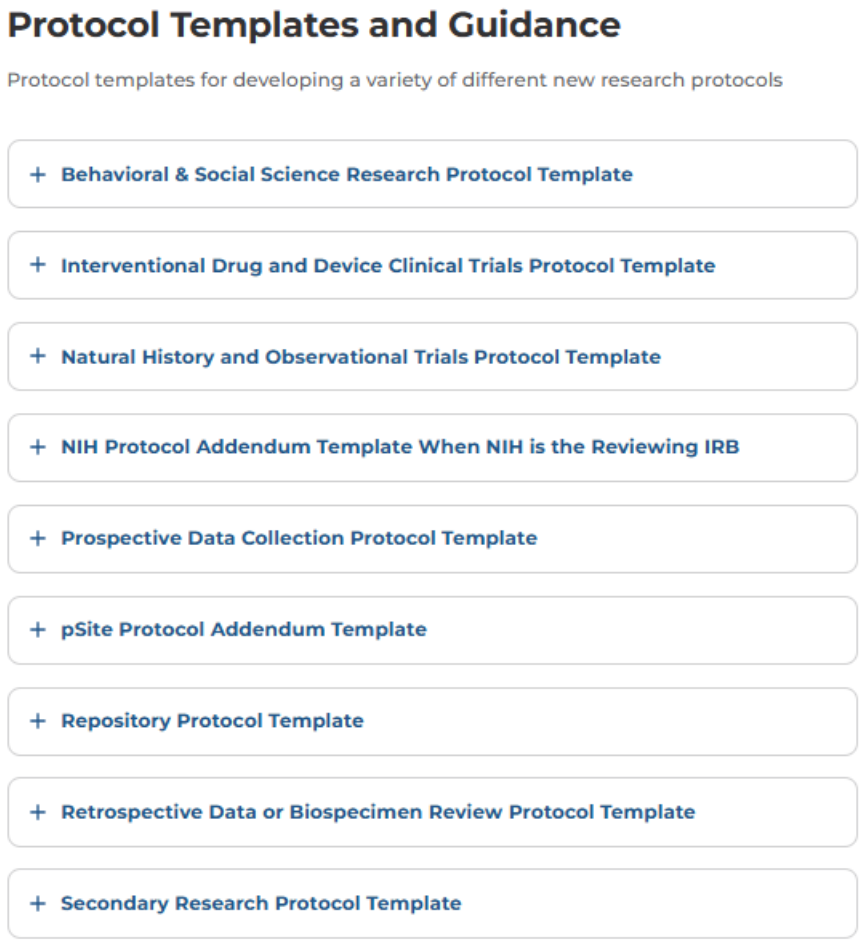
Ten simple rules on how to write a standard operating procedure

Susanne Hollmann [1,2,*,#], Marcus Frohme [3], Christoph Endrullat [3,¤], Andreas Kremer [4], Domenica D'Elia [5,#], Babette Regierer [2,#], Alina Nechyporenko [3,6,#]; on behalf of Cost Action CA15110[¶]

Templates

A pre-designed file (generally) that serves as a reusable, consistent guide for a process or activity.

- Letterhead with pre-specified formatting
- Email for consistent, complete info
- Code for routine or general tasks
- Data for data collection
- Some journals have manuscript templates



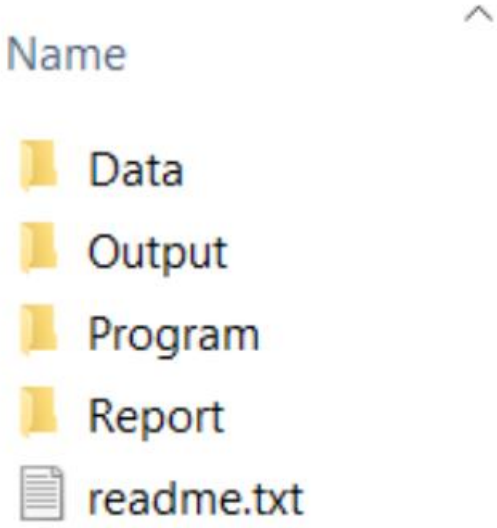
File naming, organization, version control

- Location of files and ownership/access
- File organization
  - Standard folder naming and organization
  - Know where files will be located
  - OLD or ARCHIVE folder
- File contents/naming:
  - English language
  - All tables/figures must have a title
  - All axes of a figure must be labeled and formatted
  - Date in file name for sorting, version control
  - DO NOT DELETE FILES

Project management

A .txt readme file is required to be located in the project folder.

This readme file should include the information below for each time you update this project.



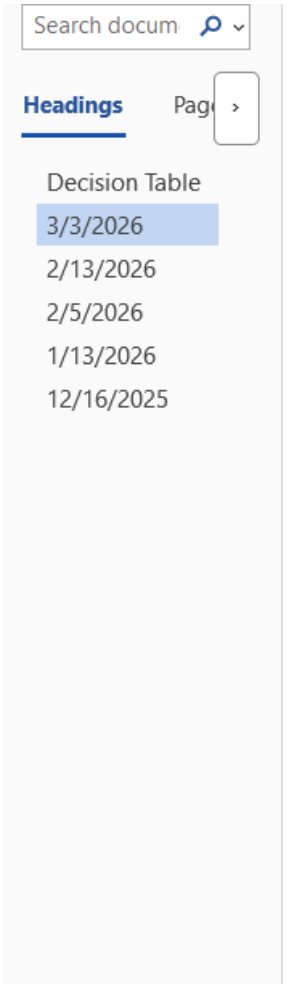
Line 1: Time (yyyy/mm/dd)

Line 2: Author

Line 3: What are the current tasks

Line 4: The work you have done

Meeting notes, decision tracking, and references



12/16/2025

Meeting Attendees listed

- discussion topics
- decisions made and why
- Items to remember for later

Action Items – next meeting 1/13/2026 10am

List out what everyone is doing for the next meeting

Decision Table

Redcap Data	As listed on clinicaltrials.gov	Planned Calculation	Adapted Calculation
Variable name	Formatted name	Equation	Final equation

Clinicaltrials.gov requires reporting both race and ethnicity so have recoded as follows. Need to confirm agreement

Original	Recoded	
Ethnic Origin	**Race**	**Ethnicity**
Hispanic	White	Hispanic
Black	Black	Non-Hispanic
African American	Black	Non-Hispanic

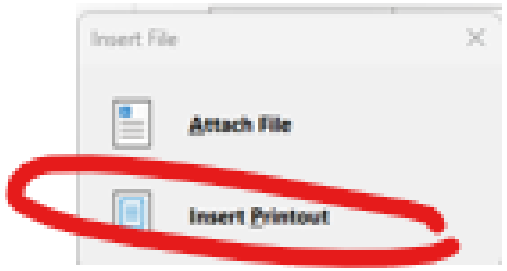
EL – relabel variables as in Excel file

EL - rerender Quarto & update table with variable names

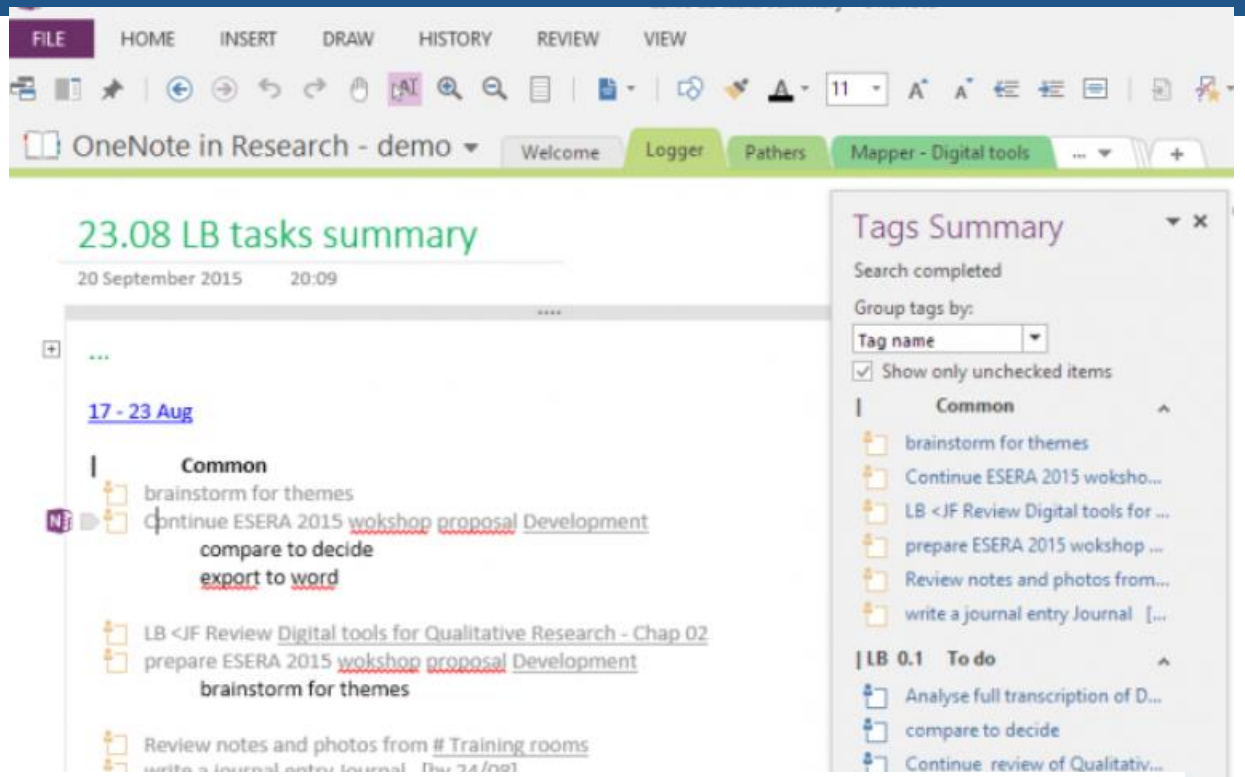
EL – resend this output

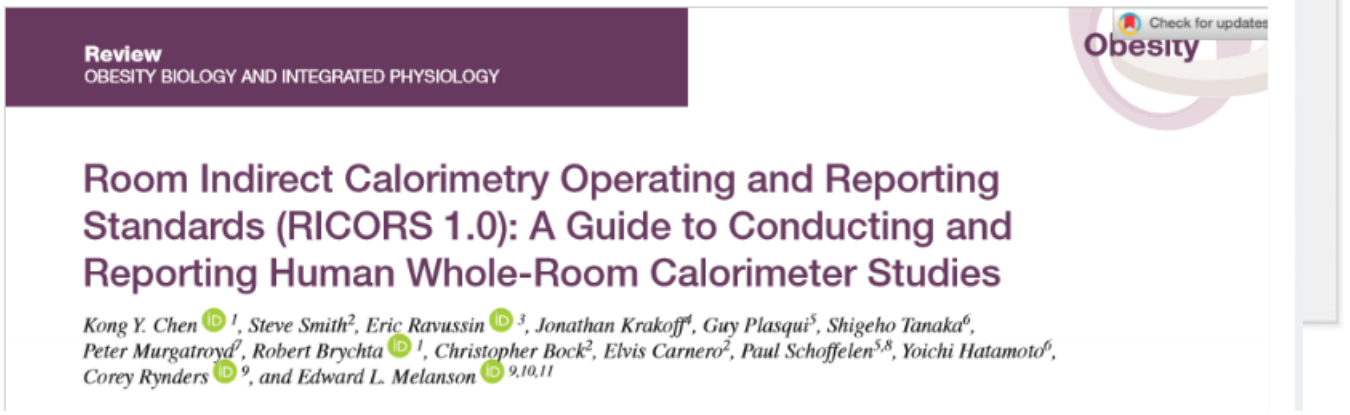
OneNote

- “electronic 3-ring binder”
- CIT approved and free
- Excellent search function
- insert files, connects with Outlook tasks
- insert images, videos, links, tags, annotation
- Easy to share but can add passwords



IRB Trainings	Emily	Emily2	Emily3
GCP	☑	Leary CITI GCP USFD...	6/10/2026
Biomedical 101		Obesity - 2020 - Che...	

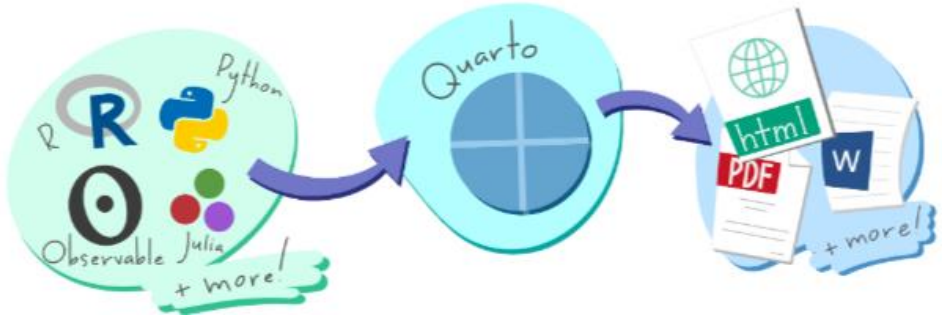




Integrated documentation

- Workflow automation
  - NIDDK-CR presentation on “Design Patterns for Reproducible Research”
- R Markdown/markdown
  - Markdown is a plain-text formatting language
  - R Markdown files are the source code for rich, reproducible documents – place code, figures, written explanations, decisions with reasoning – all in ONE place
- Quarto
  - is a multi-language version of Markdown, open-source scientific and technical publishing system

1. Literate Programming
   - Jupyter Notebooks
   - Markdown
   - Pseudocode
2. Everything as Code
   - Mermaid Diagrams
   - `gt` for Tables
   - LaTex
3. Version Control and Provenance
   - `devtools`
4. Code to Package
   - Functions
   - Auto-document
   - Unit tests
5. Social Coding



R Markdown Example

RMarkdown

Emily Leary, PhD

2025-07-02

Note that output accepts the following types:

- html_document, which will create HTML output (default)
- pdf_document, which will create PDF output
- word_document, which will create Word output

Import and Data Manipulation

Setting up readr

First, we need to be able to import data to work with it in R. We'll use the readr package to do this (part of the tidyverse). Note what I'm adding to the beginning line of where I want my R code and then what I'm adding when I'm done writing code, to denote that the code block is completed. Note that I've also included "echo=TRUE" so that the code and the results are both in the rendered report. If I used "echo=FALSE" then the code would be omitted.

```
library(readr)
```

```
## Warning: package 'readr' was built under R version 4.4.3
```

Importing data using readr

Note that unless you change your working directory or use the full file path, R will look to the location where your code is currently saved as your working directory.

```
data <- read_csv("ctg-studies CSV.csv", show_col_types = FALSE)
```

You can change your working directory using the following, updating the file path to whatever you use.
Note the way that the slashes are pointing, they are not the default when copying from a PC!

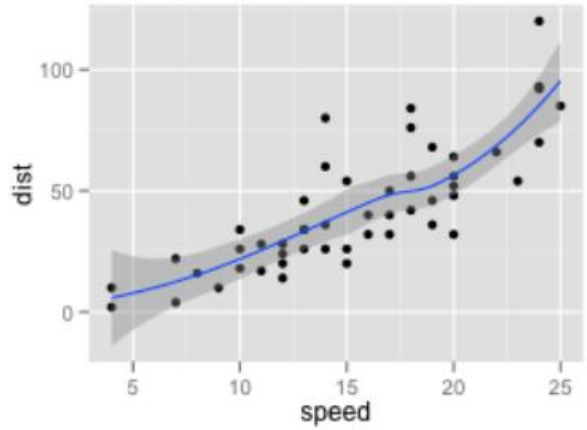
R Code Chunks

With R Markdown, you can insert R code chunks including plots:

```
# quick summary and plot
library(ggplot2)
summary(cars)
```

```
##      speed           dist
##  Min.   : 4.0   Min.   :  2
##  1st Qu.:12.0   1st Qu.: 26
##  Median :15.0   Median : 36
##  Mean   :15.4   Mean   : 43
##  3rd Qu.:19.0   3rd Qu.: 56
##  Max.   :25.0   Max.   :120
```

```
qplot(speed, dist, data = cars) + geom_smooth()
```



Meta data

Meta data = data about data

- Info to reconstruct context and evaluate if fit for purpose
- Should be complete enough that users know what can and cannot be done using the data, in a complete way
- README files – provide needed details and guidance
- Standardized formats for info – CDEs, ICD coding etc
- Standardized file naming formats to facilitate easy use/matching
  - SurveyForm_20260813.pdf with companion file SurveyForm_20260813.dat

Cornell University AUTHOR_DATASET_ReadmeTemplate.txt

README file

GENERAL INFORMATION

1. Title of Dataset: Data from: A Community of Interests: A Social History of the British in Buenos Aires, 1860-1914
2. Author Information
    A. Principal Investigator Contact Information
        Name: Deborah Jakubs
        Institution: Stanford University; Archivo General de la Nacion, Buenos Aires, Argentina
        ORCID:
        Email: deborah.jakubs@duke.edu

    B. Data Processor
        Name: John Little
        Institution: Duke University
        Contact: https://www.johnlittle.info/

Link here

Cornell University AUTHOR_DATASET_ReadmeTemplate.txt

3. Date of data collection (single date, range, approximate date) 1977-78

4. Geographic location of data collection <latitude, longitude, or city/region, State, Country, as appropriate>: Buenos Aires, Argentina

SHARING/ACCESS INFORMATION

1. Licenses/restrictions placed on the data: Data files are under a CC0 Waiver, Public domain designation; R Code and Processes CC-By 4.0 John Little

2. Links to publications that cite or use the data: A Community of Interests: A Social History of the British in Buenos Aires, 1860-1914 https://hdl.handle.net/10161/23839

DATA & FILE OVERVIEW

- census_1869.csv - 1869 Argentina National Census, Buenos Aires, British Nationality Extract, sorted by last name
- census_1895.csv - 1895 Argentina National Census, Buenos Aires, British Nationality Extract, unsorted
- census_1895_sorted.csv 1895 Argentina National Census, Buenos Aires, British Nationality Extract, sorted by last name,
- Documentation.pdf - Codebook file extracted from dissertation text located here: https://hdl.handle.net/10161/23839
-/R_Docs.zip - Documentation and files for the jakubs.Rproj used to clean and transform the original ASCII data.
    .gitignore
    .Rbuildignore
    .Rhistory
    .Rprofile
    01_ingest_explore.nb.html
    jakubs.Rproj
    LICENSE.md
    renv.lock
    README.md
    01_ingest_explore.Rmd

Research Project Management + AI

Other considerations: AI in research

Welcome to the U.S. Department of Health & Human Services workspace

As government professionals, it is critical to safeguard the confidentiality, integrity, and availability of federal information. Users must exercise caution and avoid entering information that could put government operations, employees, or the public at risk.

When using ChatGPT please do not enter:

- Sensitive Personally Identifiable Information (SPII)
- Protected Health Information (PHI)
- Classified Information/Data
- Export Controlled Data
- Confidential Commercial Information or Trade Secret Data

Users should be skeptical of everything you read, watch for potential bias, and treat answers as suggestions. Before making a significant decision, make sure you have considered original sources and counterarguments.

Other considerations: AI in research

Science

"…we are now updating our license and Editorial Policies to specify that text generated by ChatGPT (or any other AI tools) cannot be used in the work, nor can figures, images, or graphics be the products of such tools. And an AI program cannot be an author. "

Link here

nature

Risk assessment level	Green-Permitted: Assistive AI use (low risk)
Summary	AI use that supports expression, organisation, or efficiency without influencing scientific, scholarly or evaluative judgement.
Risk assessment level	Amber - Exercise caution: Evaluative or interpretive AI use (requires caution and care)
Summary	AI use that may influence interpretation, framing, emphasis, or evaluative judgement, but remains under human control.
Risk assessment level	Red - Not permitted: AI replacing scholarly judgement or lacking transparency
Summary	AI use that is opaque or replaces accountable human contribution, generates unverifiable outputs, or compromises confidentiality or integrity.

Link here

JAMA

"…authors must take responsibility for the integrity of the content generated by these tools. Authors should report the use of artificial intelligence…or similar technologies to create content or assist with writing or editing of manuscripts in the Acknowledgment section or Methods section if this is part of formal research design or methods.

Link here

Key Takeaways

- Outcomes / Deliverables
- SMART goals
- Milestones/timeline
- Resources
- People

- Roles and timing
- Structure / organization
- Training
- Continuity & resilience

- When
- How
- Who
- Why

- SOPs, notes
- Access, permissions
- File naming, templates
- Meta data
- Version control

Bonus Slide before/after using HHS ChatGPT

Threats to Goals: Scope creep

In scientific research, scope creep occurs when new questions, analyses, experiments, or project goals are added after the project has begun without formally evaluating their impact on time, budget, personnel, or the primary scientific objectives.

- Define in-scope and out-of-scope work
- Prevent scope creep
- Review and assess changes as needed
- Protocol considerations; secondary protocols

Before slide: no color or pictures!

After slide: Lots of pictures and icons! But NOT editable ☹

Final slide with editable text.
Note that some of the pictures were adjusted and don't look as nice.