

Practical Advice on Doing Ethical Research in Applied Economics

Jeffrey D. Michler

1 September 2023

With thanks to Anna Josephson (UA), April Athnos (UA), Travis Lybbert (UC Davis), and Sarah Janzen (UIUC)

Research Ethics in Applied Economics

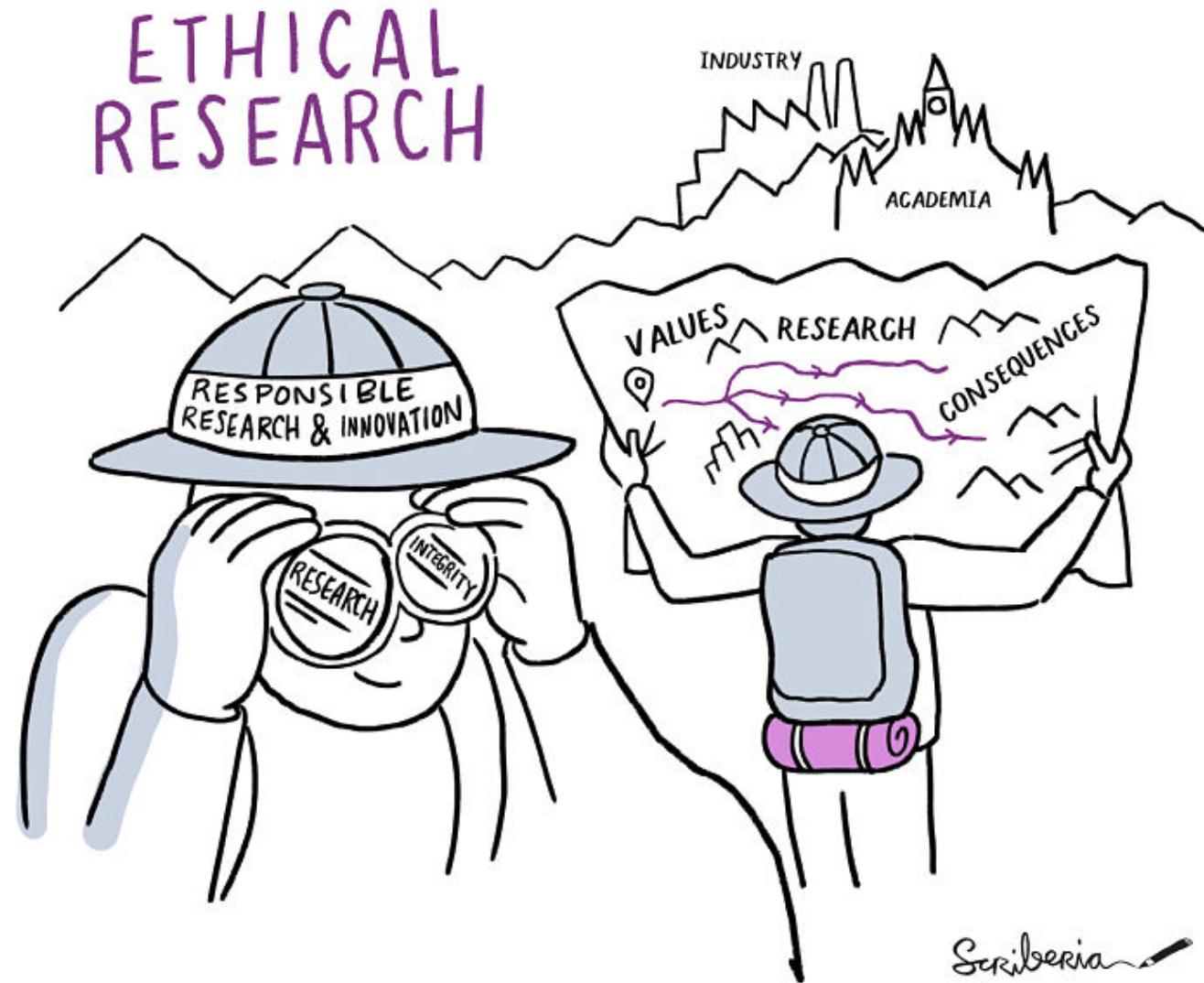
A Practical Guide

Anna Josephson and Jeffrey D. Michler



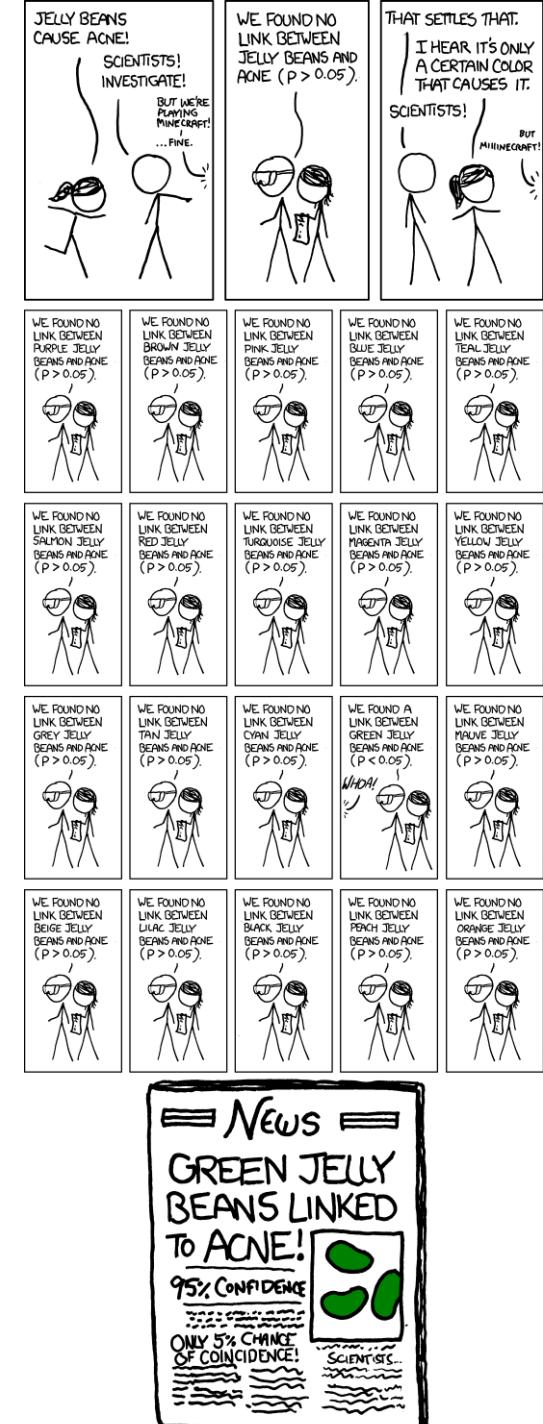
Last Week

- What is research ethics?
- Issues with data collection?
- Issues with data analysis?

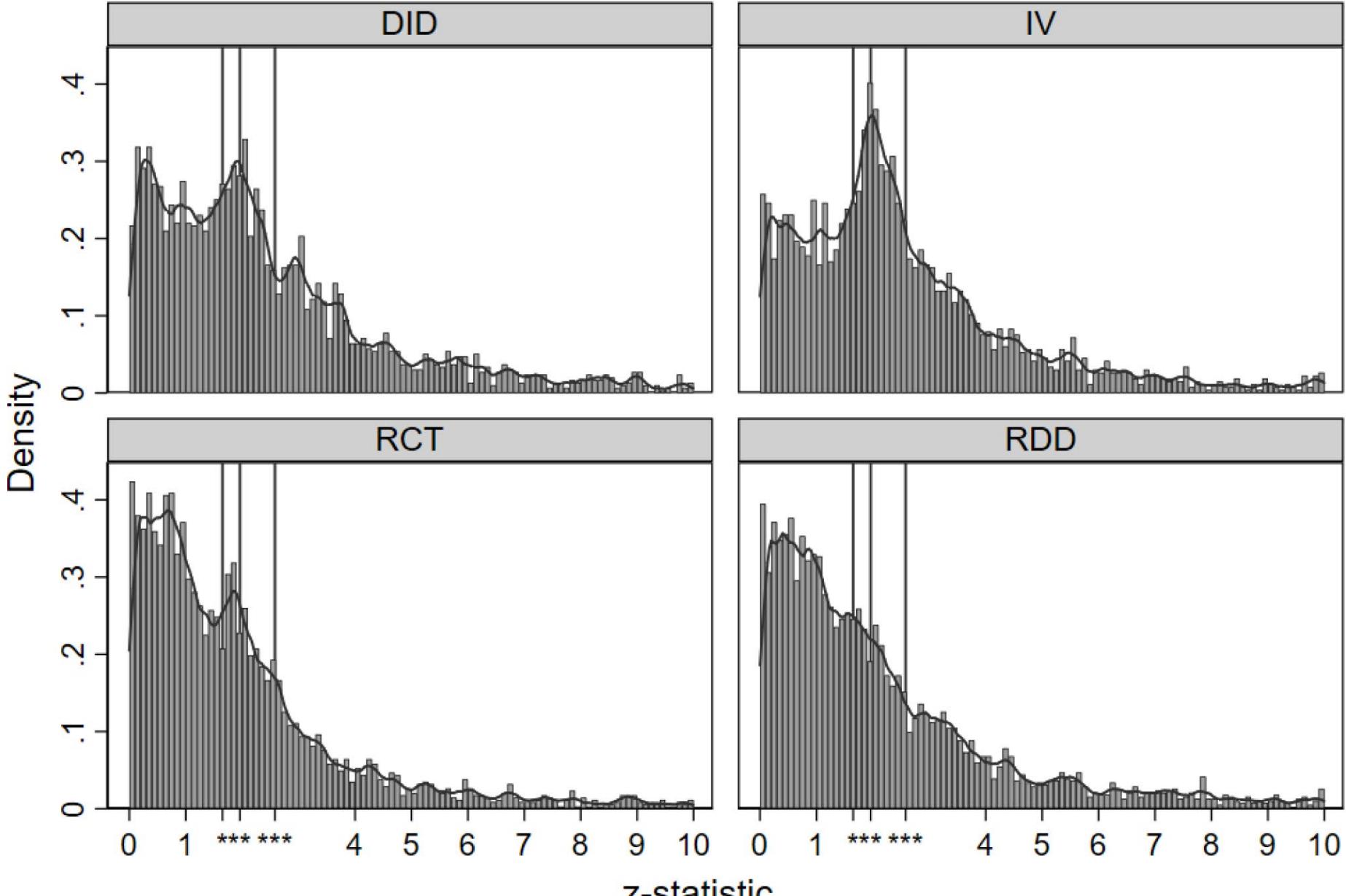


Changing Incentives

- 1) Pre-commitments to “tie hands”
 - 1) Pre-analysis plans
 - 2) Journals accept research plans then publish results
 - 3) Hypothesis registry
- 2) Transparency and Reproducibility
 - 1) Lab notebook approach & submission requirements
 - 2) Transparency when search over many models performed
 - 3) Reproducible workflow



Brodeur et al. (2016)



Graphs by Method

What Is A Pre-Analysis Plan?



They bound me hand and foot in the tight ship—
erect at the mast-block, lashed by ropes to the mast.

— Homer, The Odyssey, Book XII

A **pre-analysis plan** is a step-by-step plan setting out how a researcher will analyze data which is written in advance of them seeing this data (and ideally before collecting it in cases where the researcher is collecting the data).

What Can A Pre-Analysis Plan Accomplish?

1. Restrict **which hypotheses** to investigate.
2. Restrict **how to test** the hypotheses.
3. Pre-commit to a **data collection** plan.

The result is that a pre-analysis plan should increase the probability that a published positive result is true.

Pre-Analysis Plans in Medicine

- Plans has been required by law in the United States for all drug trials since 1997
 - Stated goal was to counter strong incentives of drug companies to find positive results
- In 2004 the twelve journals of International Committee of Medical Journal Editors announced public trial registration would be pre-requisite for publication
 - Stated goal was to counter the “file drawer problem,” where only significant results get published
- ClinicalTrials.gov is where most are registered
 - Pre-2004, registry averaged 1,879 a year
 - Post-2004, registry averaged 22,088 a year

Pre-Analysis Plans in Social Science

- Pre-analysis plans predate registry
- Most commonly cited first studies:
 - Casey, Glennerster, and Miguel (2012). “Reshaping institutions: Evidence on aid impacts using a preanalysis plan.” *QJE*.
 - Finkelstein et al. (2012). “The Oregon health insurance experiment: Evidence from the first year.” *QJE*.
- In 2012 AEA established the AEA RCT Registry
 - Allows for registration of lab and field experiments

Pre-Analysis Plans v. Registrations

- In medical science, there has been a clear distinction between:
 - What pre-analysis plans accomplish - p-hacking
 - What public registration accomplishes - file drawer problem
- In economics, these two have been muddled
 - See Olken (2015); Coffman et al. (2015); Burlig (2018); Duflo et al. (2020)
 - Making pre-analysis plans work means understanding the difference between what a plan accomplishes and what a public registry accomplishes

Basics of a Pre-Analysis Plan

- Data description
 - Describe the process by which data was generated or collected
- Hypotheses
 - Clearly state the research questions and hypotheses to be tested
- Outcome variables
 - List the primary outcome variables as well as any secondary outcomes to be tested
- Variable construction
 - Precisely define how each outcome and control variable is to be constructed
- Model specification
 - Describe the statistical method of analysis, including covariates and the approach to inference

Prevent *p*-Hacking

- Forming a Ulysses Pact is a good way to prevent p-hacking
- They bind the researcher to limit research degree of freedoms
- Important when there are strong incentives to cheat (drug trials)
- Or when there is leeway for what researchers can “discover”
 - Coffman and Niederle (2015) - economic theory limits what a research might “find”
 - Brodeur et al (2016, 2020) - little evidence of p-hacking in experiments

Solve the File Drawer Problem

- Ulysses Pacts don't solve the file drawer problem, but public registries do
- Regardless of the empirical methods a “hypotheses registry” creates database on the universe of research questions asked
- Pre-analysis plans and public registries work together but are not the same thing
 - Coffman and Niederle (2015) - promote basic plan and public registration
 - Registered reports at *Journal of Development Economics*, the *Journal of Accounting Research*, *The Review of Financial Studies*, and *PLOS One*

Changes in Culture

- Institutional responses
 - Open Science Foundation
 - Berkeley Institute for Transparency in the Social Sciences
- Journal policies
 - Transparency and Openness Promotion
 - “The *AJAE* adheres to TOP “Level 1” standards as described here...”

Data and Documentation

Data sources should be indicated clearly in the article. Upon acceptance and in keeping with evolving policies at other top economics journals, authors are expected to submit their datasets and associated documentation (perhaps in a readme file), for readers to download from the *AJAE* website. Authors are also expected to provide whatever other material is needed to ensure that their results can be replicated (this might include code or pseudo-code used in estimation). Authors are encouraged to comply with all of this policy, but the editors would prefer partial compliance over non-compliance. The submitted files will be accessible as clickable links to the online version of the published paper. It is understood that in cases where the data are proprietary or confidential authors cannot fully satisfy this policy. Such cases should be brought to the attention of the editor at the time of initial submission.

Transparency and Openness

- Code and data management have improved with this openness expectation – less data sloppiness
 - Data notebook

```
131 drop _merge
132 save "C:\Users\aljosephson\Dropbox\Out for Review\DISE2_Sorghum\newrebuild_10Jan-withstaple_rain.dta", replace
133 probit panelhh grow_sorghum tot_sorghum_area grow_maize tot_maize_area plot_area assetvalue head_age head_edu num_cattle plough rec_ext sorg
134 tab plot_area, missing
135 tab head_age, missing
136 sort yearpanel
137 sort panelhh
138 by panelhh: tab head_age
139 by panelhh: tab head_age, missing
140 *tried to come up with a better method ... but didn't find one
141 *so by hand replaced information, based on, if household head was the same, and thus could determine if someone aged, by how much, etc.
142 *still missing about 100 values for head_age and head_edu
143 sort rc id yearpanel head_age
144 order rc id yearpanel head_age head_edu
145 replace head_age = 71 in 8
146 replace head_age = 60 in 11
147 replace head_age = 54 in 13
148 replace head_edu = 5 in 13
149 replace head_age = 43 in 26
150 replace head_edu = 11 in 26
151 replace head_age = 56 in 33
152 replace head_edu = 8 in 33
```

- Will enhanced transparency provide sufficient information to detect misconduct?

Reproducible Workflow

```
1 * Project: COVID Food Security
2 * Created on: 22 September 2021
3 * Created by: lirr
4 * Edited by: jdm
5 * Last edited: 23 September 2021
6 * Stata v.17.0
7
8 * does
9     * reads in baseline burkina faso data
10    * pulls FIES data questions
11
12 * assumes
13     * raw burkina faso data
14
15 * TO DO:
16     * complete
17
18 ****
19 ***# 0 - setup
20 ****
21
22
23 * define
24     global root      = "$data/burkina_faso/raw"
25     global export    = "$data/burkina_faso/refined"
26     global logout   = "$data/burkina_faso/logs"
27
28 * open log
29     cap log          close
30     log using        "$logout/bf_fies", append
31
32 ****
33 ***# 1 - FIES data
34 ****
35
36
37 * load data
38     use              "$root/wave_00/menage/s08a_me_BFA2018", clear
39
40 * check for unique identifier
41     isid             grappe menage
42
43 * replace counts with binary indicators
44     lab def          yesno 1 "Yes" 0 "No"
45     foreach          x in 1 2 3 4 5 6 7 8{
46         replace         s08aq0`x' = 0 if s08aq0`x' > 1 & s08aq0`x' < .
47         lab val         s08aq0`x' yesno
48     }
49
50 * rename variables
51     rename           s08aq01    fies_4
52     rename           s08aq02    fies_5
53     rename           s08aq03    fies_6
54     rename           s08aq04    fies_7
55     rename           s08aq05    fies_8
56     rename           s08aq06    fies_1
57     rename           s08aq07    fies_2
58     rename           s08aq08    fies_3
59
60 * !clear
61
```

- Preamble

- Notes on when, who, and what
- Any necessary packages or data

- Set-up

- Use relative not absolute directory paths
- Keep code and data segregated

- Style

- Code should be a readable narrative
- Code for human beings not computers (Knuth, 1984)

Version Control with GitHub

Current repository **covid_food_security**

Current branch **fies** #1

Fetch origin
Last fetched 2 minutes ago

Changes **History**

Select branch to compare...

finished cleaning bf_fies_0

44iggs44 · d3edda1 ± 3 changed files +24 -49 New

Line	File	Change Type	Content
33	burkina_faso\bf_fies_0.do	+	*****
34		-	**# 1 - FIES data
35		+	*****
36	nigeria\nga_fies_ph0.do	+	*****
37		-	* load data
38		-	use "\$root/menage/s08a_me_BFA2018", clear
39		+	use "\$root/wave_00/menage/s08a_me_BFA2018", clear
40	nigeria\nga_fies_pp0.do	+	*****
41		-	* replace counts with binary indicators
42		lab def	yesno 1 "Yes" 0 "No"
43		-	foreach x in a b c d e f g h{
44		-	replace s8q02`x' = 1 if s8q02`x' > 1 & s8q02`x' < .
45		-	lab val s8q02`x' yesno
46		+	foreach x in 1 2 3 4 5 6 7 8{
47		+	replace s08aq0`x' = 0 if s08aq0`x' > 1 & s08aq0`x' < .
48		+	lab val s08aq0`x' yesno
49		}	
50		-	
51		-	
52		48	* rename variables
53		-	rename s8q01 fies_4
54		-	rename s8q02a fies_5
55		-	rename s8q02b fies_6
56		-	rename s8q02c fies_8
57		-	rename s8q02d fies_7
58		-	rename s8q02e fies_2
59		-	rename s8q02f fies_9
60		-	rename s8q02g fies_1
61		-	rename s8q02h fies_3
62		49	+ rename s08aq01 fies_4
63		+	rename s08aq02 fies_5
64		+	rename s08aq03 fies_6
65		+	rename s08aq04 fies_7
66		+	rename s08aq05 fies_8
67		+	rename s08aq06 fies_1
68		+	rename s08aq07 fies_2
69		+	rename s08aq08 fies_3

Merge branch 'main' of https://github....
jdmichler · Sep 9, 2021

update project name
jdmichler · Sep 9, 2021

Update README.md
Jeffrey D. Michler · Sep 9, 2021

update project directories
jdmichler · Sep 9, 2021

Merge branch 'main' of https://github....
jdmichler · Sep 9, 2021

update project name
jdmichler · Sep 9, 2021

Update README.md
Jeffrey D. Michler · Sep 9, 2021

Bringing “Sinning” Out of the Basement

- “Sinners are not expected to avoid sins; they need only confess their errors openly.” Edward Leamer (1978)
- The goal should be open and transparent analysis
- Pre-define a research plan, even if you don’t develop a full PAP
- Keep meticulous records
- Ask questions
- Learn by doing

Homecoming



Due: 15 September



Research Ethics and Compliance Training

A grid of four cards representing training modules. From left to right: 1. "Laboratory Chemical Safety" (NEW) - Describes fundamental safety practices for working with hazardous chemicals. Includes a hazard symbol graphic. 2. "Webinars" (NEW) - Offers topics like GDPR, FERPA, and the revised Common Rule. Includes a microphone and laptop icon. 3. "Protocol Registration and Results Summary Disclosure in ClinicalTrials.gov" (NEW) - Provides a video-enhanced practical guide to protocol registration and reporting. Includes a tablet icon. 4. "Biomedical PI" (NEW) - Reviews the role of the biomedical investigator in clinical investigation. Includes a test tube icon. Each card has a "View Course" button at the bottom. Navigation arrows are present on the left and right sides of the grid.

- Complete CITI training: “Social and Behavioral Responsible Conduct of Research”
- Send certificate to: aljosephson@arizona.edu