THE UNIVERSITY OF MELBOURNE

Faculty of Medicine, Dentistry and Health Sciences

Australian Clinical Trials Alliance

STInG
STATISTICS IN TRIALS INTEREST GROUP

Conducting efficient and reproducible research in Stata

Dr Anurika De Silva

Methods and Implementation Support for Clinical and Health (MISCH) research Hub
Centre for Epidemiology and Biostatistics
Melbourne School of Population and Global Health
The University of Melbourne

anurika.de@unimelb.edu.au @AnurikaDeSilva

Acknowledgments: Ms Sabine Braat, A/Prof Emily Karahalios, A/Prof Mark Chatfield, Dr Kristy Robledo, Prof Katherine Lee

Housekeeping

- Please keep your microphone switched off during the webinar.
- You are welcome to leave your video on or off as you prefer.
- If you have any questions, please feel free to enter them in the chat box. We will review them throughout the webinar.
- Note that this presentation will be recorded.
- A copy of the slides and Stata code will be provided after the webinar.

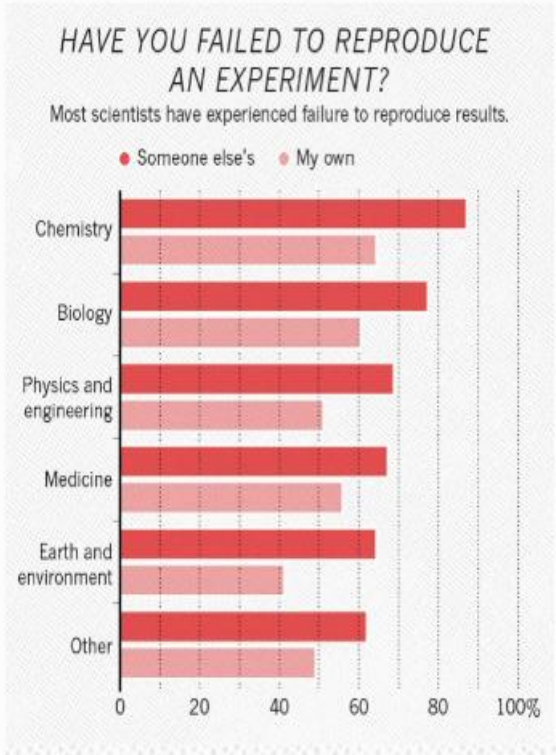
Outline

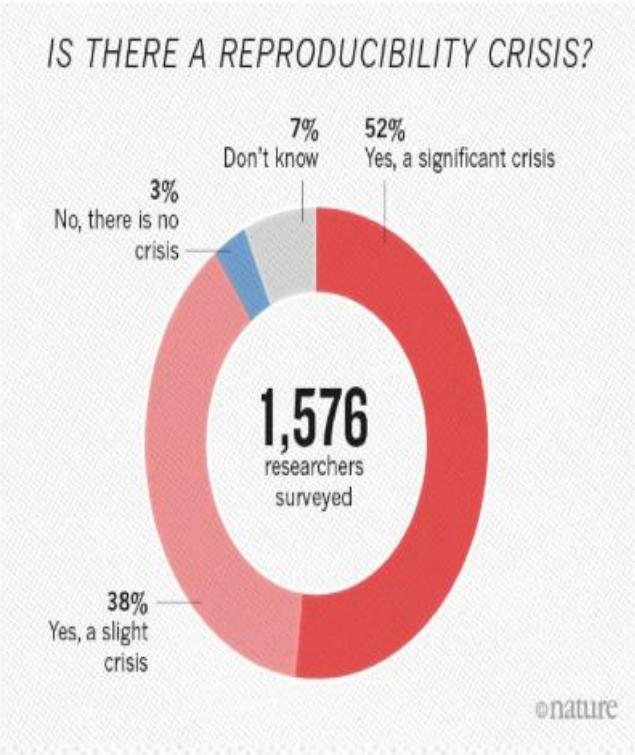
1. Importance of reproducible research
2. Tips for reproducible research
3. Motivating example
4. Organising your files
5. Conducting reproducible research in Stata
6. Discussion and Questions

Image source: https://coderefinery.github.io/reproducible-research/motivation/

Importance of reproducible research

- What is reproducible research?
- Reproducibility crisis
- Improved accuracy and efficiency
- Increased transparency
- Better collaboration
- Journals now promoting reproducible research





Baker, M. *1,500 scientists lift the lid on reproducibility*. Nature 533, 452–454 (2016).

Tips for reproducible research

- Applies to the entire study cycle
- Organising your files
- Version control
- Naming conventions
- Data management
- Communications management
- Project progress, time tracking
- Following and documenting procedures

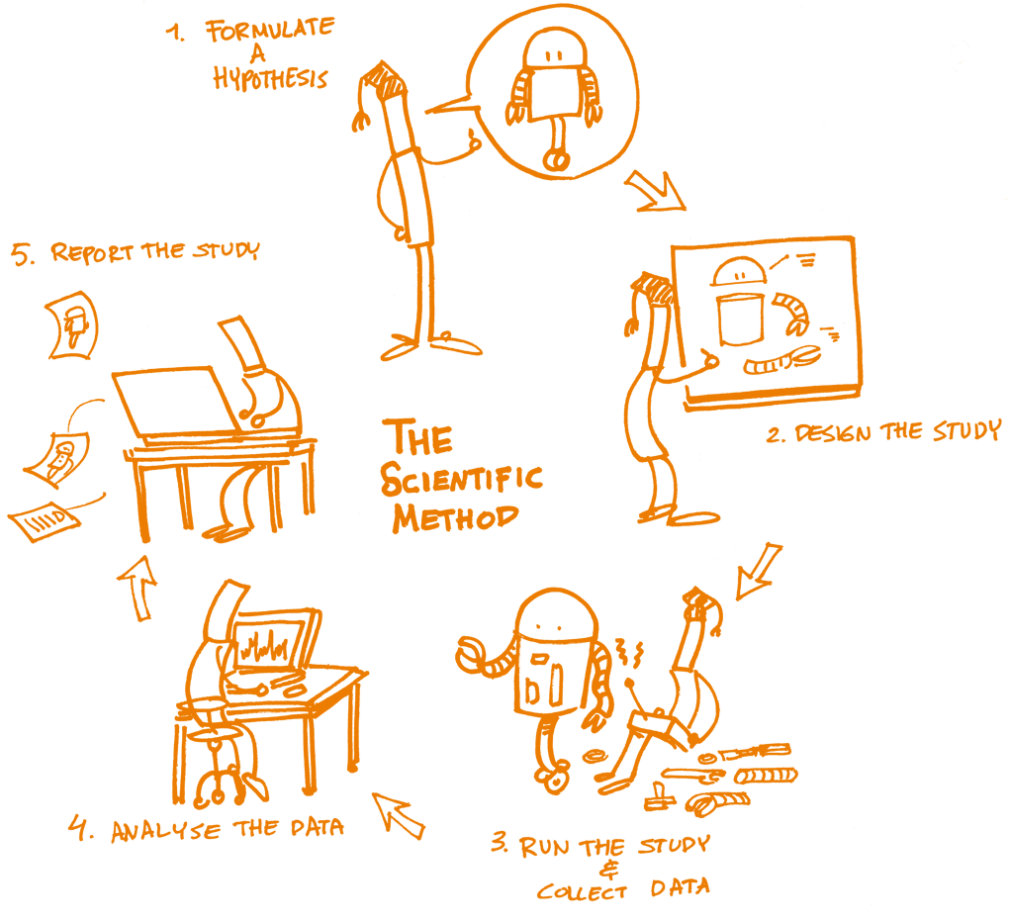


Image source: https://open-science-training-handbook.gitbook.io/book/open-science-basics/reproducible-research-and-data-analysis

Motivating example

Objective: Compare molecularly targeted therapy based on tumour profiling (MTA) versus conventional therapy (CT) for advanced cancer

Primary endpoint: Progression-free survival

Data available: https://search.r-project.org/CRAN/refmans/ipcwswitch/html/SHIdat.html

Molecularly targeted therapy based on tumour molecular profiling versus conventional therapy for advanced cancer (SHIVA): a multicentre, open-label, proof-of-concept, randomised, controlled phase 2 trial

Christophe Le Tourneau, Jean-Pierre Delord, Anthony Gonçalves, Céline Gavoille, Coraline Dubot, Nicolas Isambert, Mario Campone, Olivier Trédan, Marie-Ange Massiani, Cécile Mauborgne, Sebastien Armanet, Nicolas Servant, Ivan Bièche, Virginie Bernard, David Gentien, Pascal Jezequel, Valéry Attignon, Sandrine Boyault, Anne Vincent-Salomon, Vincent Servois, Marie-Paule Sablin, Maud Kamal, Xavier Paoletti, for the SHIVA investigators

Lancet Oncol 2015; 16: 1324–34

Published Online September 3, 2015 http://dx.doi.org/10.1016/S1470-2045(15)00188-6

See Comment page 1276

Department of Medical Oncology, Institut Curie, Paris & Saint-Cloud, France (C Le Tourneau MD); **Department of Medical Oncology, Institut Claudius Régaud, Toulouse, France** (Prof J-P Delord MD); **Department of Medical Oncology, Institut Paoli-Calmettes, Marseille, France** (Prof A Gonçalves MD); **Department of Medical Oncology, Centre Alexis Vautrin, Nancy, France** (C Gavoille MD); **Department of Medical Oncology, Institut Curie, Saint-Cloud, France** (C Dubot MD); **Department of Medical Oncology, Centre Georges-François Leclerc, Dijon, France** (N Isambert MD);

Summary

Background Molecularly targeted agents have been reported to have anti-tumour activity for patients whose tumours harbour the matching molecular alteration. These results have led to increased off-label use of molecularly targeted agents on the basis of identified molecular alterations. We assessed the efficacy of several molecularly targeted agents marketed in France, which were chosen on the basis of tumour molecular profiling but used outside their indications, in patients with advanced cancer for whom standard-of-care therapy had failed.

Methods The open-label, randomised, controlled phase 2 SHIVA trial was done at eight French academic centres. We included adult patients with any kind of metastatic solid tumour refractory to standard of care, provided they had an Eastern Cooperative Oncology Group performance status of 0 or 1, disease that was accessible for a biopsy or resection of a metastatic site, and at least one measurable lesion. The molecular profile of each patient's tumour was established with a mandatory biopsy of a metastatic tumour and large-scale genomic testing. We only included patients for whom a molecular alteration was identified within one of three molecular pathways (hormone receptor, PI3K/AKT/mTOR, RAF/MEK), which could be matched to one of ten regimens including 11 available molecularly targeted agents (erlotinib, lapatinib plus trastuzumab, sorafenib, imatinib, dasatinib, vemurafenib, everolimus, abiraterone, letrozole, tamoxifen). We randomly assigned these patients (1:1) to receive a matched molecularly targeted agent (experimental group) or treatment at physician's choice (control group) by central block randomisation (blocks of size six). Randomisation was done centrally with a web-based response system and was stratified according to the Royal Marsden Hospital prognostic score (0 or 1 *vs* 2 or 3) and the altered molecular pathway. Clinicians and patients were not masked to treatment allocation. Treatments in both groups were given in accordance with the approved product information and standard practice protocols at each institution and were continued until evidence of disease progression. The primary endpoint was progression-free survival in the intention-to-treat population, which was not assessed by independent central review. We assessed safety in any patients who received at least one dose of their assigned treatment. This trial is registered with ClinicalTrials.gov, number NCT01771458.

Le Tourneau C et al. (2015). *Molecularly targeted therapy based on tumour molecular profiling versus conventional therapy for advanced cancer (SHIVA): a multicentre, open-label, proof-of-concept, randomised, controlled phase 2 trial*. The Lancet Oncology, 16(13).

Organising your files – Folder structure

Name	Date modified	Type	Size
1. Research_Planning_and_Design	9/02/2023 8:47 PM	File folder	
2. Data_Analysis_and_Reporting	11/02/2023 11:45 PM	File folder	
3. Data_Analysis_and_Reporting_DSMB_1	9/02/2023 8:48 PM	File folder	

Name	Date modified	Type	Size
1. Literature	9/02/2023 8:49 PM	File folder	
2. Grant	9/02/2023 8:46 PM	File folder	
3. Sample_Size	9/02/2023 8:46 PM	File folder	
4. Protocol_and_CRF	9/02/2023 8:46 PM	File folder	
5. Database	9/02/2023 8:46 PM	File folder	
6. Randomization	9/02/2023 8:46 PM	File folder	
7. Stat_Plan	9/02/2023 8:47 PM	File folder	
8. DSMB_Charter	9/02/2023 8:47 PM	File folder	

Name	Date modified	Type	Size
1. Research_Planning_and_Design	9/02/2023 8:47 PM	File folder	
2. Data_Analysis_and_Reporting	11/02/2023 11:45 PM	File folder	
3. Data_Analysis_and_Reporting_DSMB_1	9/02/2023 8:48 PM	File folder	

Name	Date modified	Type	Size
1. Data_Management	9/02/2023 8:47 PM	File folder	
2. Source_Data	12/02/2023 2:01 PM	File folder	
3. Analysis_Data	12/02/2023 2:01 PM	File folder	
4. Programs	12/02/2023 2:00 PM	File folder	
5. Output	12/02/2023 2:01 PM	File folder	
6. Validation	9/02/2023 8:48 PM	File folder	
7. Report	12/02/2023 2:03 PM	File folder	

Conducting reproducible research in Stata

- Dynamic document commands – dyndoc, markdown, Stata dynamic tags (<<dd ...>>)
  - Create text files, Word documents, HTML files with Stata output
  - Use the Markdown text-formatting language to customise your report
- put* commands – putdocx, putexcel, putpdf
  - Create Word documents, Excel files, PDF documents with Stata output

Image source: https://www.istockphoto.com/photo/man-thinking-with-red-question-marks-gm518221379-49052030

Conducting reproducible research in Stata

1. Master do-file
2. Task specific do-files
3. Create a report in Microsoft Word

Image source: https://www.istockphoto.com/photo/man-thinking-with-red-question-marks-gm518221379-49052030

Discussion and Questions

- Challenges for conducting reproducible research
- Tools for reproducible research
- General tips

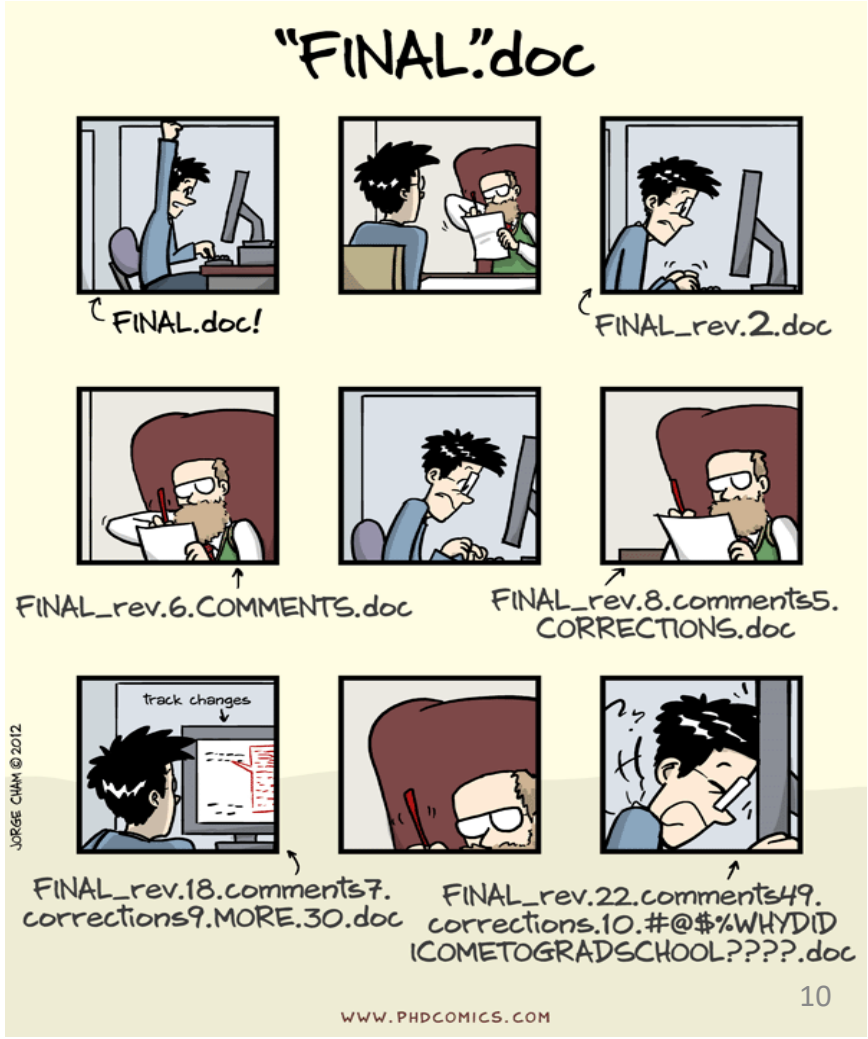


Image source: https://phdcomics.com/comics/archive_print.php?comicid=1531

THE UNIVERSITY OF MELBOURNE

Faculty of Medicine, Dentistry and Health Sciences

Thank you!

Webinar two: R Markdown

Wednesday February 22nd- 4-5pm (AEDT)

Australian Clinical Trials Alliance | STInG STATISTICS IN TRIALS INTEREST GROUP

This session will introduce R Markdown as a tool to increase efficiency and reproducibility in statistical analyses and reporting, using an example of a clinical trial.

In this presentation, Kristy will explain and provide a live demonstration on how to create a word document containing R analyses to share with other researchers. This will include how to provide a typical table one of baseline characteristics by treatment and reporting of models and provide figures, using Cox regression and Kaplan Meier curves as an example. All code will be freely available after the webinar.

Presenter- Dr Kristy Robledo

Dr Kristy Robledo is a senior research fellow in Biostatistics at the NHMRC Clinical Trials Centre, University of Sydney. She has over 14 years of experience in clinical research and provides leadership in her role in the design and analysis of large-scale randomised clinical trials in neonatology, cardiovascular and oncology. She also works on statistical methodology, particularly focused on the analysis of biomarker data. She is passionate about sharing and teaching coding, particularly to increase efficiency and reproducibility in clinical trials research.