A process for independent reproduction of computations underlying research

# Open Reproducible Research

**Reproducible research** refers to achieving the <u>same</u> results (e.g., tables, figures, numbers) as reported in the paper by using the <u>same</u> source code and data. In **Open reproducible research**, these materials are <u>publicly</u> accessible.

Goodman, S. N., Fanelli, D., & Ioannidis, J. P. (2016). What does research reproducibility mean? *Science Translational Medicine* 8(341), 341ps12-341ps12. https://doi.org/10.1126/scitranslmed.aaf5027

# The CODECHECK philosophy

- Reproducibility tools like Code Ocean and Binder set the bar high by making code reproducible for everyone.

- CODECHECK simply asks: "Was the code reproducible once for someone else?"

- CODECHECK checks if the code runs and generates the expected number of output files.

- The contents of those output files are checked visually and available for others to see.

- The validity of the code is not checked.

# CODECHECK process in four steps

Step 1: Provide me with access to code and data underlying the results in the manuscript.

# CODECHECK process in four steps

Step 1: Provide me with access to code and data underlying the results in the manuscript.

Step 2: Reproducibility attempt: Executable? Same results?

# CODECHECK process in four steps

Step 1: Provide me with access to code and data underlying the results in the manuscript.

Step 2: Reproducibility attempt: Executable? Same results?

Step 3: Create certificate, including how I ran the code, what was checked, and a copy of the outputs (plus differences), recommendations.

# CODECHECK process in four steps

Step 1: Provide me with access to code and data underlying the results in the manuscript.

Step 2: Reproducibility attempt: Executable? Same results?

Step 3: Create certificate, including how I ran the code, what was checked, and a copy of the outputs (plus differences), recommendations.

Step 4: Publication of certificate on Zenodo and a badge for your code repository.

CODE WORKS ✓

# Example

https://doi.org/10.5281/zenodo.5106408

# When is the best time to CODECHECK?



Develop idea/ hypothesis → Design study → Collect data → Analyse data → Store data + results → Write report → Submit report → Peer review → Publish → Reuse → Develop idea/ hypothesis

CODE CHECK ✓
https://codecheck.org.uk/
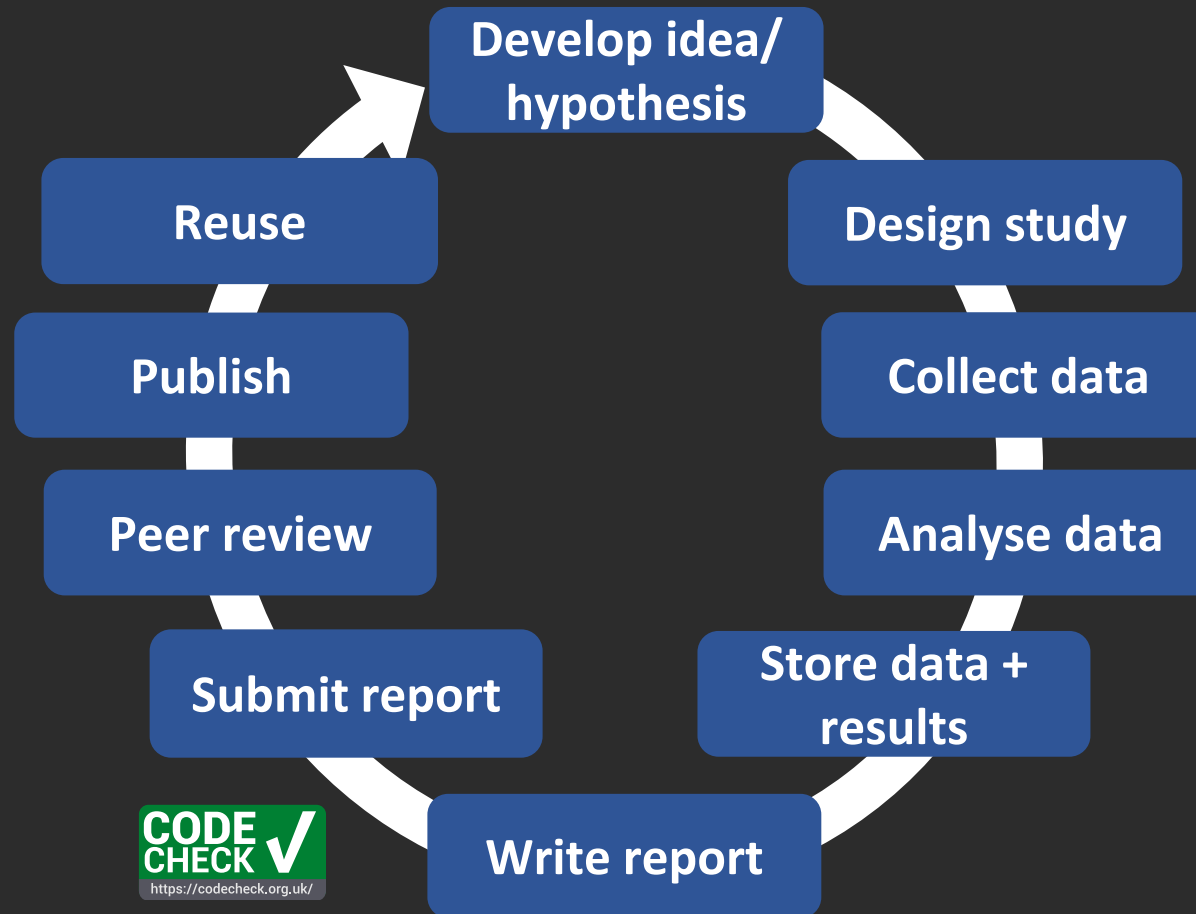
# How to prepare for a CODECHECK

**At least:  just write a bit of documentation.**

A bit more? create a so-called Research Compendium

- Create a folder that includes all materials necessary to reproduce the results
  - Manuscript
  - Data
  - Source code
  - Ideally in a notebook format (e.g., R Markdown, Jupyter notebook)
- Add a license!

# Further general remarks

- Everyone at ITC can request a CODECHECK
- Effort you need to invest is mainly about communication
- Unsuccessful CODECHECKs have no implications ("safe space")
- Issues in the code will not damage your work
- Synthetic datasets can be used in case of privacy concerns
- Also possible with licensed software (if I can get a license)
- In case of big data, data subsets can be used to reduce computation time
- It is possible to have a CODECHECK after publication

# Get started

- For more information, visit https://www.itc.nl/research/open-science/codecheck/
- You want to start your first CODECHECK? Drop me a message: m.konkol@utwente.nl