

Reproducible Manuscripts

Alexander Mark Weber

UBC and BCCHR

2024-11-27

Bridge-In

Mentimeter

Scientific Method

Scientific Method

What is the scientific method (broadly)?

1. Define a question
2. Gather information and resources (observe)
3. Form an explanatory hypothesis
4. Test the hypothesis by performing an experiment and collecting data in a reproducible manner
5. Analyze the data
6. Interpret the data and draw conclusions that serve as a starting point for a new hypothesis
7. Publish results
8. Retest (frequently done by other scientists)

Scientific Method

7. Publish Results
8. Retest (frequently done by other scientists)

Problem

- In 2011, John Ioannidis¹ published

Essay

Why Most Published Research Findings Are False

John P. A. Ioannidis

- Why?
 - Studies are underpowered
 - Current incentives lead scientists to publish quantity over quality
 - No incentives for scientists to replicate other studies
 - More...

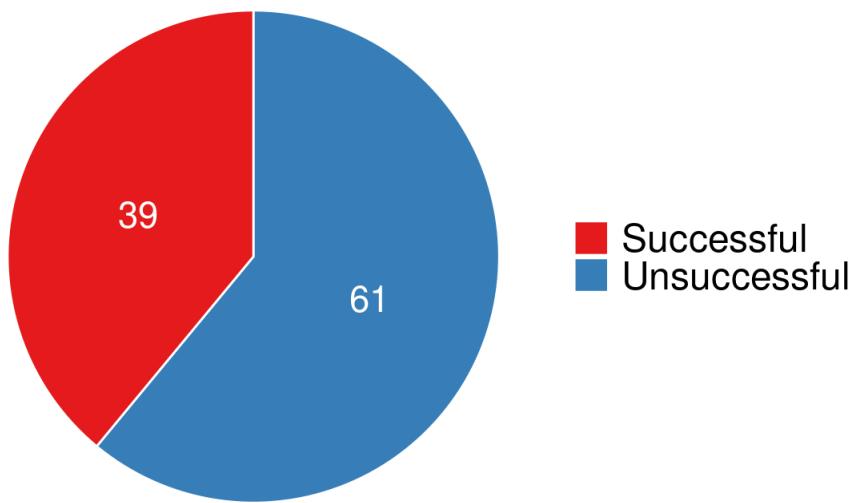
Problem

Was he right?

- In 2015, the Open Science Collaboration sampled studies from prominent journals to estimate the replicability of psychological research.²

Problem

Problem



Out of 100 independently performed replications, only 39% were subjectively labelled as successful replications, and on average, the effects were roughly half the original size ³

Problem

Not just in Psychology:

- animal behaviour⁴;
- cancer biology⁵;
- economics⁶
- pharmaceutical industry⁷
- neuroscience⁸
- neuroimaging⁹
- clinical trials¹⁰

Problem

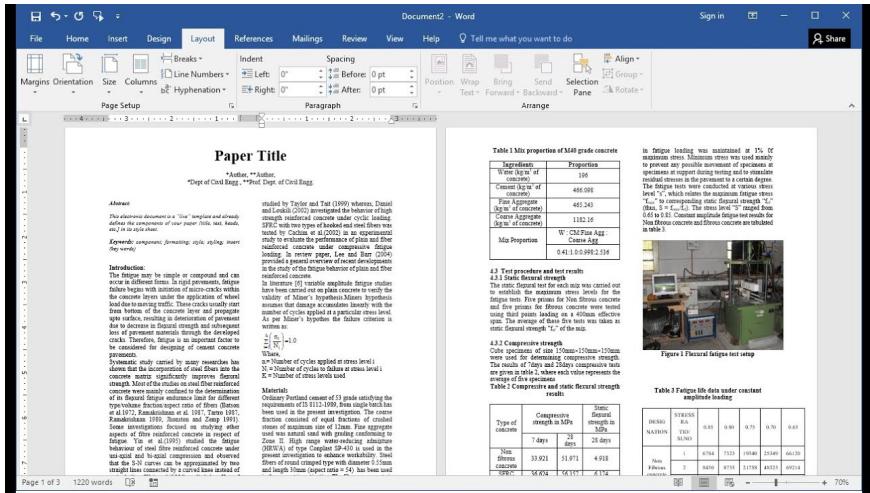
- For clinical trials: 44% contained at least some flawed data:¹¹
 - impossible statistics,
 - incorrect calculations,
 - or duplicated numbers or figures
 - 26% of trials were impossible to judge: either due to incompetence or faked data

Problem

- Publishing irreproducible results is worse than not publishing: more difficult to eliminate an idea than it is to introduce it¹²
- Spurious results can mislead other researchers who conduct follow-up investigations or try to integrate findings into broader theories.

What Can We Do?

- Many solutions are needed; far outside the scope of this talk
- One thing we can do is change the way we write papers.



- Currently, papers are written and published in a way that results in errors and the inability to computationally reproduce results.

What Can We Do?

- Errors: a 2016 paper by Nuijten et al.¹³ found that
 - nearly half of all papers had errors in them;
 - over 10% of p-values in published papers were inconsistent with the reported details of the statistical test
 - 1.6% were what they called “grossly” inconsistent, e.g. difference between the p-value and the test statistic meant that one implied statistical significance and the other did not

What Can We Do?

- Computational reproducibility: a 2021 paper by Hardwicke et al.¹⁴ attempted to reproduce results from 25 published papers that publicly shared their data and code:
 - found substantial numerical discrepancies between reported statistical values and values obtained from reproduction attempts in 64% of these papers

What Can We Do?

This is where Reproducible Papers come in...

Quarto_Chisep_CSVO2_Manuscript - main - RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

index.qmd gavin_thesis_markdown.qmd quarto.yml newdata

Source Visual

```
round(so(newdata$lnisep_intvein_CSVO2*100),2) & r signif(LLVSVO2.tsp.value,3) & r round(LLVSVO2.ts$cont.low*100,2) & r
269   \btrule
270   \end{tabular}
271   \footnotetext{QSM = quantitative susceptibility mapping; CI = confidence interval; SSS = superior sagittal sinus; CCV = central cerebral vein}
272 ...
273 Summary of acquired physiological parameters. Mean  $\pm$  SD is shown for chi and  $SvO_2$ - values. The P-value and 95% confidence interval (CI) were obtained through the comparison of values between QSM and paramagnetic maps; ( $n=19$ ).
274 ...
275
276 Region-specific  $\chi$  and  $SvO_2$ - values acquired from QSM were compared to values acquired from paramagnetic maps. In both the SSS and CCV, it was found that a significant difference existed between values acquired ( $\chi$  and  $SvO_2$ ) from QSM and paramagnetic maps ( $p < 0.05$ ). A boxplot showing the comparisons made is shown in #fig-methodplot.
277
278 {{< embed notebooks/gavin_thesis_markdown.qmd#fig-methodplot >}}
279
280 The acquired  $\chi$  and  $SvO_2$ - values were additionally compared between veins. In data created from QSM, a significant difference was found between the CCV and SSS in mean  $\chi$  ( $p < 0.05$ ; 95% CI [ $r$  round(regionQSMChi.ts$conf.low,3),  $r$  round(regionQSMChi.ts$conf.high,3)]) and mean  $SvO_2$  ( $p < 0.05$ ; 95% CI [ $r$  round(regionQSMsvo2.ts$conf.low,3),  $r$  round(regionQSMsvo2.ts$conf.high,3)]). In data acquired from paramagnetic maps, no significant difference was observed between the CCV and the SSS in either mean  $\chi$  ( $p = r$  round(regionChisepChi.ts$value,3); 95% CI [ $r$  round(regionChisepChi.ts$conf.low,3),  $r$  round(regionChisepChi.ts$conf.high,3)]) or mean  $SvO_2$  ( $p = r$  round(regionChisepSvo2.ts$value,3); 95% CI [ $r$  round(regionChisepSvo2.ts$conf.low,3),  $r$  round(regionChisepSvo2.ts$conf.high,3)]). A summary of these comparisons is represented in #fig-regionplot.
281
282 {{< embed notebooks/gavin_thesis_markdown.qmd#fig-regionplot >}}
283
284 # Discussion (#sec-discussion)
285
286 The primary objective of the present study was to assess whether the application of magnetic susceptibility separation to neonatal QSM data could provide more accurate cerebral venous oxygenation measurements. To the best of our knowledge, we are the first to test this in a neonatal cohort, as susceptibility separation has been typically evaluated as a method of imaging myelin and brain iron in adult subjects [@ShmueliSeparationMagneticSusceptibility2021; @AhmedDiagnosticComponentMap2023a]. Our results showed that the  $SvO_2$ - values of the SSS and CCV obtained from susceptibility separation are significantly lower than the respective  $SvO_2$ - values obtained from QSM alone. When our results were compared to the literature, we found that our SSS  $SvO_2$ - data from susceptibility separation agreed well with the findings of other studies measuring  $SvO_2$ - of the SSS in similar subject populations. Conversely, the paramagnetic CCV  $SvO_2$ - data saw less agreement with the existing literature than the corresponding data from QSM. However, there is reason to believe our paramagnetic CCV values may be accurate given their similarity to the paramagnetic SSS values and the limitations of the two studies that observed CCV  $SvO_2$ . Additionally, it is important to note that our  $SvO_2$ -
```

237:1 Results

Console Terminal Background Jobs

R 4.4.0 · ~/Dropbox/AssistantProf_BCCHR/Projects/Gavin_CSVO2/Quarto_Chisep_CSVO2_Manuscript/

Environment History Connections Build Git Tutorial

Files Plots Packages Help Viewer Presentation

Table of contents

- Introduction
- Methods
- Results**
- Discussion
- Conclusion
- References

C D

Paramagnetic Component Method QSM

SV_{O₂} (%)

Figure 3: Vein-specific susceptibility and oxygen saturation values by method. (A, B) contains violin plots comparing subject X (ppm) acquired from QSM (A) and its paramagnetic component (B); (C, D) contains violin plots comparing subject SvO₂ (%) acquired from QSM (C) and its paramagnetic component (D). Raw data points from paramagnetic maps are shown as filled green circles and raw data points from QSM are shown as filled orange circles. Each line connects the raw data points of a single subject. (****) indicates $P < 0.05$.

The acquired χ and SvO_2 values were additionally compared between veins. In data created from QSM, a significant difference was found between the CCV and SSS in mean χ ($p < 0.05$; 95% CI [0.017, 0.04]) and mean SvO_2 ($p < 0.05$; 95% CI [-0.052, -0.023]). In data acquired from paramagnetic maps, no significant difference was observed between the CCV and the SSS in either mean χ ($p = 0.711$; 95% CI [-0.02, 0.029]) or mean SvO_2 ($p = 0.752$; 95% CI [-0.034, 0.029]). A summary of these comparisons is represented in [Figure 4](#).

A B

Chi (ppm)

SvO₂ (%)

Learning Goals

Learning Goals

By the end of the talk, the audience should:

- Know what a reproducible manuscript is,
- Understand some reasons why scientists should be writing their manuscripts this way,
- Know what Markdown, Knitr, Pandoc, LaTeX, Jupyter Notebook, R/RMarkdown, and Quarto are,
- Know the basics of the syntax for Markdown, R and Quarto,
- See how to integrate author information, code, equations, tables, images, and citations
- Be able to start writing your next manuscript using Quarto Manuscripts

Introduction

What is a reproducible manuscript?

- Reports the scientific findings
- Provides all (or almost all) the necessary data, code, and methodologies required to create those findings (i.e. data, stats, figures, tables, etc.)
- Transparent and organized
- Enables others to **replicate** and **verify** the results of your study independently



What does it look like?

```
1  ---
2 title: "The Application of Magnetic Susceptibilit
3 titlerunning: "SVO2"
4 authors:
5   - name: Thomas Gavin Carmichael
6     orcid: 0009-0008-6849-5333
7     corresponding: false
8     email: tg.carmichael@outlook.com
9     roles:
10      - writing - original draft
11      - formal analysis
12      - methodology
13      - validation
14      - visualization
15     affiliations:
16       - ref: 1
17       - ref: 2
18     degrees:
19       - HBSc
20   - name: Alexander Rauscher
21     orcid: 0000-0002-1961-8252
22     email: rauscher@physics.ubc.ca
23     corresponding: false
24     roles:
25       - writing - review & editing
26     affiliations:
27       - ref: 3
28     degrees:
29       - PhD
30       - MSc
31   - name: Ruth E Grunau
32     orcid: 0000-0002-5428-9212
33     corresponding: false
34     email: r.grunau@mail.ubc.ca
35     roles:
36       - writing - review & editing
37       - funding acquisition
38     affiliations:
39       - ref: 2
40       - ref: 3
```

The Application of Magnetic Susceptibility Separation for Measuring Cerebral Oxygenation in Preterm Neonates

Thomas Gavin Carmichael^{1,2}, Alexander Rauscher³,
Ruth E Grunau^{2,3}, Alexander Mark Weber^{2,*}

¹, Integrated Sciences, The University of British Columbia, Vancouver,
BC, Canada.

², BC Children's Hospital Research Institute, The University of British
Columbia, Vancouver, BC, Canada.

³, Pediatrics, The University of British Columbia, Vancouver, BC, Canada.

*Corresponding author(s). E-mail(s): aweber@bccr.ca;
Contributing authors: tg.carmichael@outlook.com;
rauscher@physics.ubc.ca; r.grunau@mail.ubc.ca;

Abstract

Background: Quantitative susceptibility mapping (QSM), an magnetic resonance imaging (MRI) modality sensitive to deoxyhemoglobin, is a promising method for measuring cerebral oxygenation in human neonates. Paramagnetic sources, like deoxyhemoglobin, however, can be obscured by diamagnetic sources such as water and myelin. This study evaluated whether QSM images, or isolated paramagnetic components, are more accurate for measuring oxygenation of cerebral veins of preterm neonates, and explored oxygenation differences between the major cerebral veins.

Methods: 19 preterm neonates were scanned at term equivalent age on a 3T MRI using a multi-echo susceptibility-weighted imaging sequence. Susceptibility values were calculated from QSM images to determine oxygen saturation (SvO_2) in the superior sagittal sinus (SSS) and central cerebral veins (CCV).

What are some other benefits?

- Already mentioned:
 - reducing **errors** from copy-pasting results to paper
 - anyone can see how I obtained my results or figures by **reviewing my code** (bonus: learn how others made their figures!)
- Easy to restructure, rewrite, revise:
 - no need to tweak reported values, tables, or figures by hand
 - remove barrier to re-running analyses (thanks to Reviewer #2); speed up resubmission

What are some other benefits?

- easy cross-referencing and citations
- allow for more rigorous meta-analyses studies to be performed (raw-data)
- can be written in your favorite code editor (RStudio, VSCode, plain text editor, Neovim)
- render from `.qmd` to PDF, Word, HTML, *LATEX*, etc.
- execute code in R, Python, Julia, and more
- apply journal styles to your outputs with Quarto extensions
- publish to Github Pages, Netlify, and more

Notebooks: an Aside

Aside: What is a notebook?

A notebook is a document that contains both **code** and **narrative**:

- Jupyter notebooks
([.ipynb](#))



```
[5]: import matplotlib.pyplot as plt
plt.style.use('classic')
%matplotlib inline
import numpy as np
import pandas as pd
import seaborn as sns
sns.set()
```

```
[6]: rng = np.random.RandomState(0)
x = np.linspace(0, 10, 500)
y = np.cumsum(rng.randn(500, 6), 0)
```

Next step

Now, create a graph.

```
[7]: plt.plot(x, y)
plt.legend('ABCDEF', ncol=2, loc='upper left');
```



Aside: What is a notebook?

Aside: What is a notebook?

A notebook is a document that contains both **code** and **narrative**:

- Rmarkdown (.rmd)



The screenshot shows the RStudio interface with an Rmarkdown file open. The code in the editor is:

```
1 ---
2 title: "Viridis Demo"
3 output: html_document
4 ---
5
6 ```{r include = FALSE}
7 library(viridis)
8 ```
9
10 The code below demonstrates two color palettes in the [viridis](https://github.com/sjmgarnier/viridis) package. Each plot displays a contour map of the Maunga Whau volcano in Auckland, New Zealand.
11
12 ## Viridis colors
13
14 ```{r}
15 image(volcano, col = viridis(200))
16 ```
17
18 ## Magma colors
19
20 ```{r}
21 image(volcano, col = viridis(200, option = "A"))
22 ```

The R Markdown pane shows the command > render("1-example.Rmd", output_format = "word_document").
```

To the right of the editor, there are two plots. The top plot is titled "Viridis Demo" and shows a contour map of the Maunga Whau volcano using the Viridis color palette. The bottom plot is titled "Magma colors" and shows the same contour map using the Magma color palette.

Aside: What is a notebook?

Aside: What is a notebook?

A notebook is a document that contains both **code** and **narrative**:

- Quarto document
(.qmd)



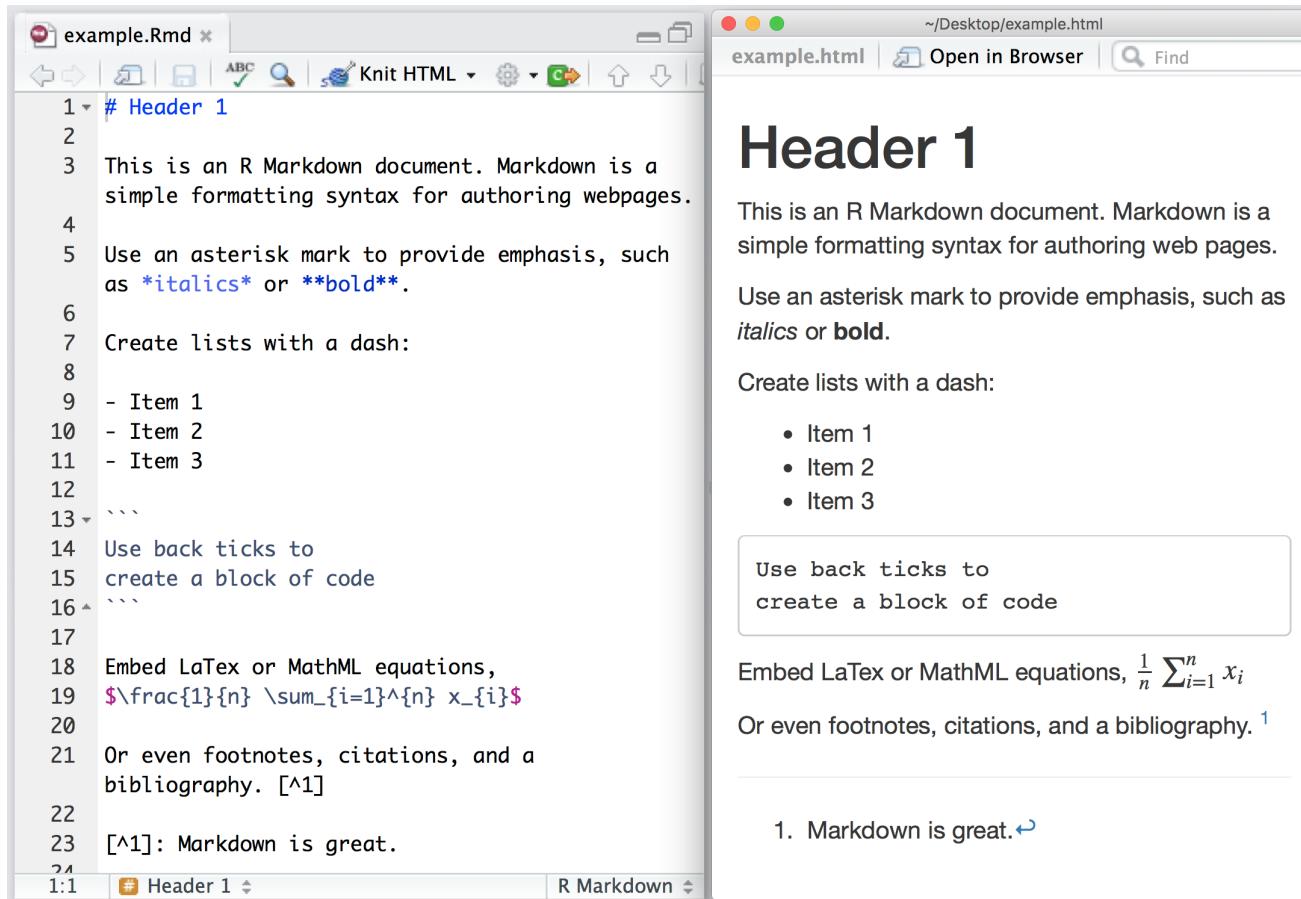
The screenshot shows the Quarto IDE interface. On the left, the code editor displays a .qmd file titled "Hello, Quarto". The code includes a YAML front matter section and R code to load the tidyverse and palmerpenguins packages. Below the code, a "Meet Quarto" section provides an introduction to Quarto, mentioning its ability to weave content and code. It also includes a small image of three penguins labeled "CHINSTRAP", "GENTOO", and "ADELIE". On the right, the rendered output shows a "Hello, Quarto" page with the same introductory text and penguin image. Further down, there is a section titled "Meet the penguins" with a scatter plot titled "Flipper and bill length" showing data for Palmer Archipelago penguins. The plot includes a legend for "Penguin species" with colored dots for Adelie (orange), Gentoo (blue), and Chinstrap (purple).

Aside: What are notebooks good for?

- By combining narrative with code, researchers can share and explain what they did, how they did it, and why they did it.
- Kind of like a research paper but with all the data, stats and computation baked in.
- Great for teaching, communicating, and/or collaborating: where you can directly see what someone did, with helpful explanations along the way.

Aside: Quick intro to R Markdown

Some markdown syntax



The image shows a comparison between an R Markdown source file and its resulting HTML output.

Left Panel (example.Rmd):

```
1 # Header 1
2
3 This is an R Markdown document. Markdown is a
4 simple formatting syntax for authoring webpages.
5
6 Use an asterisk mark to provide emphasis, such
7 as *italics* or **bold**.
8
9 Create lists with a dash:
10
11 - Item 1
12 - Item 2
13 - Item 3
14
15 Use back ticks to
16 create a block of code
17
18 Embed LaTex or MathML equations,
19 $\frac{1}{n} \sum_{i=1}^n x_i$
20
21 Or even footnotes, citations, and a
22 bibliography. [^1]
23
24 [^1]: Markdown is great.
```

Right Panel (~/Desktop/example.html):

Header 1

This is an R Markdown document. Markdown is a simple formatting syntax for authoring web pages.

Use an asterisk mark to provide emphasis, such as *italics* or **bold**.

Create lists with a dash:

- Item 1
- Item 2
- Item 3

Use back ticks to
create a block of code

Embed LaTex or MathML equations, $\frac{1}{n} \sum_{i=1}^n x_i$

Or even footnotes, citations, and a bibliography.¹

1. Markdown is great. ↪

Aside: Quick intro to R Markdown

Markdown with evaluated code

The image shows a split-screen view of RStudio. On the left is the R Markdown (.Rmd) file 'chunks.Rmd' with the following content:

```
1 R Code Chunks
2 =====
3
4 With R Markdown, you can insert R code
5 chunks including plots:
6
7 ````{r qplot, fig.width=4, fig.height=3,
8 message=FALSE}
8 # quick summary and plot
9 library(ggplot2)
10 summary(cars)
11 qplot(speed, dist, data=cars) +
12   geom_smooth()
13 ````
```

On the right is the 'Preview HTML' window showing the rendered output:

R Code Chunks

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

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

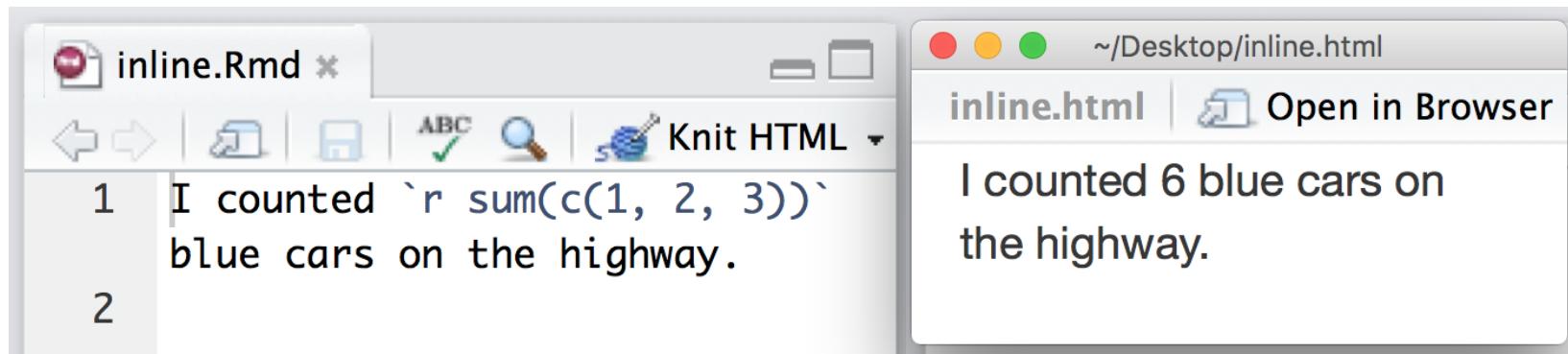
```
##      speed         dist
## Min.   : 4.0   Min.   : 2
## 1st Qu.:12.0   1st Qu.: 26
## Median :15.0   Median : 36
## Mean   :15.4   Mean   : 43
## 3rd Qu.:19.0   3rd Qu.: 56
## Max.   :25.0   Max.   :120
```

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

A scatter plot titled 'qplot(speed, dist, data = cars) + geom_smooth()' is displayed, showing the relationship between speed and distance. The x-axis is labeled 'speed' and ranges from 5 to 25. The y-axis is labeled 'dist' and ranges from 0 to 100. The plot includes a blue regression line and a light gray shaded area representing the confidence interval.

Aside: Quick intro to R Markdown

You can also include code inline (mixed in with the markdown text)



The image shows two side-by-side windows. On the left is a text editor window titled "inline.Rmd" containing R Markdown code. The code consists of two numbered lines: "1 I counted `r sum(c(1, 2, 3))` blue cars on the highway." and "2". Above the code area are standard file operations like back, forward, and save. Below the code are icons for "ABC" (auto-complete), a magnifying glass (search), and a knitting needle (Knit HTML). On the right is a web browser window titled "inline.html" with the URL " ~/Desktop/inline.html ". The browser displays the rendered content: "I counted 6 blue cars on the highway." This demonstrates how R Markdown allows for the inclusion of R code directly within the document's text.

```
1 I counted `r sum(c(1, 2, 3))` blue cars on the highway.
2
```

inline.html | Open in Browser

I counted 6 blue cars on the highway.

Overview of Quarto Manuscript

Project Files

- `index.qmd`: a notebook file where you write your article.
This file contains:
 - document metadata, including article front matter (authors, affiliations, etc.) and Quarto options,
 - the article body, written using special Quarto markdown syntax that allows you to add things like cross references and citations, and
 - optionally, code, where you control if, or how, the code and its output appear in the article.

Project Files

- `_quarto.yml`: a configuration file that identifies the project as a Quarto manuscript and controls how your manuscript is put together.

```
project:  
  type: manuscript  
  
execute:  
  freeze: auto  
  
format:  
  html:  
    toc: true  
  comments:  
    hypothesis: true  
  docx: default  
  jats: default  
  nature-pdf:  
    journal: "sn-nature"
```

index.qmd

The file `index.qmd` is a Quarto markdown file. It contains three types of content:

- Starts with a YAML header, used to set document metadata, including scholarly front matter. The YAML header starts and ends with a line of three dashes (---)

index.qmd

YAML header:

```
1  ---
2  title: La Palma Earthquakes
3  author:
4    - name: Steve Purves
5      orcid: 0000-0002-0760-5497
6      corresponding: true
7      email: steve@curvenote.com
8      roles:
9        - Investigation
10       - Project administration
11       - Software
12       - Visualization
13     affiliations:
14       - Curvenote
15   - name: Rowan Cockett
16     orcid: 0000-0002-7859-8394
17     corresponding: false
18     roles: []
19     affiliations:
20       - Curvenote
21 license: CC BY-SA 4.0
22 keywords:
23   - La Palma
24   - Earthquakes
25 date: '2022-05-11'
26 abstract: |
27   In September 2021, a significant jump in seismic activity on the island
28 keypoints:
```

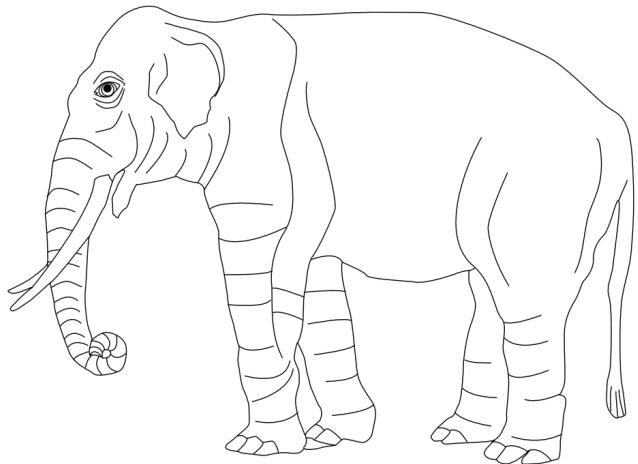

index.qmd

- `index.qmd` body may include executable code chunks: start with three backticks followed by the code language in curly braces (e.g. ````{r}` or ````{python}`).
- The rest of the document interpreted as Quarto specific markdown, allowing you to include figures, tables, equations, cross references and citations.

figures

![An elephant](/path/to/elephant.png){fig-elephant}

- Can be cross referenced with `@fig-elephant` in the document



An elephant

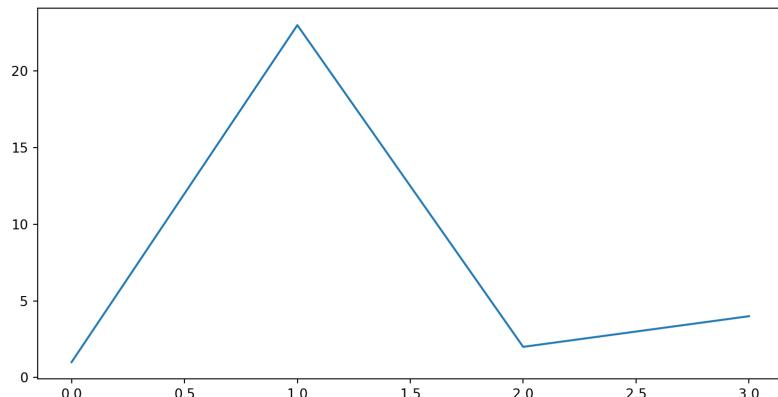
figures

- Can also be created using R or Python

```
```python
#| label: fig-plot
#| fig-cap: "Plot"
import matplotlib.pyplot as plt
plt.plot([1,23,2,4])
plt.show()
```

```

For example, see @fig-plot.



figures

- Can also embed from a `.qmd` or `.ipynb` file:
`{{< embed mycode.ipynb#fig-plot >}}`
- Which itself contains something like:

```
#| label: fig-plot
#| fig-cap: "Plot"
import matplotlib.pyplot as plt
plt.plot([1,23,2,4])
plt.show()
```

tables

- This is probably the most complex element of Quarto
- Easiest way is using markdown pipes (using RStudio Visual mode):

| Default | Left | Right | Center |
|---------|------|-------|--------|
| 12 | 12 | 12 | 12 |
| 123 | 123 | 123 | 123 |
| 1 | 1 | 1 | 1 |

: Demonstration of pipe table syntax {#tbl-numbers}

| Default | Left | Right | Center |
|---------|------|-------|--------|
| 12 | 12 | 12 | 12 |
| 123 | 123 | 123 | 123 |
| 1 | 1 | 1 | 1 |

tables

- You can also write your table in R or Python

```
#| label: tbl-planets
#| tbl-cap: Astronomical object

from IPython.display import Markdown
from tabulate import tabulate
table = [["Sun", "696,000", 1.989e30],
          ["Earth", "6,371", 5.972e24],
          ["Moon", "1,737", 7.34e22],
          ["Mars", "3,390", 6.39e23]]
Markdown(tabulate(
    table,
    headers=["Astronomical object", "R (km)", "mass (kg)"]
))
```

Table 1: Astronomical object

| Astronomical object | R (km) | mass (kg) |
|---------------------|---------|-----------|
| Sun | 696,000 | 1.989e+30 |
| Earth | 6,371 | 5.972e+24 |
| Moon | 1,737 | 7.34e+22 |
| Mars | 3,390 | 6.39e+23 |

tables

- Can also write tables in *LATEX* (but may not be converted to HTML well, and vice versa)

```
```{=latex}
\begin{table}[]
\begin{tabular}{@{}l l@{}}
\toprule
Header 1 & Header 2 \\ \midrule
A & B \\
C & C
\end{tabular}
\end{table}
```
```

| Header 1 | Header 2 |
|----------|----------|
| A | B |
| C | C |

references

- References can be accomplished using a `.bib` file
- Here is an example:

```
@article{ahmedDiamagneticComponentMap2023a,
    title = {The Diamagnetic Component Map from Quantitative Susceptibility
Mapping (\{QSM\}) Source Separation Reveals Pathological Alteration in
\{Alzheimer\}'s Disease-Driven Neurodegeneration},
    author = {Ahmed, Maruf and Chen, Jingjia and Arani, Arvin and Senjem,
Matthew L. and Cogswell, Petrice M. and Jack, Clifford R. and Liu, Chunlei},
    year = {2023},
    month = oct,
    journal = {NeuroImage},
    volume = {280},
    pages = {120357},
    issn = {1095-9572},
    doi = {10.1016/j.neuroimage.2023.120357},
    langid = {english},
    pmid = {37661080},
```

- referenced in the document with
`@ahmedDiamagneticComponentMap2023a`

in-line code

- You can also mix text with in-line code, in order to reference variables
- For example:

```
The mean SvO~2~ values for the SSS and the CCV were found to be `r  
| rndmean(newdata$Gavin_SSSVein_CSv02)` $\pm` r rndsd(newdata$Gavin_SSSVein_CSv02)`  
| and `r rndmean(newdata$Gavin_IntVein_CSv02)` $\pm` r  
| rndsd(newdata$Gavin_IntVein_CSv02)` ppm, respectively, when determined from the QSM  
| data. When determined from the paramagnetic map, the mean SvO~2~ values for the SSS  
| and the CCV were found to be `r rndmean(newdata$Chisep_SSSVein_CSv02)` $\pm` r  
| rndsd(newdata$Chisep_IntVein_CSv02)` %, respectively. A summary of the measured  
| physiological parameters, including the chi values used to calculate SvO~2~, can  
| found in @tbl-chistats.
```

The mean SvO_2 values for the SSS and the CCV were found to be 0.72 ± 0.03 and 0.69 ± 0.03 ppm, respectively, when determined from the QSM data. When determined from the paramagnetic map, the mean SvO_2 values for the SSS and the CCV were found to be 0.58 ± 0.07 %, respectively. A summary of the measured physiological parameters, including the chi values used to calculate SvO_2 , can found in [Table 3](#).

Simple Demonstration with Quarto

Let's write a manuscript

- First, install quarto <https://quarto.org/docs/download/prerelease.html>
- Approach 1: Start from scratch
 - Creating a Quarto manuscript
 - RStudio: New Project > New Directory > Quarto Manuscript
 - `quarto create project manuscript <name>`
 - Add manuscript content
- Approach 2: Start with a sample from <https://quarto.org/docs/manuscripts>

Let's write a manuscript

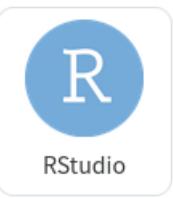
- We will then need to decide what tool we will use:



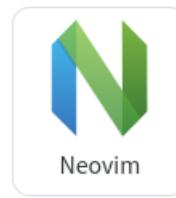
VS Code



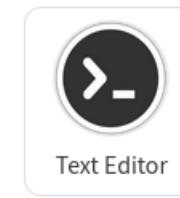
Jupyter



RStudio



Neovim



Text Editor

- Today I will use RStudio as an example

Let's write a manuscript

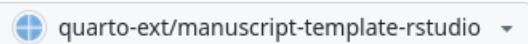
- Clone the Template Repository
 - Head to <https://github.com/quarto-ext/manuscript-template-rstudio/generate>

Create a new repository

A repository contains all project files, including the revision history. Already have a project repository elsewhere? [Import a repository](#).

Required fields are marked with an asterisk ().*

Repository template



Start your repository with a template repository's contents.

Include all branches

Copy all branches from quarto-ext/manuscript-template-rstudio and not just the default branch.

Owner *



weberam2

Repository name *

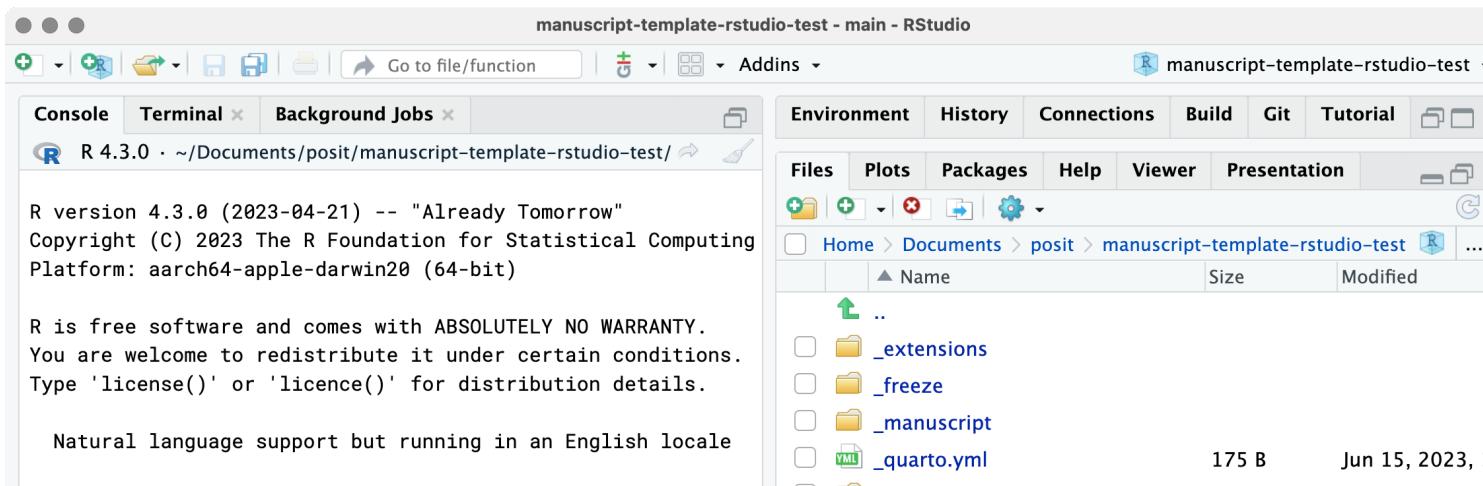
/manuscript-template-rstudio

 **manuscript-template-rstudio is available.**

Great repository names are short and memorable. Need inspiration? How about [effective-succotash](#) ?

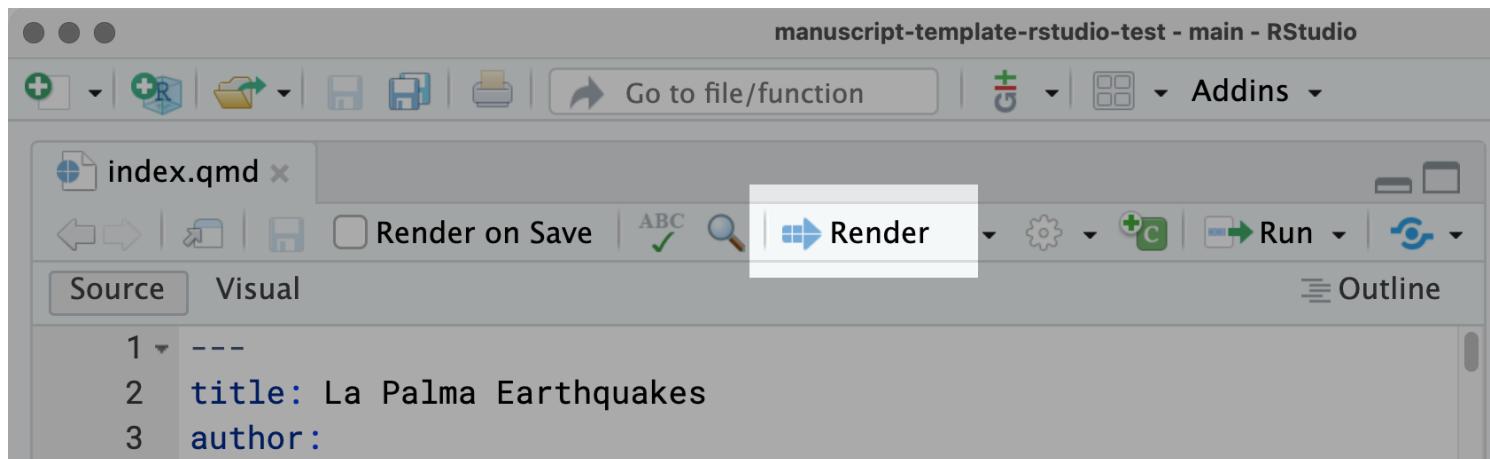
Let's write a manuscript

- Once your repository is created, clone it to your local computer.
- In RStudio, you can do: **File > New Project**.
- In the **New Project** dialog, select **From Version Control**, then **Git**, and copy and paste the repo URL from GitHub.



Workflow

- The basic workflow for writing a manuscript in Quarto is to make changes to your article content in `index.qmd`, preview the changes with Quarto, and repeat.
- Render and preview the manuscript by hitting the **Render** button located in the menu bar of the editor (RStudio in this case):



Let's write a manuscript

You'll see some output from Quarto in the Background Jobs pane and then a live preview will appear in the Viewer pane.

The screenshot shows the RStudio interface with a Quarto manuscript titled "index.qmd". The left pane displays the QMD source code, which includes metadata like title, author, and abstract, along with a timeline plot. The right pane shows a live preview of the manuscript, including sections for AUTHORS, AFFILIATION, PUBLISHED, ABSTRACT, and an INTRODUCTION. Below the preview is a timeline plot titled "Figure 1: Timeline of recent earthquakes on La Palma". The timeline spans from 1500 to 2000, with vertical tick marks indicating seismic events. The RStudio status bar at the bottom indicates the file is "Running".

manuscript-template-rstudio-test - main - RStudio

index.qmd x

Source Visual

1 ---
2 **title:** La Palma Earthquakes
3 **author:**
4 - **name:** Steve Purves
5 **orcid:** 0000-0002-0760-5497
6 **corresponding:** true
7 **email:** steve@curvenote.com
8 **roles:**
9 - Investigation
10 - Project administration
11 - Software
12 - Visualization
13 **affiliations:**
14 - Curvenote
15 - **name:** Rowan Cockett
16 **orcid:** 0000-0002-7859-8394
17 **corresponding:** false
18 **roles:** []
19 **affiliations:**
20 - Curvenote
21 **keywords:**
22 - La Palma
23 - Earthquakes
24 **abstract:** |
25 In September 2021, a significant jump in seismic activity on

6:22 La Palma Earthquakes Quarto

Console Terminal x Background Jobs x

Preview: index.qmd Running 42:28

Output created: index.html

Watching files for changes
GET: /index.html

manuscript-template-rstudio-test

Environment History Connections Build Git Tutorial

Files Plots Packages Help Viewer Presentation

La Palma Earthquakes

AUTHORS

Steve Purves

AFFILIATION

Curvenote

Rowan Cockett

AFFILIATION

Curvenote

PUBLISHED

June 23, 2023

ABSTRACT

In September 2021, a significant jump in seismic activity on the island of La Palma (Canary Islands, Spain) signaled the start of a volcanic crisis that still continues at the time of writing. Earthquake data is continually collected and published by the Instituto Geográfico Nacional (IGN). ...

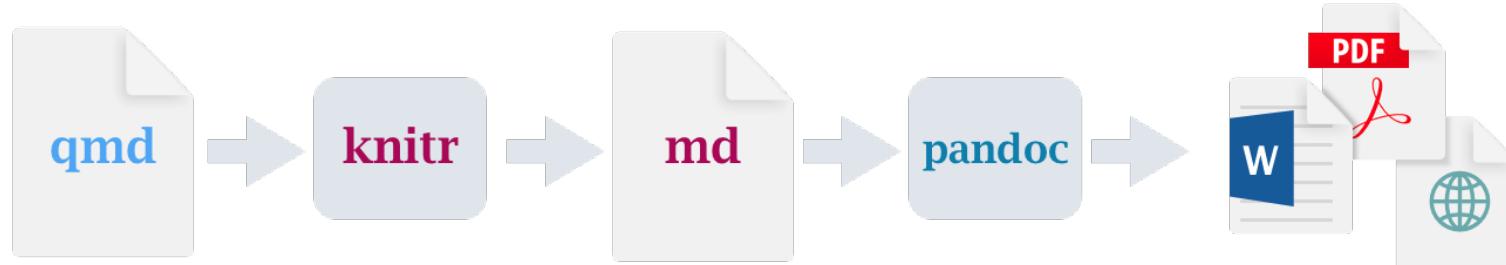
1 Introduction

Figure 1: Timeline of recent earthquakes on La Palma

1500 1600 1700 1800 1900 2000

Let's write a manuscript

What does Render do?



- **knitr** executes the code chunks and creates a new markdown (`.md`) document, which includes the code and its output
- the `.md` file is then processed by **pandoc**, which translates markdown/HTML/*LATEX* into various output formats

Let's write a manuscript

Open RStudio

[https://github.com/quarto-ext/manuscript-template-rstudio/
generate](https://github.com/quarto-ext/manuscript-template-rstudio/generate)

Real life demonstration

Open RStudio

What's Next?

Collaborating

- Most people will be happy to use the `.docx` and track changes
- Others may wish to publish their page on Github and use the annotation tool

The Application of Magnetic Susceptibility Separation for Measuring Cerebral Oxygenation in Preterm Neonates

AUTHORS
Thomas Gavin Carmichael, HBSc  Alexander Rauscher, PhD, MSc  Ruth E Grunau  Alexander Mark Weber, PhD, MSc 

AFFILIATIONS
The University of British Columbia
The University of British Columbia

PUBLISHED
June 26, 2024

ABSTRACT
Background: Quantitative susceptibility mapping (QSM) is a magnetic resonance imaging (MRI) modality proposed to be a viable method of measuring cerebral oxygenation in neonates given its sensitivity to deoxyhemoglobin, a paramagnetic molecule. During QSM, however, paramagnetic sources can be obscured by opposing diamagnetic sources such as water and myelin. We sought to evaluate whether QSM images alone, or an algorithm that attempts to isolate their paramagnetic components, are more accurate in measuring oxygenation of the major cerebral veins in a cohort of neonates born preterm. Additionally, we aimed to determine whether a difference in oxygenation existed between the major cerebral veins.

Methods: 19 neonates born preterm were scanned on a 3T research MRI at term equivalent age. The protocol included a multi-echo susceptibility-weighted imaging sequence. The acquired imaging data were processed as QSM images to

WeberLab     

Annotations 1 Page Notes

weberam2 Just now

ing (QSM) is a magnetic resonance imaging (MRI) modality proposed to be a viable method of measuring cerebral 

Some comments here

Collaborating

- Brave souls may wish to make changes to the `.tex` file and track changes with `latexdiff`

~~Draft~~Revision Title

Pratik Patel ~~and Another Author~~

February 9, 2013

This is an example of a ~~draft-revision~~ article. These are some types of changes to ~~expect~~be expected. Here is how it deals with equations:

$$y = \int_{\textcolor{blue}{0}}^{\textcolor{red}{x}} (x^2 + \textcolor{blue}{3}x) dx \quad (1)$$

When you do not include your collaborator's name in the document, they might get upset with you. But inclusion of their name in the final version will settle all scores.

Publishing

- Journal will usually want the file in `.docx`, sometimes they accept `.tex`, and rarely they will take `.pdf` for reviewing and then require the `.docx` or `.tex` file once accepted
- But `.pdf`s are static... how will people be able to see your source code once your paper is published?
- Publish your HTML manuscript on Github (or other alternatives) and cite your page in your pdf, perhaps in the **Data Availability** section
- Can add a **binder** using `quarto use binder` and including `code-links: binder` in your `_quarto.yml` file

Publishing

- Neuroibre
- Notebooks Now!
- eLife example

Criticisms

What are some barriers?

- Conversion to different outputs (.docx .html .pdf) is not always straightforward
- Getting complex tables to work in HTML and *LATEX*
- Steep learning curve
- No career incentive to try this new approach (i.e. “who has the time?”)
- Slow to preview/render (although you can set `execute: freeze: auto` so that computational documents are only re-rendered if their source file has changed)
- Errors/bugs can be frustrating (this is where a good text editor can be incredibly handy)

What are some criticisms?

- “But people will take my analysis, run it another way, and say my results aren’t valid”
- “This looks hard, I don’t want to have to learn *another* new thing”

What are some criticisms?

Summary

Summary

In this talk, we learned about

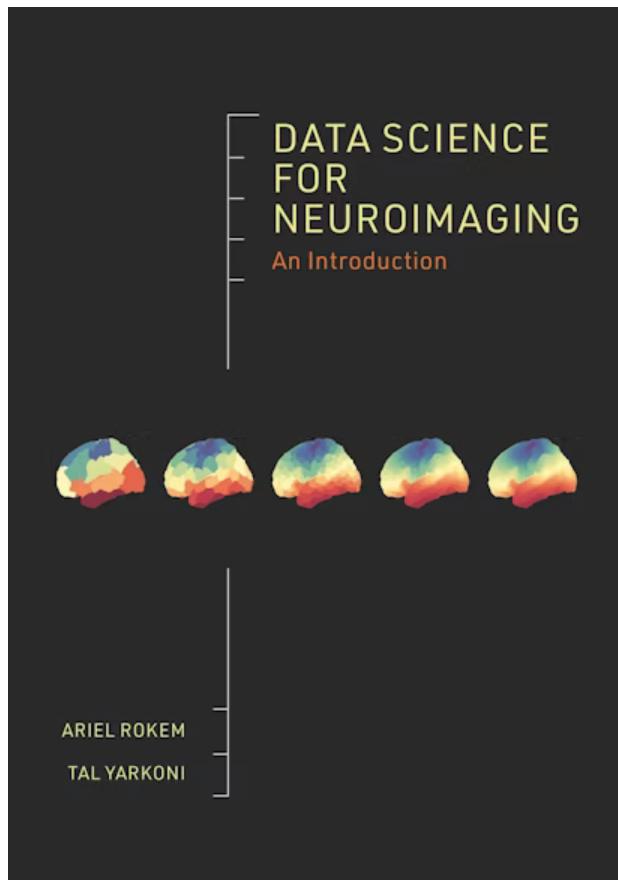
- What a reproducible manuscript is,
- Some reasons why scientists should be writing their manuscripts this way,
- What Markdown, Knitr, Pandoc, LaTeX, Jupyter Notebook, R/RMarkdown, and Quarto are,
- The basics of the syntax for Markdown, R and Quarto,
- How to integrate author information, code, equations, tables, images, and citations
- How to start writing your next manuscript using Quarto Manuscripts

Resources

-  <https://quarto.org/docs/get-started/>
-  <https://quarto.org/docs/guide/>
-  <https://quarto.org/docs/manuscripts>
-  <https://www.learnlatex.org/en/>
-  <https://www.markdownguide.org/>

Resources

-  <https://press.princeton.edu/books/hardcover/9780691222738/data-science-for-neuroimaging>



Thank you

⌚ <https://github.com/weberam2/>

ReproducibleManuscriptTalk

✉️ aweber@bcchr.ca

Ⓜ️ @weberam2@mastodon.social

🦋 @weberam2.bsky.social

𝕏 [delete your Twitter account](#)



Footnotes

1.
Ioannidis JPA (2005) Why most published research findings are false. PLoS Med 2(8): e124.
2.
Open Science Collaboration. Estimating the reproducibility of psychological science. Science 349, aac4716 (2015).
3. <https://www.nature.com/articles/s44271-023-00003-2>
4.
Farrar, B. G., Boeckle, M. & Clayton, N. S. Replications in comparative cognition: what should we expect and how can we improve? Anim. Behav. Cognit. 7, 1 (2020).
5.
Errington, T. M. et al. Investigating the replicability of preclinical cancer biology. Elife 10, e71601 (2021).
6.
Camerer, C. F. et al. Evaluating replicability of laboratory experiments in economics. Science 351, 1433–1436 (2016).
7.
Begley CG, Ellis LM (2012) Drug development: Raise standards for preclinical cancer research. Nature 483: 531–533. doi: 10.1038/483531a PMID: 22460880
8.
K.S. Button, J.P.A. Ioannidis, C. Mokrysz, B.A. Nosek, J. Flint, E.S.J. Robinson, M.R. Munafò. Power failure: Why small sample size undermines the reliability of neuroscience. Nat Rev Neurosci, 14 (2013), pp. 365-376
- 9.

