**1 Background**

Last week at the [MSACL](https://www.msacl.org/) conference [Dr. Keith Baggerly](https://faculty.mdanderson.org/profiles/keith_baggerly.html) from [MD Anderson Cancer Centre’s Bioformatics and Computational Biology Group](https://www.mdanderson.org/research/departments-labs-institutes/departments-divisions/bioinformatics-and-computational-biology.html) spoke about the importance of reproducible research using the [Duke University ovarian cancer biomarker scandal](http://www.nytimes.com/2011/07/08/health/research/08genes.html) as a backdrop. The talk…was…incredible and illustrated how easy it is to introduce catastrophic errors into your research papers through the use of GUI analytical tools. The now retracted article that Baggerly dismantled is [here](https://www.nature.com/articles/nm1491.epdf?referrer_access_token=GgTgb-wUTsKNRw8OVEiON9RgN0jAjWel9jnR3ZoTv0NW9rGpa6nJwwOU6XHtudagmmBJv2qoOFXJDOLGy16Gidwm_R8x_nqWC3-114LE5H3SScoWgmIQdaw7K2tHqMs9GpXdkbxm5UnBCXTU7ML2PxT0oPNvGTCRf77UNxBWFRsFi3ZFes90JzqRk43GrQU3wnzBirtqhyyW2gWDk0xLm8XrIKW7RAZF3B6m2iJBSWs%3D&tracking_referrer=www.nytimes.com). I urge everyone in our field to watch similar talks from Keith discussing biomarker analysis in [mass spectrometric proteomic data](https://www.youtube.com/watch?v=j1MT0oZqPXY) and [microarray data](https://www.youtube.com/watch?v=7gYIs7uYbMo). [Shannon Haymond](http://www.feinberg.northwestern.edu/faculty-profiles/az/profile.html?xid=17992) and I were then discussing how to make a submission in the field of clinical chemistry that is reproducible. While this article will not discuss the [basics of R and R Markdown](http://rmarkdown.rstudio.com/authoring_basics.html), it will serve as a guide for those who know a little about these and give you a working YAML header and get the citations and cross-references correct.

**2 Overhead**

**2.1 YAML**

RMarkdown articles require a [YAML](https://en.wikipedia.org/wiki/YAML) header to instruct R Markdown how to process your article. This is the YAML code that worked for me. I am sure there are other ways to do this:

---

title: "Reproducible Research: Write your Clinical Chemistry paper using R Markdown"

author:

- Daniel Holmes, MD

date: "February 04, 2018"

documentclass: "article"

header-includes:

- \usepackage{amsmath}

output:

#bookdown::pdf\_document2:

#bookdown::word\_document2

bookdown::html\_document2:

toc: no

pandoc\_args: [

"--csl", "clinical-chemistry.csl" , "--citation-abbreviations", "abbreviations.json"

]

bibliography: bibliography.bib

abstract: |

\*\*Background:\*\* Put

\*\*Methods:\*\* Your

\*\*Results:\*\* Abstract

\*\*Conclusions:\*\* Here

keywords: "bla bla bla"

---

Indenting and spacing really matter a lot in YAML so don’t mess with them. You can generate PDF, MS Word or HTML output as needed by uncommenting and commenting out the output type as appropriate in the YAML above.

**2.2 CSL Files**

[CSL files](http://citationstyles.org/) take care of citation formatting for you. Depending on what journal you are making submission to, you will need a different .csl file. They exist for every conceivable journal. In the world of non-reproducible reports, this process is taken care of by reference managers in GUI word processors but since GUI word processors do not produce reproducible research, we must break up with them.

From the YAML, you will see that you need a file called [“clinical-chemistry.csl”](https://github.com/citation-style-language/styles/blob/5e38d2aabba638edabfb9603388ea3d425d63fd1/dependent/clinical-chemistry.csl) which I downloaded from [here](https://github.com/citation-style-language/styles). Put this file in the same folder as your R Markdown file. The [Clinical Chemistry](http://clinchem.aaccjnls.org/) .csl depends on [the .csl file of the American Association for Cancer Research](https://github.com/citation-style-language/styles/blob/master/american-association-for-cancer-research.csl) but it will be downloaded for you on the fly. If you need a different reference format, search the [CSL GitHub repository](https://github.com/citation-style-language/styles) for the appropriate file.

**2.3 BibTeX**

Reference management in R Markdown is taken care of by [BibTeX](http://www.bibtex.org/). You can see from the YAML that we need a bibliography text file called “bibliography.bib”. You can name it whatever you like but you will need to change the YAML accordingly. In any case, any citation you intend to make will have to be in the .bib file. I am going to cite [Shannon Haymond](http://www.feinberg.northwestern.edu/faculty-profiles/az/profile.html?xid=17992) because this article was her idea and I will toss in a couple of other references so you can see that they get cited in order as we would like.

Below is my bibliography.bib file. I put it in the same folder as my R Markdown file. You make a .bib file using a text file editor (or RStudio) by cutting and pasting the BibTex citations from [Google Scholar](http://labrtorian.com/2018/02/05/reproducible-research-write-your-clinical-chemistry-paper-using-r/scholar.google.com)

@article{shannon2017,

title={Contribution of symmetric dimethylarginine to GFR decline in pediatric chronic kidney disease},

author={Brooks, Ellen R and Haymond, Shannon and Rademaker, Alfred and Pierce, Christopher and Helenowski, Irene and Passman, Rod and Vicente, Faye and Warady, Bradley A and Furth, Susan L and Langman, Craig B},

journal={Pediatric Nephrology},

pages={1--8},

year={2017},

publisher={Springer}

}

@article{li2017wellness,

title={Wellness Initiatives: Benefits and Limitations},

author={Li, Michelle and Diamandis, Eleftherios P and Paneth, Nigel and Yeo, Kiang-Teck J and Vogt, Henrik and Master, Stephen R},

journal={Clinical Chemistry},

volume={63},

number={6},

pages={1063--1068},

year={2017},

publisher={Clinical Chemistry}

}

@article{holmes2005preanalytical,

title={Preanalytical influences on DPC IMMULITE 2000 intact PTH assays of plasma and serum from dialysis patients},

author={Holmes, Daniel T and Levin, Adeera and Forer, Barry and Rosenberg, Frances},

journal={Clinical Chemistry},

volume={51},

number={5},

pages={915--917},

year={2005},

publisher={Clinical Chemistry}

}

**2.4 Journal Abbreviations**

Now, the fussiest thing I had to do was get the references abbreviating properly. We need an abbreviation database. Fortunately, I could download the abbreviation database from the Web of Science as a .csv file and then convert it to a JSON file. This was Stephen’s idea. I had to deal with a couple of badly behaving characters from some journal titles. This script, if embedded in your document, will download the .csv for you and then make the abbreviation database for you. That way your citations will say, “J Clin Pathol” and not “Journal of Clinical Pathology” etc.

```{r, echo = FALSE}

library(RJSONIO)

if(!file.exists("abbreviations.json")){

download.file("<https://ndownloader.figshare.com/files/5212423>","wos\_abbrev\_table.csv")

abbrev <- read.csv("wos\_abbrev\_table.csv", sep = ";", header = TRUE, stringsAsFactors = FALSE)

abbrev$full <- gsub("\\", "\\\\",abbrev$full, fixed = TRUE)

abbrev.list <- list('default' = list('container-title' = abbrev$abbrev.dots))

names(abbrev.list$default$`container-title`) = abbrev$full

write(toJSON(abbrev.list), "abbreviations.json")

rm(abbrev)

rm(abbrev.list)

}

```

**2.5 Citation Management**

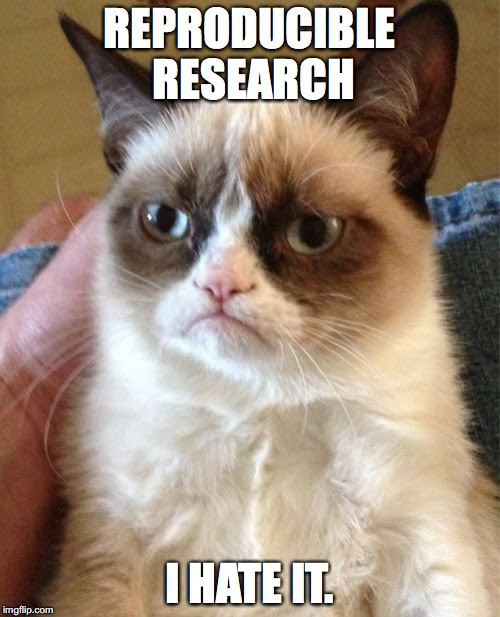
Now we can proceed to cite articles from our .bib file freely inserting syntax like this: [@shannon2017]. Shannon wrote this interesting article on symmetric dimethylarginine (1), Stephen wrote about wellness initiatives with Dr. Diamandis (2) and Dan wrote a paper about PTH when he was a resident (3). That makes for a total of 3 citations in this manuscript. (1–3)

**3 Reporting Your Results**

**3.1 Figures**

You can embed figures in R Markdown as local images or hyperlinks as follows:

![Grumpy Cat does not care about reproducibility](grumpy.jpg)



Grumpy Cat does not care about reproducibility

If you need to do cross-referencing of figures in your document which will change automatically for you if you insert another figure, you can do this by inserting your figure with an R code-chunk and giving the code-chunk a name.

```{r ancient-aliens, fig.height=3, fig.width=2, fig.cap="This is the ancient aliens guy.", echo = FALSE}

knitr::include\_graphics('ancient\_aliens.jpg')

```

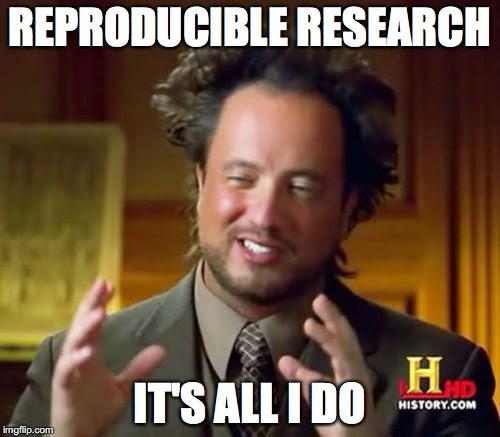


Figure 3.1: This is the ancient aliens guy.

Then you can reference your code chunk using syntax like this: See figure \@ref(fig:ancient-aliens)

and you will automatically create appropriately cross-referenced figures that get automatically numbered like this: See figure [3.1](http://labrtorian.com/2018/02/05/reproducible-research-write-your-clinical-chemistry-paper-using-r/#fig:ancient-aliens). Of course the hallmark of the reproducibility to embed R code right into the document. See for example figure [3.2](http://labrtorian.com/2018/02/05/reproducible-research-write-your-clinical-chemistry-paper-using-r/#fig:example-code)

```{r example-code, fig.cap = "This is a reproducible figure"}

set.seed(10)

x <- runif(100,0,100)

y <- x + rnorm(100,0,0.10)\*x

plot(x,y,

main = "Reproducible Figure",

pch = 16,

col = "blue",

xlab = "Current Method (mmol/L)",

ylab = "New Method (mmol/L)")

abline(lm(y~x), col = "red", pch = 2)

```

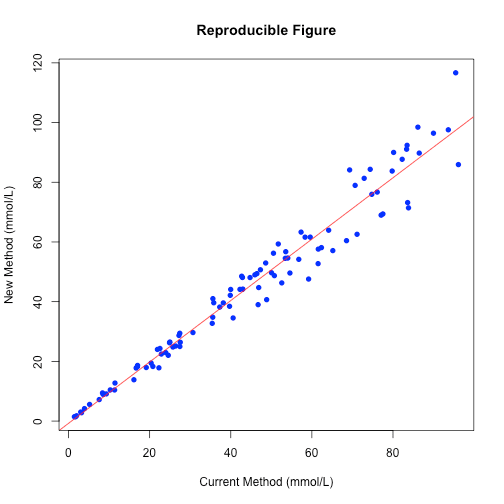


Figure 3.2: This is a reproducible figure

**3.2 Inline Calculations**

When you are reporting your amazing results you can have inline code calculations by syntax that looks like this: &#96r round(median(x),1)&#96 and would result in the median value of \(x\) being reported as 43.9 mmol/L.

**3.3 Tables**

Tables are not a problem either and can be made with the kable() function of the [knitr](https://yihui.name/knitr/) package or with the [xtable](https://cran.r-project.org/web/packages/xtable/index.html) package. Tables can be crossreferenced analogously to figures. See table [3.1](http://labrtorian.com/2018/02/05/reproducible-research-write-your-clinical-chemistry-paper-using-r/#tab:example-table)

```{r example-table, echo = FALSE, results = 'asis'}

library(knitr)

a <- 1:5

b <- 2:6

c <- a\*b

z <- data.frame(a,b,c)

kable(z,

caption = "This is a great table",

col.names = c("First","Second", "Third"))

```

| Table 3.1: This is a great table | | |
| --- | --- | --- |
| **First** | **Second** | **Third** |
| 1 | 2 | 2 |
| 2 | 3 | 6 |
| 3 | 4 | 12 |
| 4 | 5 | 20 |
| 5 | 6 | 30 |

**3.4 Math**

Math works pretty magically using \(\LaTeX\) syntax. For example inline math can be done like so $\sin^2x + \cos^2x = 1$. This will result in: \(\sin^2x + \cos^2x = 1\). And math can also be done as a code block like so:

$$

\oint\_S {E\_n dA = \frac{1}{{\varepsilon \_0 }}} Q\_{inside}

$$

Gauss’ Law says: \[  
\oint\_S {E\_n dA = \frac{1}{{\varepsilon \_0 }}} Q\_{inside}  
\]

Equations can be cross-referenced just like tables and figures.

**4 Conclusion**

I hope this makes writing a reproducible paper easier for you. A minimal template to produce output in PDF is [here](http://labrtorian.com/wp-content/uploads/2018/02/Template.zip). And the PDF output itself is [here](http://labrtorian.com/wp-content/uploads/2018/02/Rmd_paper_template.pdf). You’ll need \(\LaTeX\) installed of course.

**Parting Thought**

You can cite books too and get Greek letters too:

“Very truly I tell you,” Jesus answered, “before Abraham was born, \(\varepsilon\gamma\omega\) \(\varepsilon\iota\mu\iota\) (I Am)!” (4)

**References**

1. Brooks ER, Haymond S, Rademaker A, Pierce C, Helenowski I, Passman R, et al. Contribution of symmetric dimethylarginine to gfr decline in pediatric chronic kidney disease. Pediatr Nephrol. Springer; 2017;1–8.

2. Li M, Diamandis EP, Paneth N, Yeo K-TJ, Vogt H, Master SR. Wellness initiatives: Benefits and limitations. Clin Chem. Clinical Chemistry; 2017;63:1063–8.

3. Holmes DT, Levin A, Forer B, Rosenberg F. Preanalytical influences on DPC IMMULITE 2000 intact PTH assays of plasma and serum from dialysis patients. Clin Chem. Clinical Chemistry; 2005;51:915–7.

4. John the A, Chist J, Spirit H, God F. The Gospel According to John, 8:58. 1 Heavenly Way: Hosannah Press; 80AD.