RMarkdown Template for Managing Academic Affiliations

Also Deals with Cross References and
Reference Abbreviations for PDF Output

Duke A Caboom, MD[1] and Justin d'Ottawa, PhD[2]

[1]University of Tuktoyaktuk, CXVG+62 Tuktoyaktuk, Inuvik, Unorganized, NT Canada *
[2]University of Ottawa, 75 Laurier Ave E, Ottawa, ON K1N 6N5, Canada †

Abstract

Introduction: There's a big scientific problem out there. I know how to fix it.
Methods: My experiments are pure genius.
Results: Now you have your proof.
Conclusion: Give me more grant money.

1 Background

I like writing my academic papers in RMarkdown because it allows reproducible research. The cleanest way to submit a manuscript made in RMarkdown is using the LaTeX code that it generates using the YAML switch `keep_tex = true`. A minimalist YAML header would look like so:

```
---
title: The document title
author:
  - Duke A Caboom, MD
  - Justin d'Ottawa, PhD:
output:
  pdf_document:
    keep_tex: true
---
```

2 Introduction

However, when you want mutliple authors affiliations you discover that you can't do as you would in LaTeX because Pandoc does not know what to do with the affiliations and you end out a dishearting PDF that looks like the output shown in figure 1 below:

The situation worsens if you want MS-Word output. As those of us in medical fields know, most journals, with some notable exceptions like the Clinical Mass Spectrometry Journal and other Elsevier journals like Clinical Biochemistry and Clinica Chimica Acta, require submission of a document in MS-Word format which goes against all that Data Science and Reprodicible Research stands for. Parenthetically, it is my hope that

---

*dtholmes@mail.ubc.ca
†justin@uottawa.ca

The document title

true
true

Abstract

Introduction: There's a big scientific problem out there. I know how to fix it.
Methods: My experiments are pure genius.
Results: Now you have your proof.
Conclusion: Give me more grant money.

1   Introduction:

Figure 1: This is so sad.

since AACC has indicated to make Data Science a strategic priority for Lab Medicine, that they will soon accept submissons to Clinical Chemistry and Journal of Applied Laboratory Medicine written reproducibly in RMardown or LaTeX.

In the mean time, here are the workarounds for getting the affiliations to display correctly along with all the other stuff we want, namely, cross referencing of figures and tables and correct reference formatting and abbreviation of journal names. This allows you to avoid the horror of manually fixing your Word document after it generated from RMarkdown. In any case, let's start with MS-Word.

3   Dependencies for MS-Word and the Associated YAML

It goes without saying that you need to install LaTeX. LaTeX markup language is available here: Mac, Windows. For Linux, just install from the command line with your package manager. Do a full install with all the glorious bloat of all LaTeX packages. This saves many headaches in the future.

You will also need to install Pandoc which is the Swiss Army Knife of document conversion. It's going to turn LaTeX code into a .docx file for you. Mac users can do this with Homebrew on the terminal command line:

```
brew install pandoc
brew install pandoc-citeproc
brew install pandoc-crossref
```

There are some extra installs required to help Pandoc do its job. Install the prebuilt binaries if you can.

- pandoc-citeproc
- pandoc-crossref

Finally, you need to use some scripts written in the Lua scripting language which means you will need the language itself: + lua language

And you will need two Lua scripts:

These are in Pandoc github repository:

- scholarly-metadata.lua
- author-info-blocks.lua

You want the files named scholarly-metadata.lua and author-info-blocks.lua.

You will need to choose a .csl file for your journal. This will tell Pandoc how to format the references. You can download the correct .csl file here. You will also need a journal abbreviations database. I have made one

for you from the Web of Science list and you can download it here.

You will need to creat a .bibtex database which is just your list of references. This can be exported from various reference managers or built by hand. Name the file `mybibfile.bib`.

Now follow the bouncing ball:

1. Go to the directory containing your .Rmd file.
2. Create a directory in it called "Extras"
3. Put the two Lua scripts, the Bibtex database, the abbreviations database and the .csl file into the "Extras" folder.
4. If you want to avoid Pandoc's goofy default .docx formatting, then put this word document in the same folder.

OR

Download the contents of this folder from my github folder that has everything set up as I describe above.

For two authors, your YAML will need to look like this:

```
title: |
  RMarkdown Template for Managing
  Academic Affiliations
subtitle: |
  Also Deals with Cross References and
  Reference Abbreviations for MS-Word Output
author:
  - Duke A Caboom, MD:
      email: duke.a.caboom@utuktoyaktuk.edu
      institute: [UofT]
      correspondence: true
  - Justin d'Ottawa, PhD:
      email: justin@neverready.ca
      institute: [UofO]
      correspondence: false
institute:
  - UofT: University of Tuktoyaktuk, CXVG+62 Tuktoyaktuk, Inuvik, Unorganized, NT Canada
  - UofO: University of Ottawa, 75 Laurier Ave E, Ottawa, ON K1N 6N5, Canada
abstract: |
  **Introduction**: There's a big scientific problem out there. I know how to fix it.
  **Methods**: My experiments are pure genius.
  **Results**: Now you have your proof.
  **Conclusion**: Give me more grant money.
journal: "An awesome journal"
date: ""
toc: false
output:
 bookdown::word_document2:
    pandoc_args:
      - --csl=Extras/clinical-biochemistry.csl
      - --citation-abbreviations=Extras/abbreviations.json
      - --filter=pandoc-crossref
      - --lua-filter=Extras/scholarly-metadata.lua
      - --lua-filter=Extras/author-info-blocks.lua
      - --reference-doc=Extras/Reference_Document.docx
bibliography: "Extras/mybibfile.bib"
keywords: "CRAN, R, RMarkdown, RStudio, YAML"
```

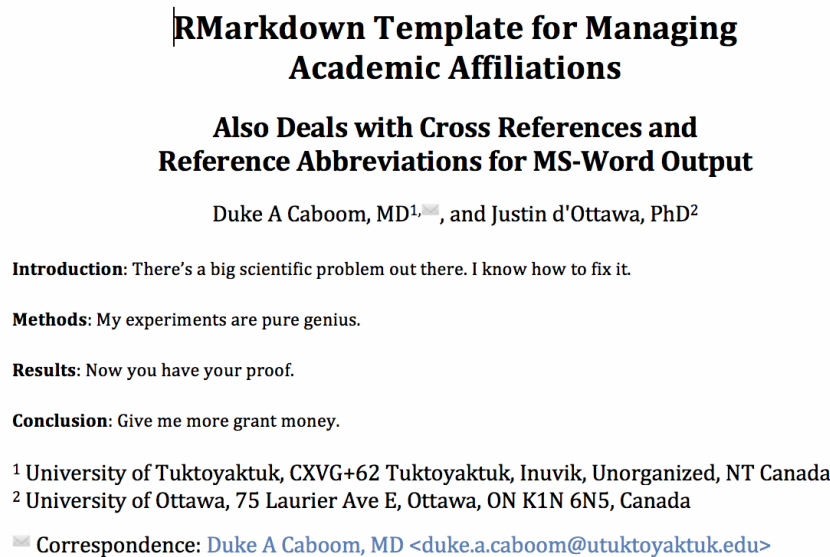Et voila! Figure 2 shows that we have something reasonable.

RMarkdown Template for Managing Academic Affiliations

Also Deals with Cross References and Reference Abbreviations for MS-Word Output

Duke A Caboom, MD[1,✉], and Justin d'Ottawa, PhD[2]

Introduction: There's a big scientific problem out there. I know how to fix it.

Methods: My experiments are pure genius.

Results: Now you have your proof.

Conclusion: Give me more grant money.

[1] University of Tuktoyaktuk, CXVG+62 Tuktoyaktuk, Inuvik, Unorganized, NT Canada
[2] University of Ottawa, 75 Laurier Ave E, Ottawa, ON K1N 6N5, Canada

✉ Correspondence: Duke A Caboom, MD <duke.a.caboom@utuktoyaktuk.edu>

Figure 2: This is so great

4 Dependencies for LaTeX and the Associated YAML

You don't need the lua scripts for LaTeX although you *can* use them. The issue with LaTeX is that the .tex template that Pandoc uses for generating LaTeX files does not support author affiliations as descibed in the Pandoc documentation. So what you need to do is modify the Pandoc LaTeX template. To get your current working copy of the Pandoc LaTeX template open up a terminal (Mac/Linux) and type:

```
pandoc -D latex > mytemplate.tex
```

This will push the contents to a file. Move the file to the "Extras" folder discussed above. If that seems difficult, you can also download it here. Now you have to edit it. Open it up in a text editor and find the section that reads:

```
$if(author)$
\author{$for(author)$$author$$sep$ \and $endfor$}
$endif$
```

Replace this with this code that will invoke the LaTeX authblk package.

```
$if(author)$
    \usepackage{authblk}
    $for(author)$
        $if(author.name)$
            $if(author.number)$
                \author[$author.number$]{$author.name$}
            $else$
                \author[]{$author.name$}
            $endif$
            $if(author.affiliation)$
                $if(author.email)$
```

```
                    \affil{$author.affiliation$ \thanks{$author.email$}}
                $else$
                    \affil{$author.affiliation$}
                $endif$
            $endif$
            $else$
            \author{$author$}
        $endif$
    $endfor$
$endif$
```

Then make your YAML header look like this:

```
---
title: |
  RMarkdown Template for Managing
  Academic Affiliations
subtitle: |
  Also Deals with Cross References and
  Reference Abbreviations for PDF Output
author:
- name: Duke A Caboom, MD
  affiliation: University of Tuktoyaktuk, CXVG+62 Tuktoyaktuk, Inuvik, Unorganized, NT Canada
  email: dtholmes@mail.ubc.ca
  number: 1
- name: Justin d'Ottawa, PhD
  affiliation: University of Ottawa, 75 Laurier Ave E, Ottawa, ON K1N 6N5, Canada
  email: justin@neverready.ca
  number: 2
abstract: |
  **Introduction**: There's a big scientific problem out there. I know how to fix it.

  **Methods**: My experiments are pure genius.

  **Results**: Now you have your proof.

  **Conclusion**: Give me more grant money.
toc: false
output:
  bookdown::pdf_document2:
    pandoc_args:
      - --filter=pandoc-crossref
      - --csl=Extras/clinical-biochemistry.csl
      - --citation-abbreviations=Extras/abbreviations.json
      - --template=Extras/mytemplate.tex
bibliography: "Extras/mybibfile.bib"
keep-latex: true
```

And as you can see in 3 you get a correctly list of authors.

4.1 Cross Reference of a Table

Of course, tables can be cross referenced in the same manner as figures. Here is a cross reference to table 1 using the code `\@ref(tab:mytable)` .

RMarkdown Template for Managing

Academic Affiliations

Also Deals with Cross References and
Reference Abbreviations for PDF Output

Duke A Caboom, MD[1] and Justin d'Ottawa, PhD[2]

[1]University of Tuktoyaktuk, CXVG+62 Tuktoyaktuk, Inuvik, Unorganized, NT Canada *
[2]University of Ottawa, 75 Laurier Ave E, Ottawa, ON K1N 6N5, Canada †

Abstract

Introduction: There's a big scientific problem out there. I know how to fix it.
Methods: My experiments are pure genius.
Results: Now you have your proof.
Conclusion: Give me more grant money.

Figure 3: This is also great.

Table 1: A short table

term	estimate	std.error	statistic	p.value
(Intercept)	36.908	2.191	16.847	0.000
hp	-0.019	0.015	-1.275	0.213
cyl	-2.265	0.576	-3.933	0.000

4.2 This Template also Takes Care of Reference Abbreviation.

As usual, you can make a citation with the code `[@bibtexname]`, where `bibtexname` is the articles's abbreviated handle in your bibtex database. Here is a great resource on the bookdown package [1] and reproducible research [2] and here are references where the journal title is longer [3,4]. The references in your documnent (and shown below) will have appropriate abbreviations based on the .json abbreviations database I have provided. In this case, I have chosen the .csl file for Clinical Mass Spectrometry–'cause MSACL.

5 Other Ways to Skin the YAML Cat

I came across some other ways to deal with this that I did not like as much but they are simpler. Here is one using a footnote.

```
title: The document title
author:
- [Duke A Caboom, MD]^(University of Tuktoyaktuk, CXVG+62 Tuktoyaktuk, Inuvik, Unorganized, NT Canada)
- [Justin d'Ottawa, PhD]^(University of Ottawa, 75 Laurier Ave E, Ottawa, ON K1N 6N5, Canada)
output: pdf_document
```

And you can also misuse the `date` variable:

```
title: The document title
author:
- Duke A Caboom, MD [1]
- Justin d'Ottawa, PhD [2]
date: 1. University of Tuktoyaktuk, CXVG+62 Tuktoyaktuk, Inuvik, Unorganized, NT Canada \newline 2. Uni
output: pdf_document
```

6 Conclusion

This concludes my long personal struggle to get a completely reproducible .docx manusript genereated by RMardown and Pandoc.

6.1 Parting Thought

Let us not become weary in doing good, for at the proper time we will reap a harvest if we do not give up.

Galations 6:9

References

[1] Y. Xie, J.J. Allaire, G. Grolemund, R markdown: The definitive guide, Chapman; Hall/CRC, 2018. https://bookdown.org/yihui/bookdown.

[2] R.D. Peng, Reproducible research in computational science, Science. 334 (2011) 1226–1227.

[3] G. Eisenhofer, C. Durán, T. Chavakis, C.V. Cannistraci, Steroid metabolomics: Machine learning and multidimensional diagnostics for adrenal cortical tumors, hyperplasias, and related disorders, Curr. Opin. Endocr. Metab. Res. 8 (2019) 40–49. doi:https://doi.org/10.1016/j.coemr.2019.07.002.

[4] F.B. Vicente, D.C. Lin, S. Haymond, Automation of chromatographic peak review and order to result data transfer in a clinical mass spectrometry laboratory, Clin. Chim. Acta. 498 (2019) 84–89. doi:https://doi.org/10.1016/j.cca.2019.08.004.