



Theory to Practice Harnessing R to Enhance Workflows

Matthew Brierley
Joana Marques Barros





Congratulations! You've completed your **Intro to R** programming course.



You are **eager** to show off your new `{dplyr}` skills!



But - oh, no! Your study is not part of the R pilot project.

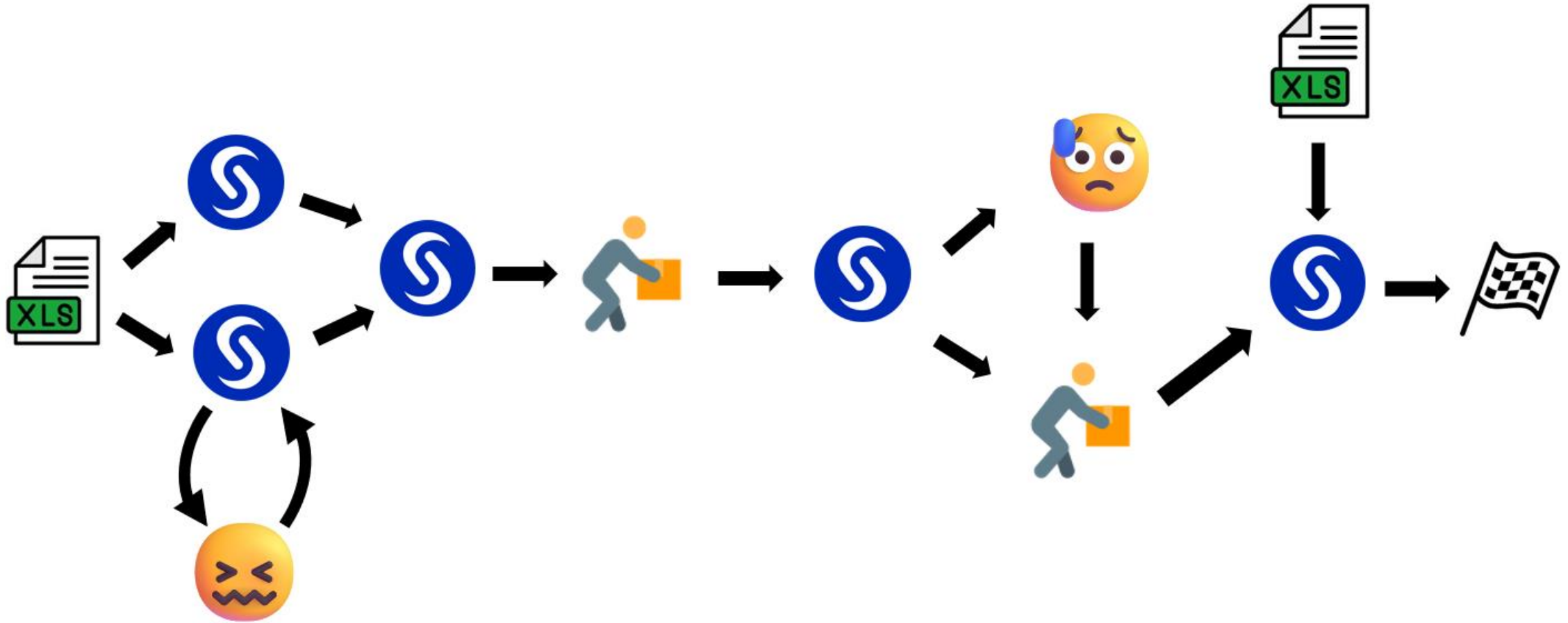
Our aim is to spark your curiosity and stimulate ideas to keep your current enthusiasm for R alive!

Outline

- 💡 Bringing Order to **Configuration** Chaos with **YAML**
- 💡 Improving **User Experience** with **{cli}**
- 💡 Quick Wins with **Quarto**

Case Study: Releasing Statistical Outputs

Old way



Case Study: Releasing Statistical Outputs

New way



What is YAML?

```
1 ♥ main:
2   project: wonderdrug
3   study: trial_001
4   delivery: CSR
5   blinded: True
6 ♥ team:
7   trial_statistician: "Penelope Value"
8   programmers: [Mack Rose, Libby Path]
9 ♥ data_managers:
10   - Lotta Outliers
11   - Audrey Trail
12   - Paige Missing
13 |
```

- **Simple**, text-based format to represent data structures
- Key-value pairs, lists, and nested objects
- Ideal for **configuration** information

Reading .yaml Files in R

```
1 ▾ main:
2   project: wonderdrug
3   study: trial_001
4   delivery: CSR
5   blinded: True
6 ▾ team:
7   trial_statistician: "Penelope Value"
8   programmers: [Mack Rose, Libby Path]
9 ▾ data_managers:
10   - Lotta Outliers
11   - Audrey Trail
12   - Paige Missing
13
```

```
> library(yaml)
> config <- read_yaml("yaml_example.yaml")
> str(config)
List of 2
 $ main:List of 4
  ..$ project : chr "wonderdrug"
  ..$ study   : chr "trial_001"
  ..$ delivery: chr "CSR"
  ..$ blinded : logi TRUE
 $ team:List of 3
  ..$ trial_statistician: chr "Penelope Value"
  ..$ programmers       : chr [1:2] "Mack Rose" "Libby Path"
  ..$ data_managers     : chr [1:3] "Lotta Outliers" "Audrey Trail" "Paige Missing"
>
> config$main$project
[1] "wonderdrug"
> config$team$programmers
[1] "Mack Rose" "Libby Path"
>
```

- Use the `{yaml}` package and the `read_yaml()` function
- It is just a `nested list` in R

The `{config}` Package

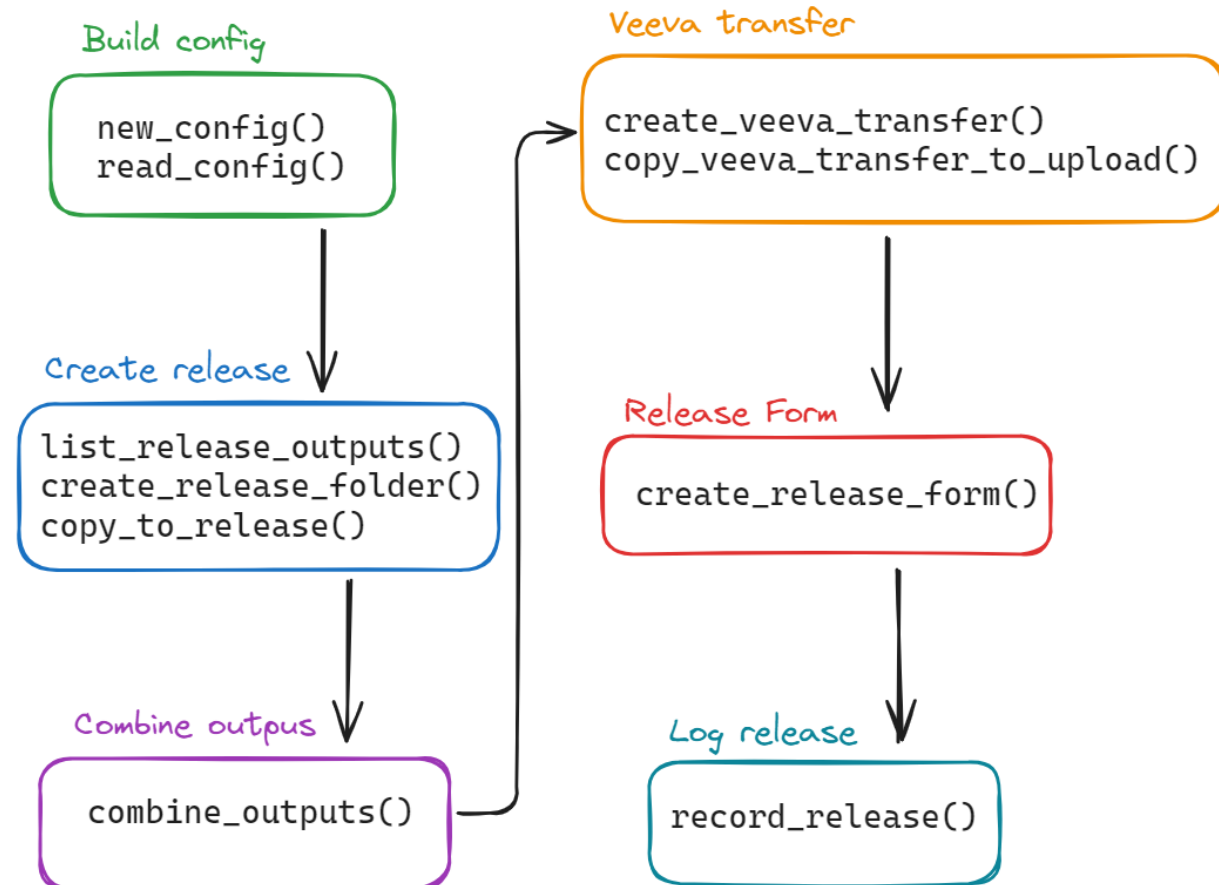


- ✓ Utilizes the `{yaml}` package
- ✓ Allows **development and production** configurations
- ✓ Integrates with **Posit Connect**

config.yml to Streamline Processes

```
1 ▾ default:
2 ▾   main:
3     project:
4     study:
5     delivery:
6     trial_statistician:
7     protocol_id:
8     protocol_version:
9     sap_veeva_id:
10 ▾  paths:
11    study_path:
12    setup_path:
13    tracker_path:
14    toc_path:
15    release_path:
16 ▾  release:
17    type:
18    reason:
19    name:
```

```
20 ▾ tracker:
21    batch_no:
22    priority_no:
23    filter:
24 ▾  veeva:
25    product_external_id:
26    clinical_studies_name:
27    transfer_area:
28    delivery_type:
29    pool_type:
30    administrative_groups:
```



YAML: Other Uses

- Programming environment set-up
- Shiny app config
- Quarto document headers
- CI/CD processes
- CDISC Analysis Results Standards



Improving User Experience with {cli}

What a user might want to know:

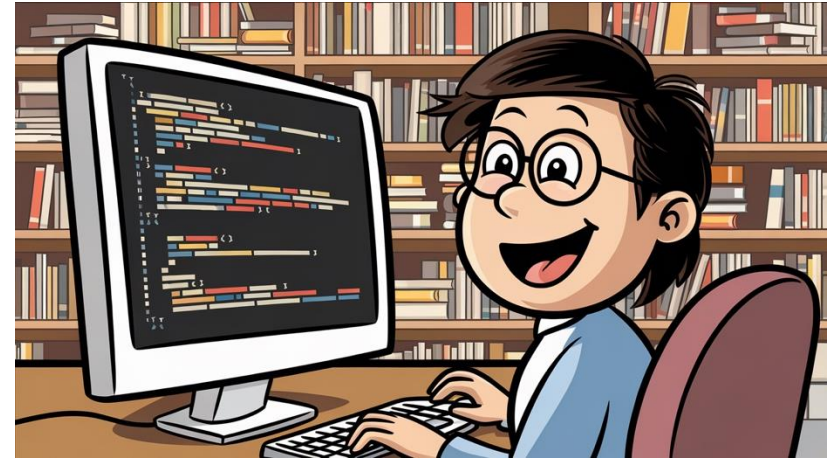
- **Environment** set-up variables
 - Data and metadata **pathnames**
 - Parsed **config** information
- **Files/directories** created/copied
- TLFs **created/modified**
- **Git** operations performed



Why Add Console Messaging?

Providing clear messaging:

- ✓ Adds **transparency**
- ✓ **Reassures** less experienced R users
- ✓ **Accelerates** debugging/corrective actions



The {cli} Package

```
> cli_alert_danger("This should not have happened!")  
✗ This should not have happened!
```

```
> cli_alert_warning("You might want to check this!")  
! You might want to check this!
```

```
> cli_alert_info("This is useful information!")  
i This is useful information!
```

```
> cli_alert_success("Wow, exactly as expected!")  
✓ Wow, exactly as expected!
```

The `{cli}` Package

Inline Markup



```
> n_outputs <- 32
> glue("Found {n_outputs} outputs to release.")
Found 32 outputs to release.
```

```
> cli_text("Something happened here: {.path path/to/event}")
Something happened here: path/to/event
```

```
> path <- file.path("path", "to", "event")
> cli_text("Something happened here: {.path {path}}")
Something happened here: path/to/event
```

Use Case: Enhancing Function Transparency

```
> outputs <- list_release_outputs(config = config)
i [redacted] was found.

— Loading and cleaning tracker —
i Reading tracker.
i Cleaning header and sub header.
i Augmenting tracker with TOC from [redacted].
i Including 4 in-text rows to the tracker.
✓ Tracker successfully loaded, cleaned and augmented with the TOC!

— Filtering tracker based on config —

— Applying filters —

batch number: >=2
priority number: ==9
filter expression: str_detect(output_title, 'DB-WD baseline')

— Outcome —

i Found 1 output.
i Found 1 program.

! Found 6 additional files to copy. These include sas7bdat, sas7bitm, sas, html, and log files.
Please inspect the results for more details.
```

Striking a balance!



Overly chatty functions are **overwhelming**!



Consider adding arguments to **suppress** or **reduce** messaging



Messages should be **helpful** and **concise**

Quick Wins with Quarto

Overview

- Successor to RMarkdown
- Seamlessly combine code, text, and images



★ Templates

★ Parameterised Reports

★ Automated Emails

Quarto Templates

- Templates incorporate **predefined styles** such as fonts, color schemes, and layouts
- Focus on **content** rather than styling

```
1 ---  
2 title: "An Example Report"  
3 format:  
4   docx:  
5     reference-doc: report_template.docx  
6 ---  
7  
8 # Overview  
9  
10 This report was made with a  
11 Microsoft Word template.  
12
```



Quarto Templates

- Templates incorporate **predefined styles** such as fonts, color schemes, and layouts
- Focus on **content** rather than styling

```
1 ---
2 title: "An Example Report"
3 format:
4   docx:
5     reference-doc: report_template.docx
6 ---
7
8 # Overview
9
10 This report was made with a
11 Microsoft Word template.
12
```

- Seamlessly switch between templates based on the desired output

- Supports    

Parameterised reports

- Enables a **programmatic** approach to **document generation**
- Computational **parameters** are defined in the **YAML** header

```
1 ---
2 title: "Patient Profile"
3 date: today
4 date-format: long
5 format: html
6 params:
7     usubjid: "01-701-1097"
8 ---
```

```

1 ---
2 title: "Patient Profile"
3 date: today
4 date-format: long
5 format: html
6 execute:
7   echo: false
8 params:
9   usubjid: "01-701-1097"
10 ---
11
12 ```{r}
13 #/ message: false
14 library(tidyverse)
15
16 load("/home/marqujo1/BCE/STD/RTRAINING/TEST_DATA/CDISC_PILOT/adsl.Rda")
17 load("/home/marqujo1/BCE/STD/RTRAINING/TEST_DATA/CDISC_PILOT/adv.s.Rda")
18 adsl_filtered <- filter(adsl, USUBJID == params$usubjid)
19 advs_filtered <- filter(advs, USUBJID == params$usubjid,
20                           PARAMN == 1, ATPTN == 816)
21 ```
22
23 # Demographics
24 **Patient:** `r params$usubjid`
25
26 **Age:** `r adsl_filtered$AGE`
27
28 **Treatment Arm:** `r adsl_filtered$ARM`
29
30 # Vital Signs
31 ```{r}
32 ggplot(data = advs_filtered, aes(x = ADY, y = AVAL)) +
33   geom_line(linetype = "dashed")+
34   geom_point() +
35   labs(title = "Systolic Blood Pressure (mmHg) after standing for 1 minute",
36        x = "Systolic Blood Pressure (mmHg)",
37        y = "Analysis Relative Day")
38 ```

```

Patient Profile

PUBLISHED

October 20, 2024

Demographics

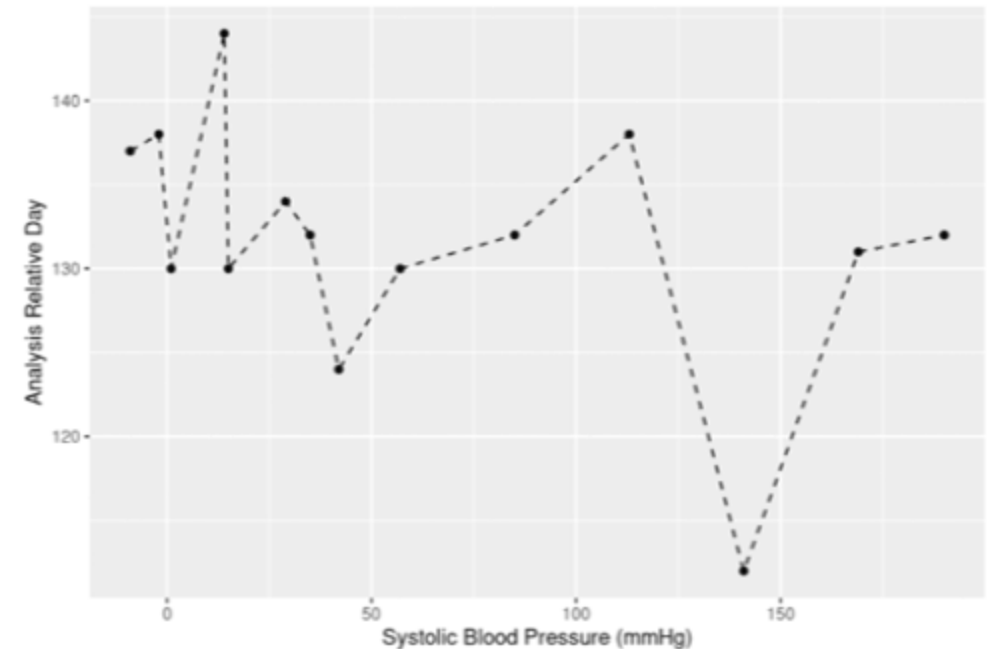
Patient: 01-701-1097

Age: 68

Treatment Arm: Xanomeline Low Dose

Vital Signs

Systolic Blood Pressure (mmHg) after standing for 1 minute



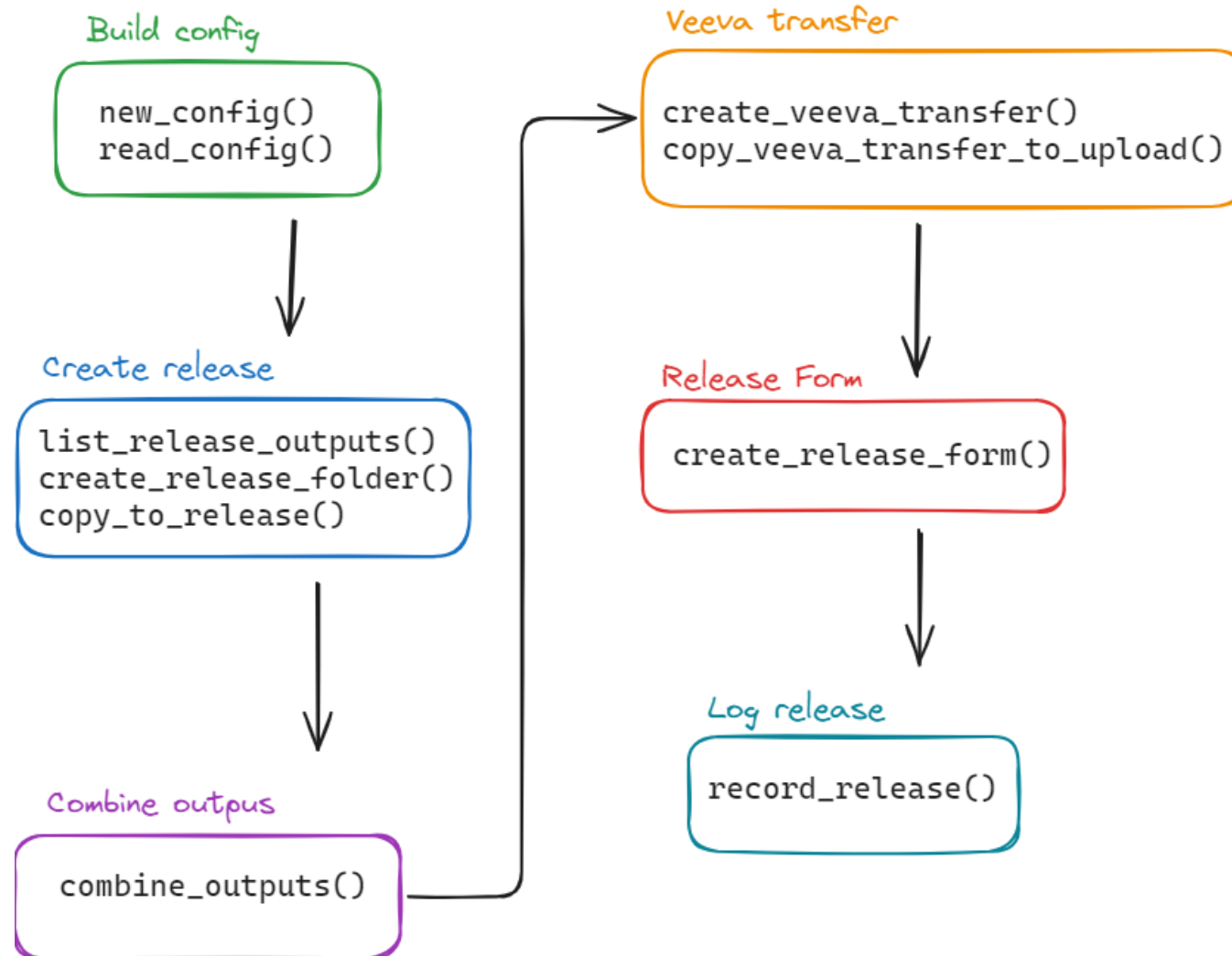
Parameterised reports

- Supports **dynamic inputs** via `quarto_render()`

```
1 criteria_usubjid <- find_usubj_matching_criteria()  
2  
3 quarto_render(  
4   input = "patient_profiles_source.qmd",  
5   output_file = "patient_profiles.html",  
6   execute_params =  
7     list(usubjid = criteria_usubjid)  
8 )  
9
```

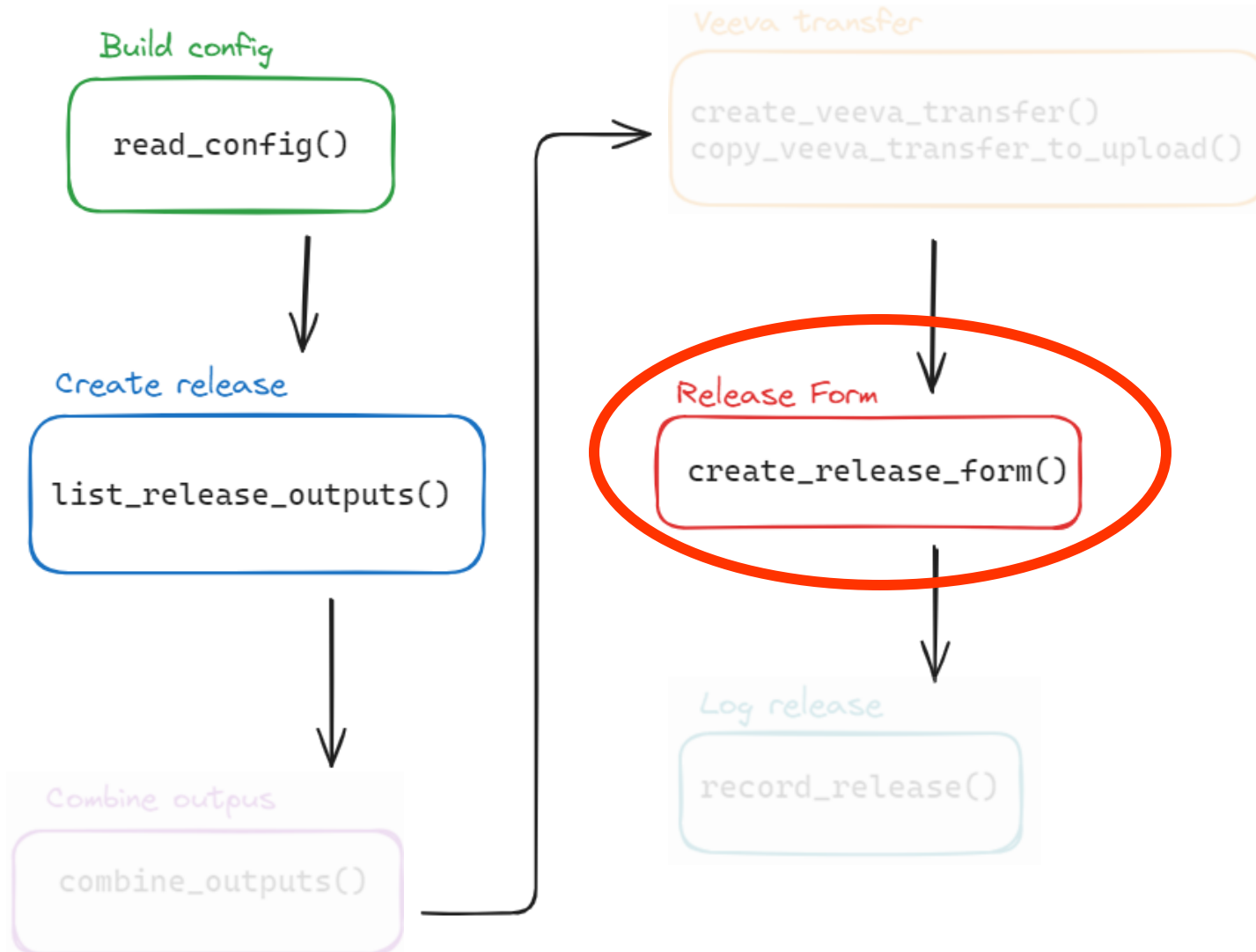
Putting it together

To create the release form, we combine templates and parameterized reports



Putting it together

To create the release form, we combine templates and parameterized reports



Putting it together

To create the release form, we combine templates and parameterized reports

Build config

`read_config()`

Reads the configuration details

Create release

`list_release_outputs()`

Returns the list of outputs and programmers

Release Form

`create_release_form()`

Takes the configuration details and the list of outputs.

Extracts the programmers and statisticians involved in the release

Renders Quarto document using a predefined template and parameters

Use Case: Generating a Release Form

Release Form

create_release_form()

```
107 quarto render(  
108   input = source_release_form,  
109   output_file = output_file_name,  
110   execute_params = list(  
111     "config" = config,  
112     "programmers" = progs,  
113     "statisticians" = stats,  
114     "signatures" = signers  
115   ),  
116   quiet = FALSE  
117 )
```

source_release_form.qmd

(i.e., the document with release form content)

```
1 ---  
2 title: ""  
3 format:  
4   docx:  
5     reference-doc: template_FRM101027.docx  
6 execute:  
7   echo: false  
8   params:  
9     config: ""  
10    programmers: ""  
11    statisticians: ""  
12    signatures: ""  
13 ---  
14  
15  
16 ```{r}  
17 #/ label: load-libs  
18  
19 library(dplyr)
```

Use Case: Generating a Release Form



Statistical Programming Release Form

FRM-101027

1. RELEASE PACKAGE KEY INFORMATION

Compound/Study	COMPOUND\STUDY\EXAMPLE01A
Release type/Purpose	Updated Tables - Descriptive stats tables needed updating due to errors (19 tables impacted). - Added new table for Q187 - Q164 tables of ABPM needed updating due to incorrect data being selected – updated to add more data points – p-values not impacted
Source data location(s)	X:\COMPOUND\STUDY\ADAM\RELEASE\ADAM X:\COMPOUND\STUDY\SDTM\CURRENT\SDTM
Output location	X:\COMPOUND\STUDY\CSR\CURRENT\RELEASE\EXAMPLE01A\TLF

2. PROGRAMMING SPECIFICATIONS AND DOCUMENTATION

Document Name	Version or Date	File Name and Location
Protocol ID-010A301	V1.1	Path or link to protocol
Statistical Analysis Plan - methods	Version 2.0	Stored in Veeva under ID: VV-CLIN-012345

3. QUALITY CONTROL

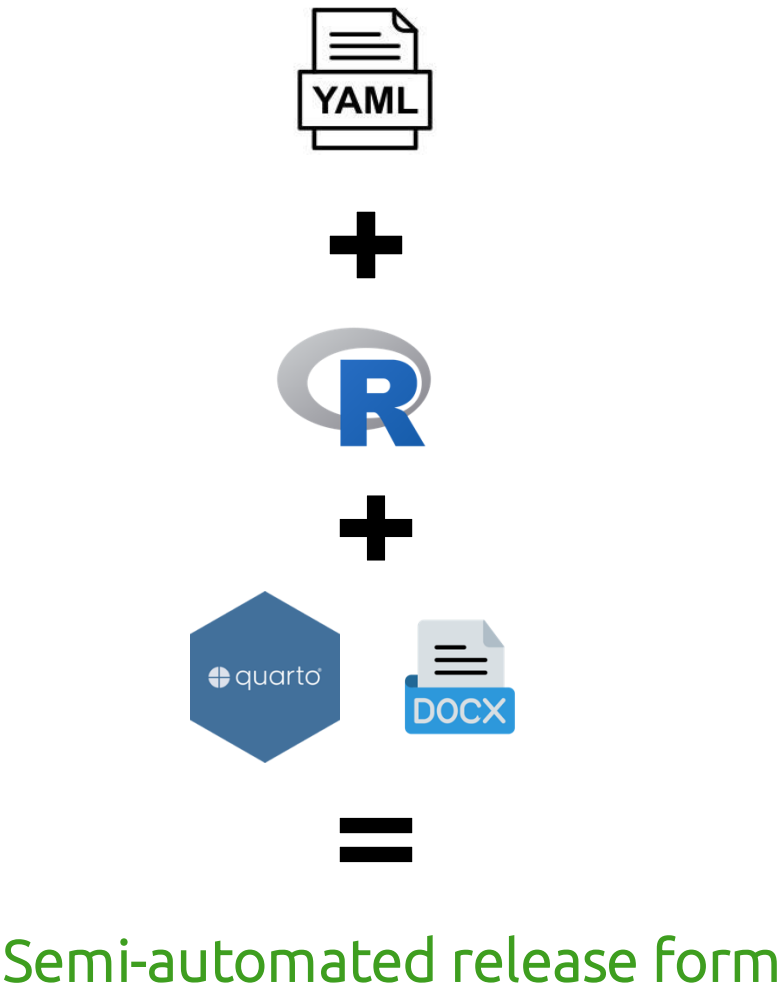
QC Components	Documentation and Comments
Risk levels definitions	X:\COMPOUND\STUDY\CSR\CURRENT\DOC\example_01A_tracker.xlsm Tracker details approach to outputs. Most QC'd by double programming using Double Programming

4. RELEASE PACKAGE SIGN-OFF

Author(s)			
Name	Role	Signature Date [dd-mmm-yyyy]	Signature
Sam A Smith	Main Programmer	{{SigB_es_;signer1:signatureblock}}	{{Dte_es_;signer1:date}}
Mary Jones	Main Programmer	{{SigB_es_;signer2:signatureblock}}	{{Dte_es_;signer2:date}}

FRM-101027_01

PAGE 1 OF 2



Automated Emails

[Posit Connect] Weekly Report

JM

Joana Marques Barros <noreply-rstudio@idorsia.com>

To

Retention Policy

Auto deletion after 6 months (6 months)

Expires

21/04/2025

Weekly_Report.xlsx

60 KB

The full Weekly Report, including the Cleaning Status, is [available here](#) and sent in attachment.

Executive Summary on 11/01/2024

Summary of Data entry	Comment
Missing visits for more than 21 days	40 visits/ 10,000 visits in total (completed + ready to complete)
Missing pages for more than 21 days	20 pages/ 40,000 pages in total (completed + missing)
Visits not SDVed	9.1 %/ 5,000 SDV in total (completed + pending)
Queries	10 open for more than 14 days. 200 outstanding queries

Location: STUDY/EDC

Folder name: STUDY_DATA_TRANSFER_2024_01_01

File name: Clean_Patient_Tracker.xlsx

File sheet: sheet 2

Access date: 10/10/2024

This message was generated on 2024-10-23 17:42:22.

This Version: https://rstudio.idorsia.com/content/2e0f676e-d85c-4779-9d77-52c7a1df133c/_rev3080/

Latest Version: <https://rstudio.idorsia.com/content/2e0f676e-d85c-4779-9d77-52c7a1df133c/>

If you wish to stop receiving emails for this document, you may [unsubscribe here](#).

28 PHUSE EU Connect 2024 - Open Source Technologies Stream - OS16 - 13 November 2024

Summary

Many opportunities for applying R beyond ADaM & TLF programming

Think about pain points in your own processes

Stay curious!



Thank you!

Any Questions?

Matthew Brierley
matthew.brierley@idorsia.com

Joana Marques Barros
joana.marques-barros@idorsia.com

