

Reproducible Clinical Data Review: A Modular R and Quarto Workflow for Transparent Reporting

Winkle Lu

Clinical Statistical Programming · Data Visualization

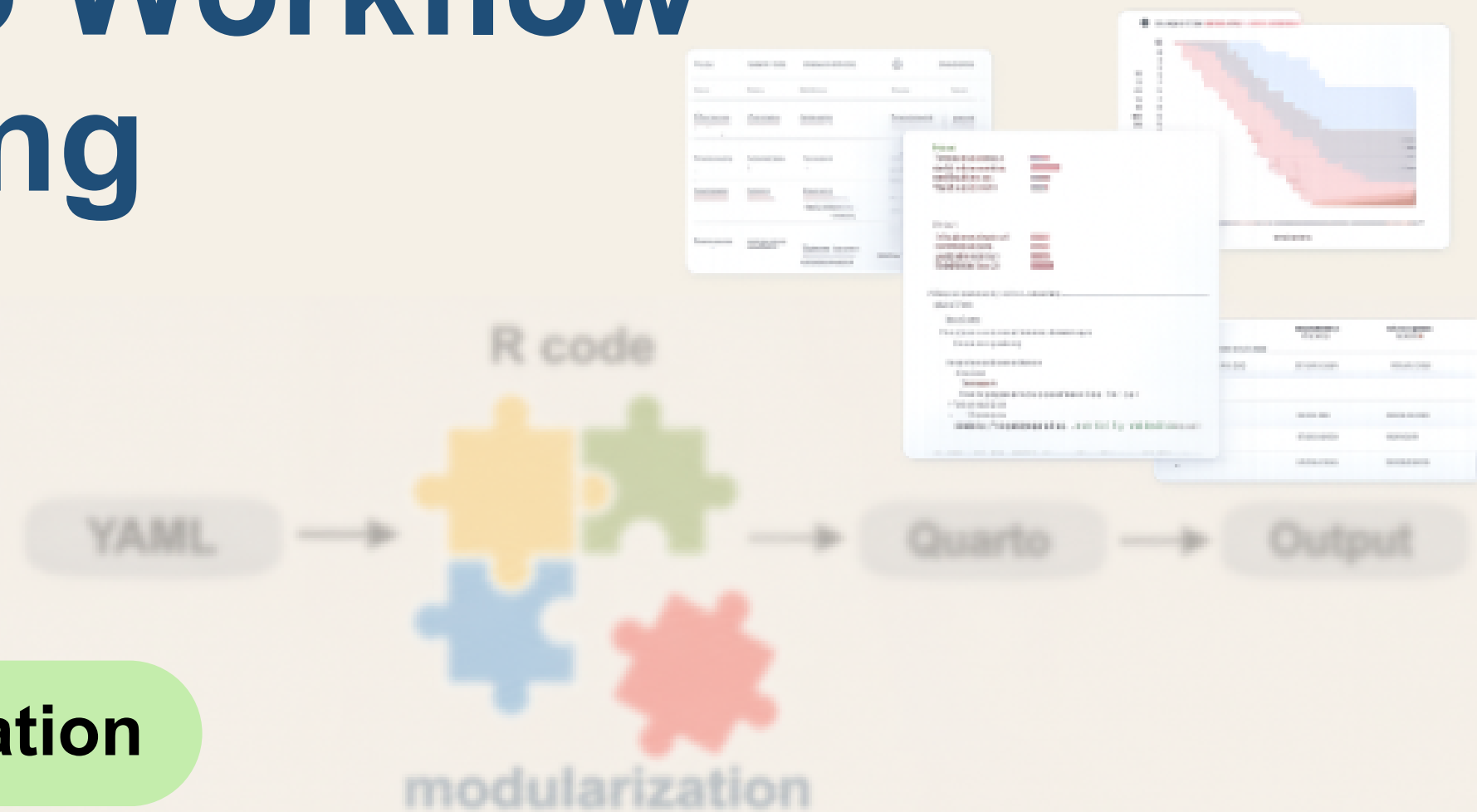
R + Quarto + YAML

Reproducible Clinical Data Review: A Modular R and Quarto Workflow for Transparent Reporting

Winkle Lu

Clinical Statistical Programming · Data Visualization

R + Quarto + YAML



**Reduced data
monitoring report
time by**

60%

with a modular workflow

**Every new cut of the data
is a multi-day round trip**

Extract the data

Map the variable

Write the code

Output

Send it back

Extract the data

Map the variable

Write the code

Output

Send it back

Extract the data

Map the variable

Write the code

Output

Send it back

Extract the data

Map the variable

Write the code

Output

Send it back

Extract the data

Map the variable

Write the code

Output

Send it back

The back-and-forth
takes days

***The back-and-forth
takes days***

***The output often
can't be traced back***

A modular workflow

Built in R + Quarto

```
1 # 1. Field Mapping
2 fields:
3     dm:
4         subject_col: PT
5         arm_col: ARM
6         site_col: SITE_ID
7         trt_start_col: TRTSDTC
8         pop_flag_col: SAFFL
9         age_col: AGE_FUL
10        sex_col: SEX_FUL
11        race_col: RACE_FUL
12        ethnic_col: ETHNIC_FUL
13        ecog_col: ECOG
14
15    ae:
16        subject_col: PT
17        term_col: AETERM_FUL
18        pt_col: AEPT_FUL
19        soc_col: AESOC_FUL
20        grade_col: AETOXGR_FUL
21        rel_col: AEREL_FUL
22        serious_col: AESER_FUL
23        ongoing_col: AEONGO_FUL
24        outcome_col: AEOUT_FUL
25        action_col: AEACN_FUL
26        start_col: AESTDAT_FUL
27        end_col: AEENDAT_FUL
28
```

```
# 2. Filters
```

```
filters:
```

```
  global:
```

```
    population: "SAFFL == 'Y'"
```

```
    sites: all
```

```
    subjects: all
```

```
  subject_criteria:
```

```
    combine: "OR"
```

```
    criteria:
```

```
      - label: "Has AE Grade >= 3"
```

```
        domain: ae
```

```
        condition: "as.numeric(GRADE_NUM) >= 3"
```

```
      - label: "Has abnormal ALT or AST"
```

```
        domain: lb
```

```
        condition: "LBTEST_FUL %in% c('ALT', 'AST') & LBNRIND_FUL == 'HIGH'"
```

```
98 # 4. Labels
99 listings:
100 - domain: ae
101   sort_by: [PT, AESTDAT_FUL]
102   columns:
103     - {col: PT,          label: "Subject ID"}
104     - {col: AETERM_FUL,  label: "AE Term"}
105     - {col: AEPT_FUL,    label: "Preferred Term"}
106     - {col: AESOC_FUL,   label: "SOC"}
107     - {col: AETOXGR_FUL, label: "Grade"}
108     - {col: AEREL_FUL,   label: "Relationship"}
109     - {col: AESER_FUL,   label: "Serious"}
110     - {col: AEACN_FUL,   label: "Action Taken"}
111     - {col: AEOUT_FUL,   label: "Outcome"}
112     - {col: AESTDAT_FUL, label: "Start Date"}
113     - {col: AEENDAT_FUL, label: "End Date"}
114 - domain: lb
115   flag_col: LBNRIND_FUL
116   sort_by: [PT, LBDAT_FUL, LBTEST_FUL]
117   columns:
118     - {col: PT,          label: "Subject ID"}
119     - {col: CPEVENT,     label: "Visit"}
120     - {col: LBCAT_FUL,   label: "Category"}
121     - {col: LBTEST_FUL,  label: "Lab Test"}
122     - {col: LBORRES_FUL, label: "Result"}
123     - {col: LBUNIT_FUL,  label: "Unit"}
124     - {col: LBNRIND_FUL, label: "Flag"}
125     - {col: LBCLSIG_FUL, label: "CS"}
126     - {col: LBDAT_FUL,   label: "Date"}
```

A self-contained report

A self-contained report

- Filter conditions
- Tables
- Figures
- Listings

A new domain or chart ?

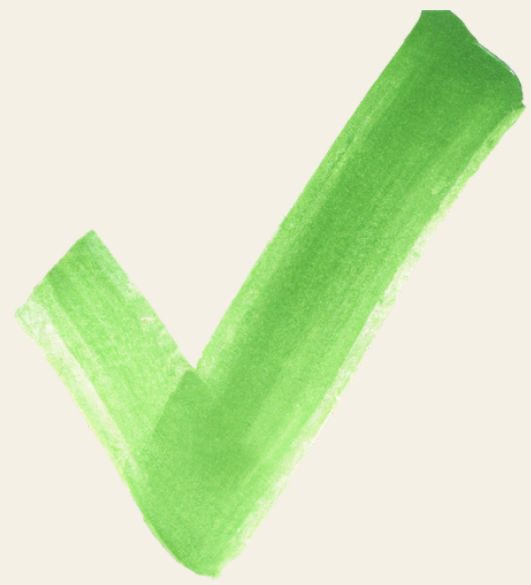
A new domain or chart ?

No changes to the main code

Just adding a module

time
saved

60%+



Traceability



What did we
see at last check?

What conditions
were set?

NO
reconstruction

Clinical Data Review Report

PUBLISHED
June 25, 2026

Review Conditions

Table of contents

[Review Conditions](#)

- 1. Summary Statistics
- 2. Patient Timeline
- 3. CRF Listings

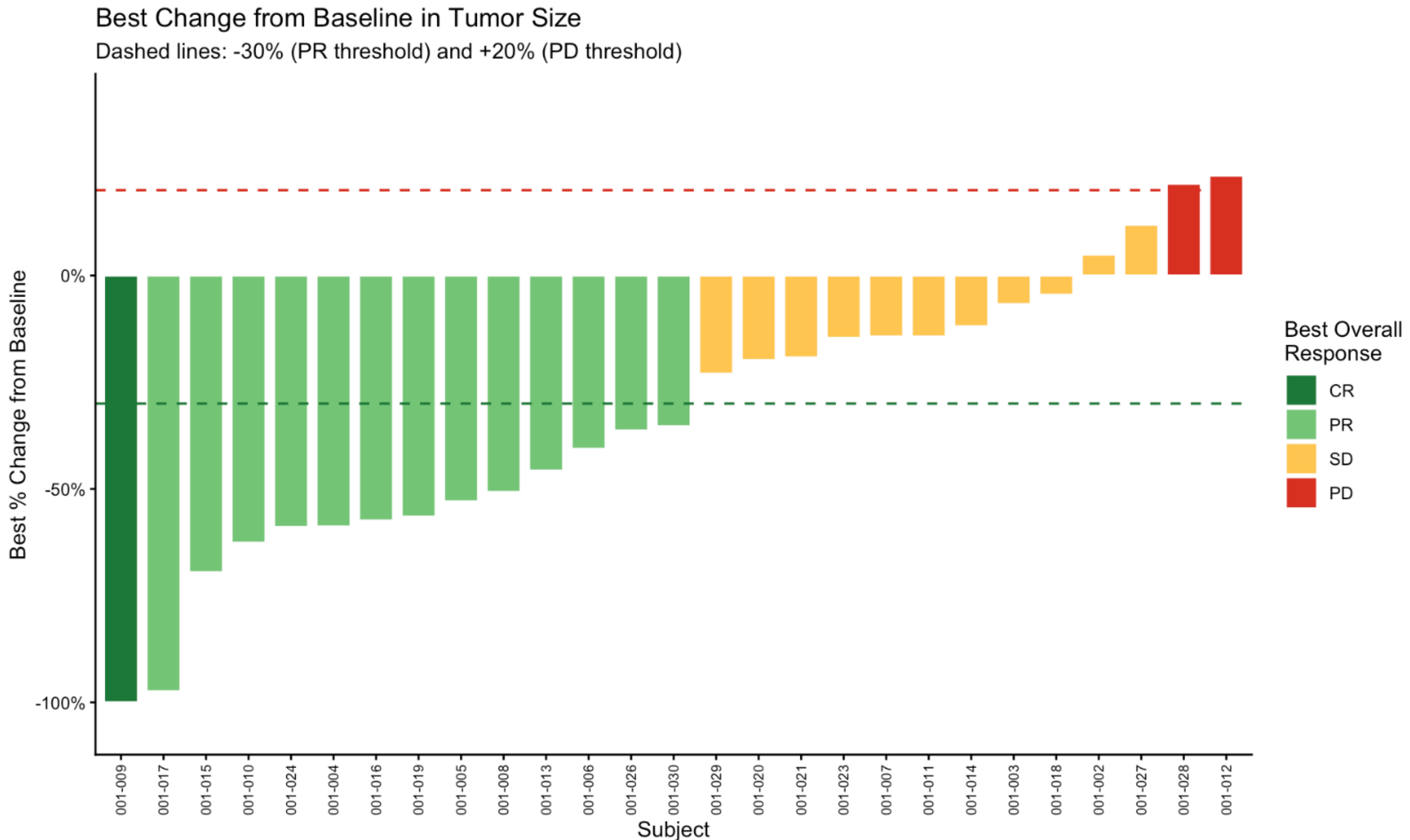
Active Review Conditions

Base Population: `SAFFL == 'Y'` — 30 subjects **Subject Criteria (OR):** - Has AE Grade $\geq 3 \rightarrow$ 16 subjects - Has abnormal ALT or AST $\rightarrow$ 27 subjects **Final Subject List:** 27 subjects *Generated: 2026-06-25 19:22:56 | Config: /Users/winkle/_Conference/A Modular R and Quarto Workflow for Transparent Reporting/config/config.yaml*

1.1 Demographic Characteristics

Characteristic	Treatment A N = 14 ¹	Treatment B N = 13 ¹	Overall N = 27 ¹
Age, years — Median (Min, Max)	60 (48, 69)	62 (53, 73)	61 (48, 73)
Sex			
F	2 (14%)	4 (31%)	6 (22%)
M	12 (86%)	9 (69%)	21 (78%)
Race			
ASIAN	4 (29%)	7 (54%)	11 (41%)
BLACK OR AFRICAN AMERICAN	2 (14%)	1 (7.7%)	3 (11%)
OTHER	1 (7.1%)	1 (7.7%)	2 (7.4%)
WHITE	7 (50%)	4 (31%)	11 (41%)
ECOG PS			
0	5 (36%)	5 (38%)	10 (37%)
1	7 (50%)	6 (46%)	13 (48%)

1.3 Best Change from Baseline in Tumor Size



3. CRF Listings

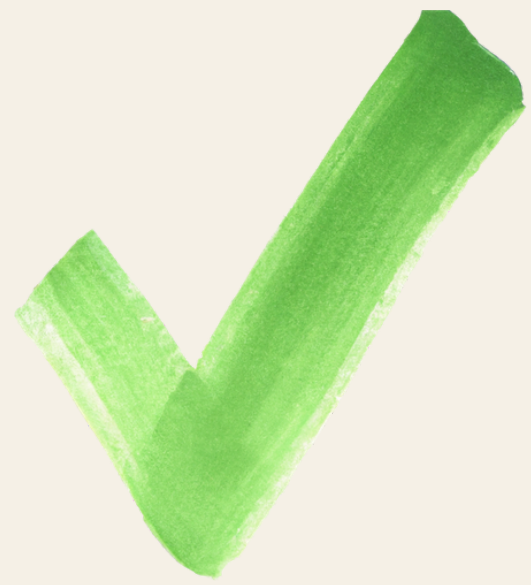
3.1 Adverse Event Listing

Copy

CSV

Search:

Subject ID	Treatment	Sex	AE Term	Preferred Term	SOC	Grade	Relationship
AI	All		All	All	All	All	All
001-002	Treatment A	M	Nausea	Nausea	Gastrointestinal disorders	Grade 1 - Mild	PROBABLY RELATED
001-002	Treatment A	M	Arthralgia	Arthralgia	Musculoskeletal and connective tissue disorders	Grade 1 - Mild	PROBABLY RELATED
001-003	Treatment A	M	Fatigue	Fatigue	General disorders and administration site conditions	Grade 2 - Moderate	PROBABLY RELATED
001-004	Treatment A	M	Arthralgia	Arthralgia	Musculoskeletal and connective tissue disorders	Grade 1 - Mild	PROBABLY RELATED

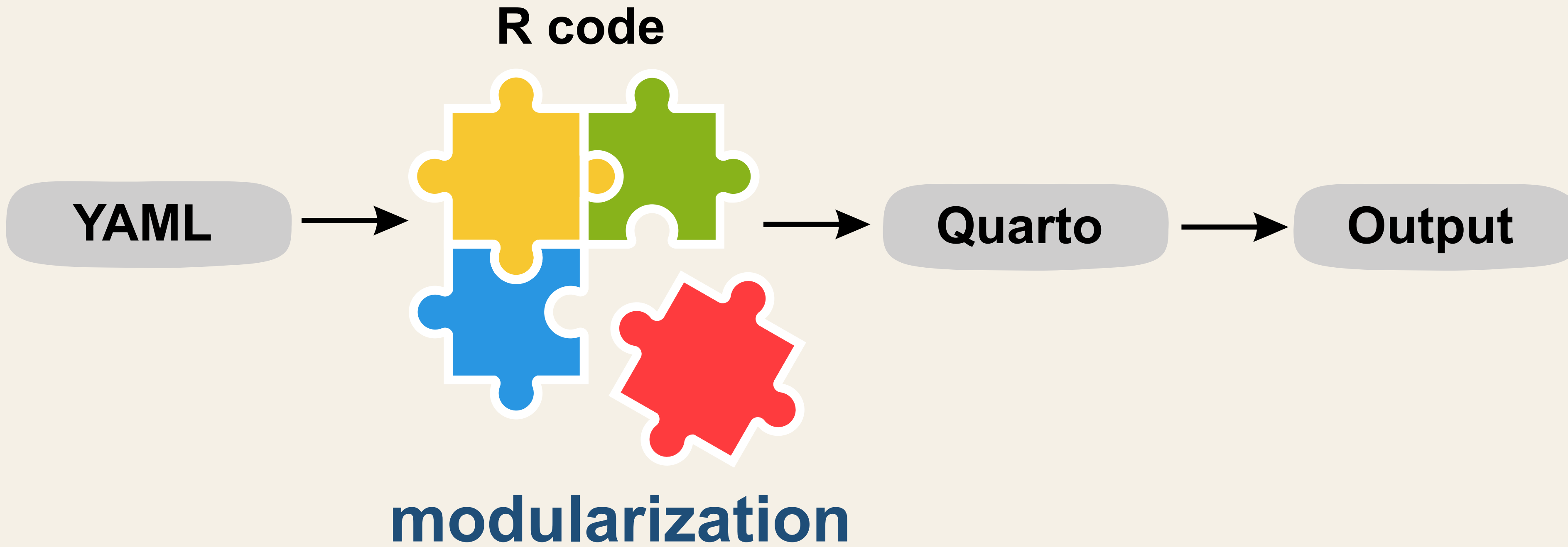


Reusable

Minimal code changes

Mostly YAML configuration

To Sum Up



Full article by Winkle

