

# An open source NEST R package Chevron for streamlining generation of standard TLFs

In both static and interactive formats

Lena Wang, Analytical Data Scientist



# Table of contents

1. Executive Summary
2. Background and Introduction
3. Automation and Workflow
4. Business Impact

# Executive Summary

# Executive Summary

- At Roche we are currently transitioning to an R backbone and open source movement to enable our clinical trial reporting
  - Have regulatory/exploratory analyses in the same language
  - R is a more modern, functional programming language (Shiny - interactive apps & R markdown - reproducible docs)
  - IT infrastructure: R allows us to run containers in different environments/computing platforms
- In addition to a new computing platform and a series of R packages and tools (e.g. NEST), we have also been investing in products to enable automation of our clinical trial reporting workflow
- Chevron is a collection of R packages under the NEST umbrella that contains standard TLF templates and tools to pipeline the creation of outputs with automation
- We are rolling Chevron out to data scientists at Roche this year and hope it will help improve efficiency, reproducibility and user experience

# Background and Introduction

# Chevron is focusing on clinical trial regulatory reportings



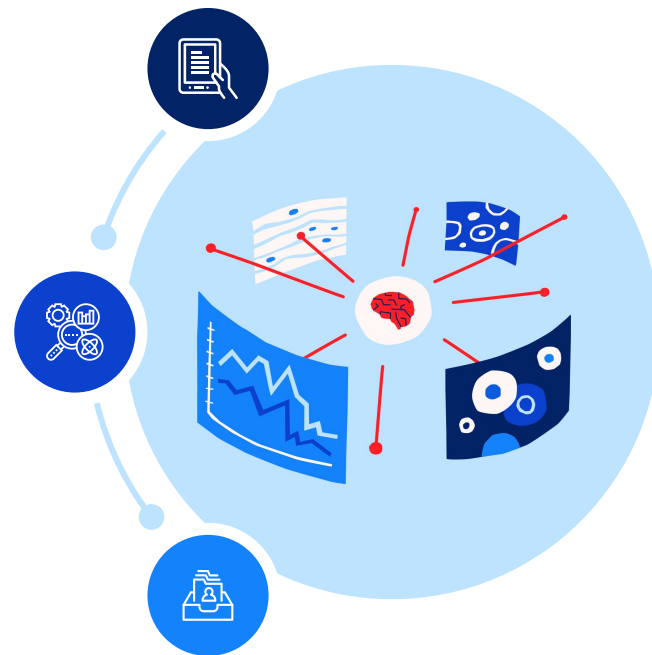
**Main task:  
reporting  
clinical trials**

**Summarize safety and  
efficacy data**

**Provide an accurate  
picture of trial outcomes**

**Follow regulatory medical  
writing guidelines**

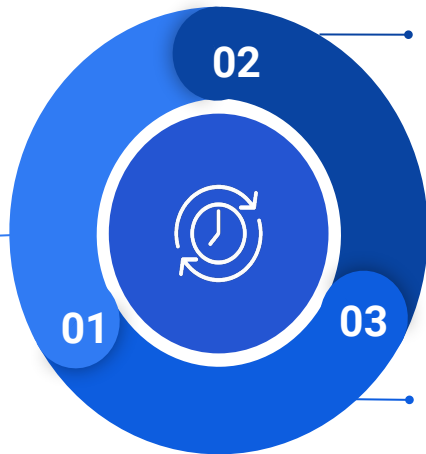
**Enable automation using  
modern tools**



# Challenges with our legacy approach

Transparency and complexity to **debugging** with licensed programming languages.

It's much easier to debug with open-source programming language.



Downstream **automation** challenges.

Source code is:

- Not available to the public
- Not free to use
- Hard to customize

Traditional workflow is **not the most efficient** way to do.

- Challenges with legacy platform and lack of standard solutions (leading to teams developing different workarounds)
- Restrictive platform limited ability to develop efficient solutions

# Good timing for making a change



**Talent out  
of university  
more likely  
to know R**



**Modern tools  
offered  
opportunities:**

**01**

New approach to validate R packages - Follow R validation hub guidance([PharmaR.org](https://pharmar.org)) which is implemented in Roche

**02**

More open source R/python workflow management systems like snakemake

**03**

Shift toward language-agnostic programming & collaboration([Pharmaverse.org](https://pharmaverse.org)) in the pharmaceutical industry

# Our Goal of Chevron

Reproducibility

Generate  
**reproducible and  
scalable** data  
analyses.

Standardization

Provide an R-based  
**workflow of  
standardization.**  
Consistent across  
**over 300 Roche  
molecule projects**  
each year.

Efficiency

**Improve the  
day-to-day work  
efficiency** of data  
science users.

Automation

Enable **automation**  
and **reduce manual  
steps** by using a  
Python based tool.

# **Automation and Workflow**

# NEST Package Relationships

Tables & Listings



Figures



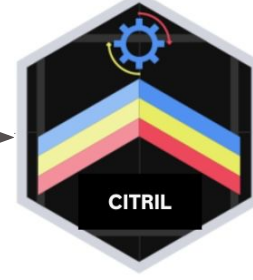
TLF (Table Listing & Figures) Toolkit



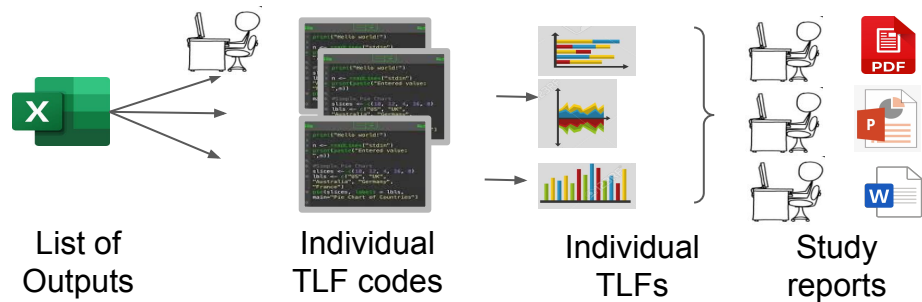
Table Templates



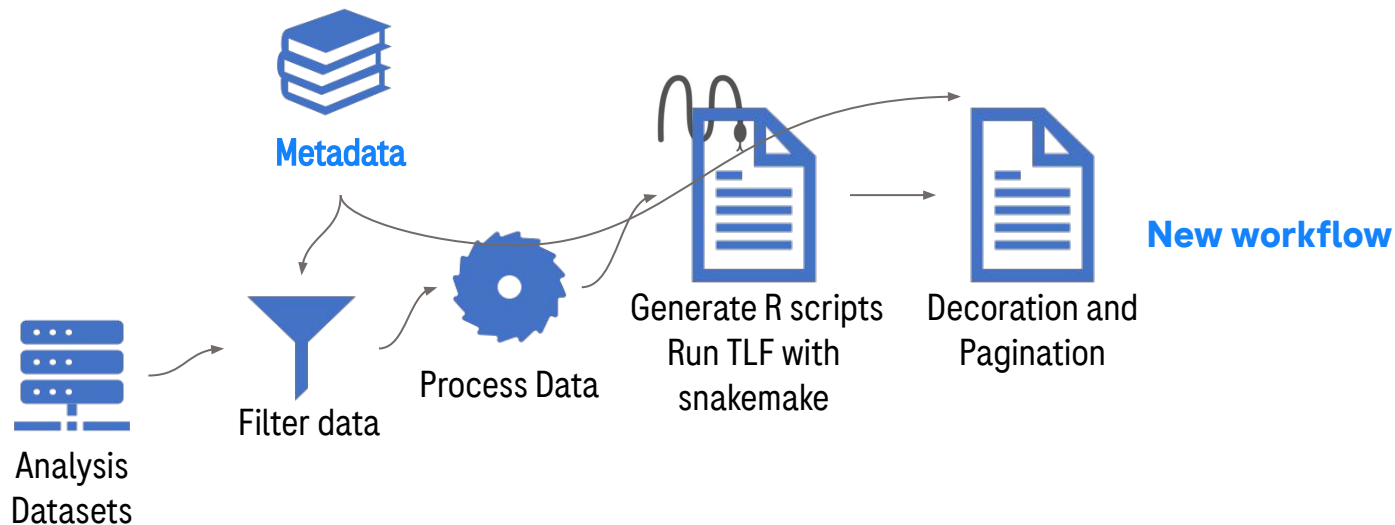
Workflow Toolkit



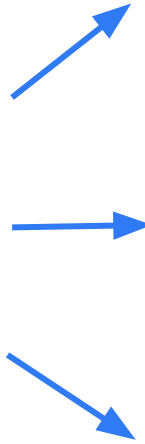
# New metadata-driven TLF workflow



Old workflow

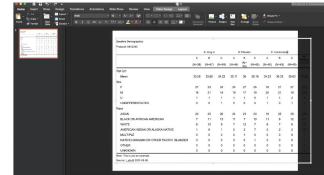


# Flexible Formats for different purposes



| r.fische.com / 33570767/3466619f9a...    |           |       |       |       |       |           |       |       |       |       |               |       |       |       |  |
|------------------------------------------|-----------|-------|-------|-------|-------|-----------|-------|-------|-------|-------|---------------|-------|-------|-------|--|
| Business Development<br>Project - 401200 |           |       |       |       |       |           |       |       |       |       |               |       |       |       |  |
|                                          | A: Single |       |       |       |       | B: Pooled |       |       |       |       | C: Comparison |       |       |       |  |
|                                          | A         | B     | C     | A     | B     | C         | A     | B     | C     | A     | B             | C     | A     | B     |  |
| Age (Y)                                  |           |       |       |       |       |           |       |       |       |       |               |       |       |       |  |
| Sex                                      | 10.08     | 10.08 | 10.11 | 10.11 | 10.11 | 10.11     | 10.11 | 10.11 | 10.11 | 10.11 | 10.11         | 10.11 | 10.11 | 10.11 |  |
| Age                                      | 21        | 25    | 28    | 28    | 27    | 28        | 18    | 21    | 27    | 22    | 21            | 27    | 22    | 21    |  |
| W                                        | 10        | 10    | 10    | 10    | 10    | 10        | 10    | 10    | 10    | 10    | 10            | 10    | 10    | 10    |  |
| U                                        | 1         | 1     | 1     | 1     | 1     | 1         | 1     | 1     | 1     | 1     | 1             | 1     | 1     | 1     |  |
| UNIDENTIFIED                             | 0         | 0     | 0     | 0     | 0     | 0         | 0     | 0     | 0     | 0     | 0             | 0     | 0     | 0     |  |
| Race                                     |           |       |       |       |       |           |       |       |       |       |               |       |       |       |  |
| WHITE                                    | 10        | 10    | 10    | 10    | 10    | 10        | 10    | 10    | 10    | 10    | 10            | 10    | 10    | 10    |  |
| BLACK OR AFRICAN AMERICAN                | 7         | 11    | 13    | 11    | 7     | 13        | 11    | 9     | 12    | 10    | 12            | 10    | 12    | 10    |  |
| ASIAN                                    | 8         | 10    | 9     | 7     | 10    | 7         | 9     | 7     | 8     | 7     | 8             | 7     | 8     | 7     |  |
| AMERICAN INDIAN OR ALASKA NATIVE         | 1         | 0     | 1     | 2     | 2     | 7         | 2     | 2     | 2     | 2     | 2             | 2     | 2     | 2     |  |
| MULTIPLE                                 | 0         | 0     | 0     | 0     | 1     | 0         | 0     | 0     | 0     | 0     | 0             | 0     | 0     | 0     |  |
| AMERICAN INDIAN OR ALASKA NATIVE         | 0         | 0     | 0     | 0     | 0     | 1         | 0     | 0     | 0     | 0     | 0             | 0     | 0     | 0     |  |
| UNIDENTIFIED                             | 0         | 0     | 0     | 0     | 0     | 0         | 0     | 0     | 0     | 0     | 0             | 0     | 0     | 0     |  |
| UNIDENTIFIED                             | 0         | 0     | 0     | 0     | 0     | 0         | 0     | 0     | 0     | 0     | 0             | 0     | 0     | 0     |  |

HTML/Shiny app



|                                  | A: Single |       |       |       |       | B: Pooled |       |       |       |       | C: Comparison |       |       |       |  | Total |
|----------------------------------|-----------|-------|-------|-------|-------|-----------|-------|-------|-------|-------|---------------|-------|-------|-------|--|-------|
|                                  | A         | B     | C     | A     | B     | C         | A     | B     | C     | A     | B             | C     |       |       |  |       |
| Age (Y)                          |           |       |       |       |       |           |       |       |       |       |               |       |       |       |  |       |
| Sex                              | 10.08     | 10.08 | 10.11 | 10.11 | 10.11 | 10.11     | 10.11 | 10.11 | 10.11 | 10.11 | 10.11         | 10.11 | 10.11 | 10.11 |  |       |
| Age                              | 21        | 25    | 28    | 28    | 27    | 28        | 18    | 21    | 27    | 22    | 21            | 27    | 22    | 21    |  |       |
| W                                | 10        | 10    | 10    | 10    | 10    | 10        | 10    | 10    | 10    | 10    | 10            | 10    | 10    | 10    |  |       |
| U                                | 1         | 1     | 1     | 1     | 1     | 1         | 1     | 1     | 1     | 1     | 1             | 1     | 1     | 1     |  |       |
| UNIDENTIFIED                     | 0         | 0     | 0     | 0     | 0     | 0         | 0     | 0     | 0     | 0     | 0             | 0     | 0     | 0     |  |       |
| Race                             |           |       |       |       |       |           |       |       |       |       |               |       |       |       |  |       |
| WHITE                            | 10        | 10    | 10    | 10    | 10    | 10        | 10    | 10    | 10    | 10    | 10            | 10    | 10    | 10    |  |       |
| BLACK OR AFRICAN AMERICAN        | 7         | 11    | 13    | 11    | 7     | 13        | 11    | 9     | 12    | 10    | 12            | 10    | 12    | 10    |  |       |
| ASIAN                            | 8         | 10    | 9     | 7     | 10    | 7         | 9     | 7     | 8     | 7     | 8             | 7     | 8     | 7     |  |       |
| AMERICAN INDIAN OR ALASKA NATIVE | 1         | 0     | 1     | 2     | 2     | 7         | 2     | 2     | 2     | 2     | 2             | 2     | 2     | 2     |  |       |
| MULTIPLE                         | 0         | 0     | 0     | 0     | 1     | 0         | 0     | 0     | 0     | 0     | 0             | 0     | 0     | 0     |  |       |
| AMERICAN INDIAN OR ALASKA NATIVE | 0         | 0     | 0     | 0     | 0     | 1         | 0     | 0     | 0     | 0     | 0             | 0     | 0     | 0     |  |       |
| UNIDENTIFIED                     | 0         | 0     | 0     | 0     | 0     | 0         | 0     | 0     | 0     | 0     | 0             | 0     | 0     | 0     |  |       |
| UNIDENTIFIED                     | 0         | 0     | 0     | 0     | 0     | 0         | 0     | 0     | 0     | 0     | 0             | 0     | 0     | 0     |  |       |

PowerPoint Presentation

| Table: Comparison                |           |       |       |       |       |           |       |       |       |       |               |       |       |       |  |       |
|----------------------------------|-----------|-------|-------|-------|-------|-----------|-------|-------|-------|-------|---------------|-------|-------|-------|--|-------|
|                                  | A: Single |       |       |       |       | B: Pooled |       |       |       |       | C: Comparison |       |       |       |  | Total |
|                                  | A         | B     | C     | A     | B     | C         | A     | B     | C     | A     | B             | C     |       |       |  |       |
| Age (Y)                          |           |       |       |       |       |           |       |       |       |       |               |       |       |       |  |       |
| Sex                              | 10.08     | 10.08 | 10.11 | 10.11 | 10.11 | 10.11     | 10.11 | 10.11 | 10.11 | 10.11 | 10.11         | 10.11 | 10.11 | 10.11 |  |       |
| Age                              | 21        | 25    | 28    | 28    | 27    | 28        | 18    | 21    | 27    | 22    | 21            | 27    | 22    | 21    |  |       |
| W                                | 10        | 10    | 10    | 10    | 10    | 10        | 10    | 10    | 10    | 10    | 10            | 10    | 10    | 10    |  |       |
| U                                | 1         | 1     | 1     | 1     | 1     | 1         | 1     | 1     | 1     | 1     | 1             | 1     | 1     | 1     |  |       |
| UNIDENTIFIED                     | 0         | 0     | 0     | 0     | 0     | 0         | 0     | 0     | 0     | 0     | 0             | 0     | 0     | 0     |  |       |
| Race                             |           |       |       |       |       |           |       |       |       |       |               |       |       |       |  |       |
| WHITE                            | 10        | 10    | 10    | 10    | 10    | 10        | 10    | 10    | 10    | 10    | 10            | 10    | 10    | 10    |  |       |
| BLACK OR AFRICAN AMERICAN        | 7         | 11    | 13    | 11    | 7     | 13        | 11    | 9     | 12    | 10    | 12            | 10    | 12    | 10    |  |       |
| ASIAN                            | 8         | 10    | 9     | 7     | 10    | 7         | 9     | 7     | 8     | 7     | 8             | 7     | 8     | 7     |  |       |
| AMERICAN INDIAN OR ALASKA NATIVE | 1         | 0     | 1     | 2     | 2     | 7         | 2     | 2     | 2     | 2     | 2             | 2     | 2     | 2     |  |       |
| MULTIPLE                         | 0         | 0     | 0     | 0     | 1     | 0         | 0     | 0     | 0     | 0     | 0             | 0     | 0     | 0     |  |       |
| AMERICAN INDIAN OR ALASKA NATIVE | 0         | 0     | 0     | 0     | 0     | 1         | 0     | 0     | 0     | 0     | 0             | 0     | 0     | 0     |  |       |
| UNIDENTIFIED                     | 0         | 0     | 0     | 0     | 0     | 0         | 0     | 0     | 0     | 0     | 0             | 0     | 0     | 0     |  |       |
| UNIDENTIFIED                     | 0         | 0     | 0     | 0     | 0     | 0         | 0     | 0     | 0     | 0     | 0             | 0     | 0     | 0     |  |       |

Word Document

# **Business Impact of developing and using Chevron**

# Business Impact of developing and using Chevron

- Chevron is enabling Roche data scientists to:
  - **Automate** the generation of standard TLFs
  - Increases **consistency** across studies
  - Improves **efficiency** and allow data scientists to spend their time on other important tasks
  - Enables **reproducibility** if/when studies need to be re-run
  - Increased **user satisfaction**
  
- In 2023 we will be rolling out this product and ultimately it will be scaled and adopted across the portfolio (~300 studies per year, across all disease areas and phases)

*“Automation is cost cutting by tightening the corners and not cutting them.”*

- Haresh Sippy

**Doing now what patients need next**

# What is Snakemake?



Snakemake is a workflow orchestration tool



Operates using “Snakefile” where workflow steps and required data sets are described in a readable way



Rules in Snakefile can feed output as input to other rules

# Package API



- rtables:
  - Expressive, modern API for table generation in R
  - Open source and available from GitHub
  - [rtables package repo](#)
- Chevron:
  - This package contains a collection of standard TLF templates.
  - [CHEVRON package repo](#)
- Citril
  - This package contains a collection of tools to pipeline the creation of chevron outputs
  - [CITRIL package repo](#)

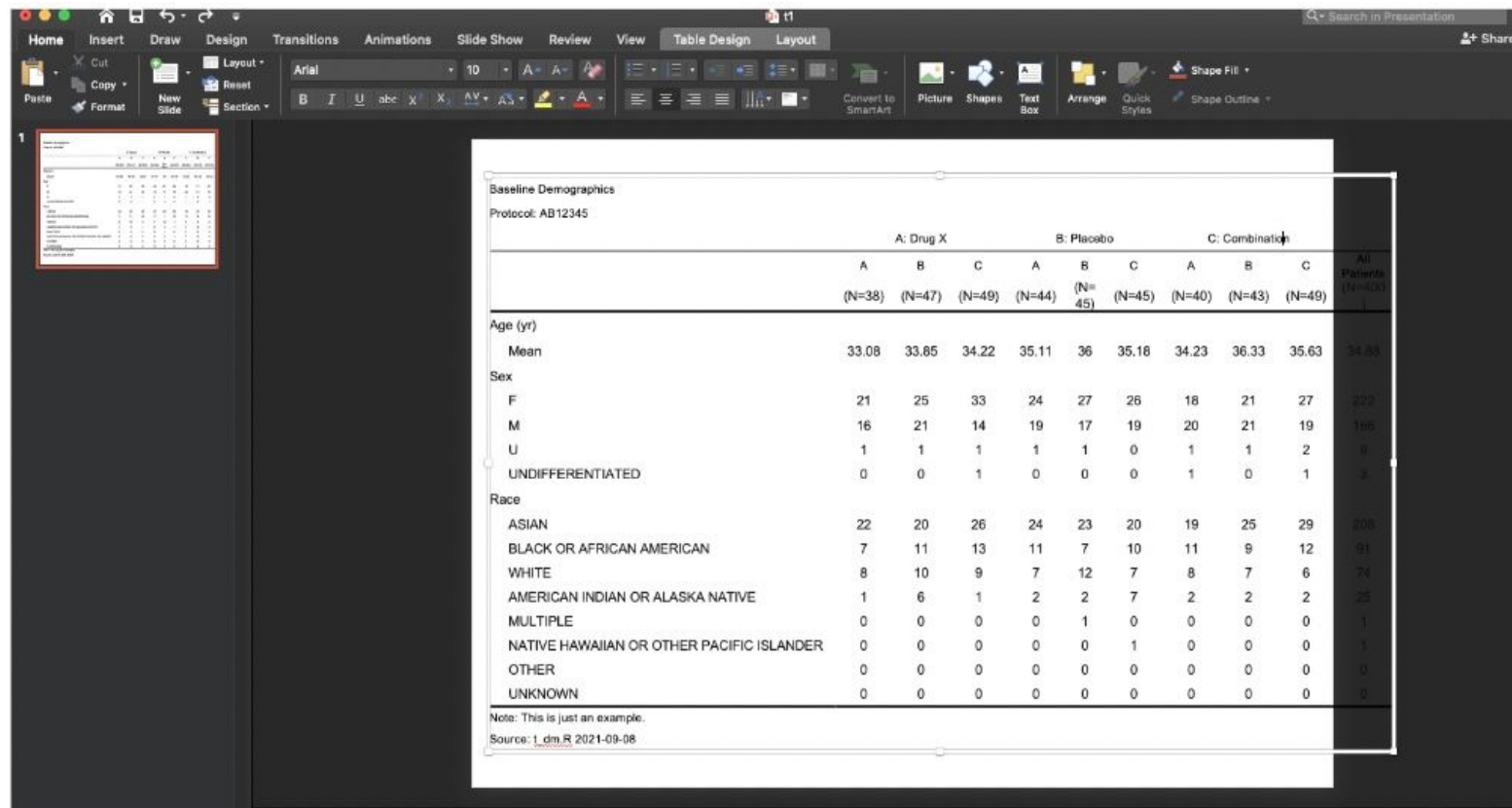
# HTML

Baseline Demographics  
Protocol: AB12345

|                                           | A: Drug X |        |        | B: Placebo |        |        | C: Combination |        |        | All Patients |
|-------------------------------------------|-----------|--------|--------|------------|--------|--------|----------------|--------|--------|--------------|
|                                           | A         | B      | C      | A          | B      | C      | A              | B      | C      |              |
|                                           | (N=38)    | (N=47) | (N=49) | (N=44)     | (N=45) | (N=45) | (N=40)         | (N=43) | (N=49) | (N=400)      |
| Age (yr)                                  |           |        |        |            |        |        |                |        |        |              |
| Mean                                      | 33.08     | 33.85  | 34.22  | 35.11      | 36     | 35.18  | 34.23          | 36.33  | 35.63  | 34.88        |
| Sex                                       |           |        |        |            |        |        |                |        |        |              |
| F                                         | 21        | 25     | 33     | 24         | 27     | 26     | 18             | 21     | 27     | 222          |
| M                                         | 16        | 21     | 14     | 19         | 17     | 19     | 20             | 21     | 19     | 166          |
| U                                         | 1         | 1      | 1      | 1          | 1      | 0      | 1              | 1      | 2      | 9            |
| UNDIFFERENTIATED                          | 0         | 0      | 1      | 0          | 0      | 0      | 1              | 0      | 1      | 3            |
| Race                                      |           |        |        |            |        |        |                |        |        |              |
| ASIAN                                     | 22        | 20     | 26     | 24         | 23     | 20     | 19             | 25     | 29     | 208          |
| BLACK OR AFRICAN AMERICAN                 | 7         | 11     | 13     | 11         | 7      | 10     | 11             | 9      | 12     | 91           |
| WHITE                                     | 8         | 10     | 9      | 7          | 12     | 7      | 8              | 7      | 6      | 74           |
| AMERICAN INDIAN OR ALASKA NATIVE          | 1         | 6      | 1      | 2          | 2      | 7      | 2              | 2      | 2      | 25           |
| MULTIPLE                                  | 0         | 0      | 0      | 0          | 1      | 0      | 0              | 0      | 0      | 1            |
| NATIVE HAWAIIAN OR OTHER PACIFIC ISLANDER | 0         | 0      | 0      | 0          | 0      | 1      | 0              | 0      | 0      | 1            |
| OTHER                                     | 0         | 0      | 0      | 0          | 0      | 0      | 0              | 0      | 0      | 0            |
| UNKNOWN                                   | 0         | 0      | 0      | 0          | 0      | 0      | 0              | 0      | 0      | 0            |

Note: This is just an example.  
Source: t\_dm.R 2021-09-08

# PowerPoint Presentation

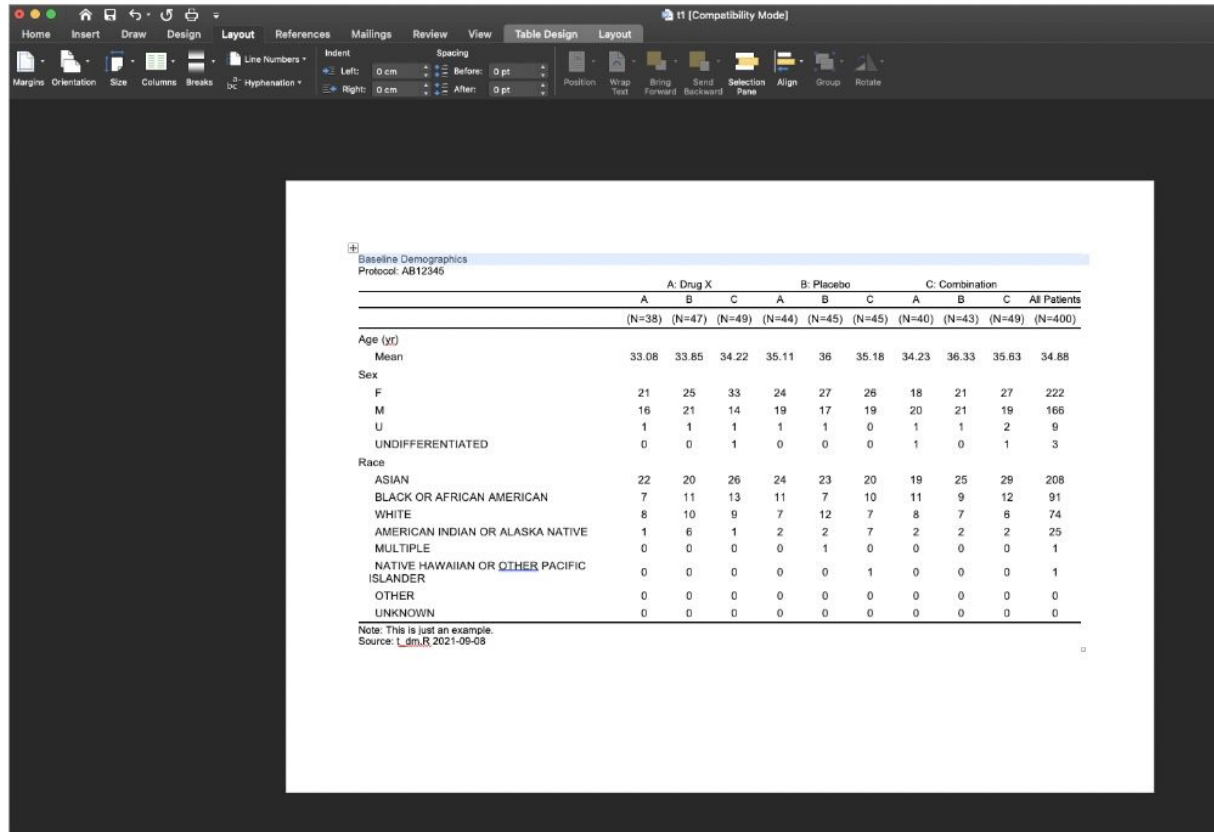


Baseline Demographics  
Protocol: AB12345

|                                           | A: Drug X   |             |             | B: Placebo  |             |             | C: Combination |             |             | All Patients<br>(N=103) |
|-------------------------------------------|-------------|-------------|-------------|-------------|-------------|-------------|----------------|-------------|-------------|-------------------------|
|                                           | A<br>(N=38) | B<br>(N=47) | C<br>(N=49) | A<br>(N=44) | B<br>(N=45) | C<br>(N=45) | A<br>(N=40)    | B<br>(N=43) | C<br>(N=49) |                         |
| Age (yr)                                  |             |             |             |             |             |             |                |             |             |                         |
| Mean                                      | 33.08       | 33.85       | 34.22       | 35.11       | 36          | 35.18       | 34.23          | 36.33       | 35.63       | 34.88                   |
| Sex                                       |             |             |             |             |             |             |                |             |             |                         |
| F                                         | 21          | 25          | 33          | 24          | 27          | 26          | 18             | 21          | 27          | 222                     |
| M                                         | 16          | 21          | 14          | 19          | 17          | 19          | 20             | 21          | 19          | 188                     |
| U                                         | 1           | 1           | 1           | 1           | 1           | 0           | 1              | 1           | 2           | 8                       |
| UNDIFFERENTIATED                          | 0           | 0           | 1           | 0           | 0           | 0           | 1              | 0           | 1           | 3                       |
| Race                                      |             |             |             |             |             |             |                |             |             |                         |
| ASIAN                                     | 22          | 20          | 26          | 24          | 23          | 20          | 19             | 25          | 29          | 208                     |
| BLACK OR AFRICAN AMERICAN                 | 7           | 11          | 13          | 11          | 7           | 10          | 11             | 9           | 12          | 91                      |
| WHITE                                     | 8           | 10          | 9           | 7           | 12          | 7           | 8              | 7           | 6           | 74                      |
| AMERICAN INDIAN OR ALASKA NATIVE          | 1           | 6           | 1           | 2           | 2           | 7           | 2              | 2           | 2           | 23                      |
| MULTIPLE                                  | 0           | 0           | 0           | 0           | 1           | 0           | 0              | 0           | 0           | 1                       |
| NATIVE HAWAIIAN OR OTHER PACIFIC ISLANDER | 0           | 0           | 0           | 0           | 0           | 1           | 0              | 0           | 0           | 1                       |
| OTHER                                     | 0           | 0           | 0           | 0           | 0           | 0           | 0              | 0           | 0           | 0                       |
| UNKNOWN                                   | 0           | 0           | 0           | 0           | 0           | 0           | 0              | 0           | 0           | 0                       |

Note: This is just an example.  
Source: 1\_dm\_R 2021-09-08

# Word Document



11 [Competibility Mode]

Home Insert Draw Design Layout References Mailings Review View Table Design Layout

Margins Orientation Size Columns Breaks Hyphenation

Indent Spacing

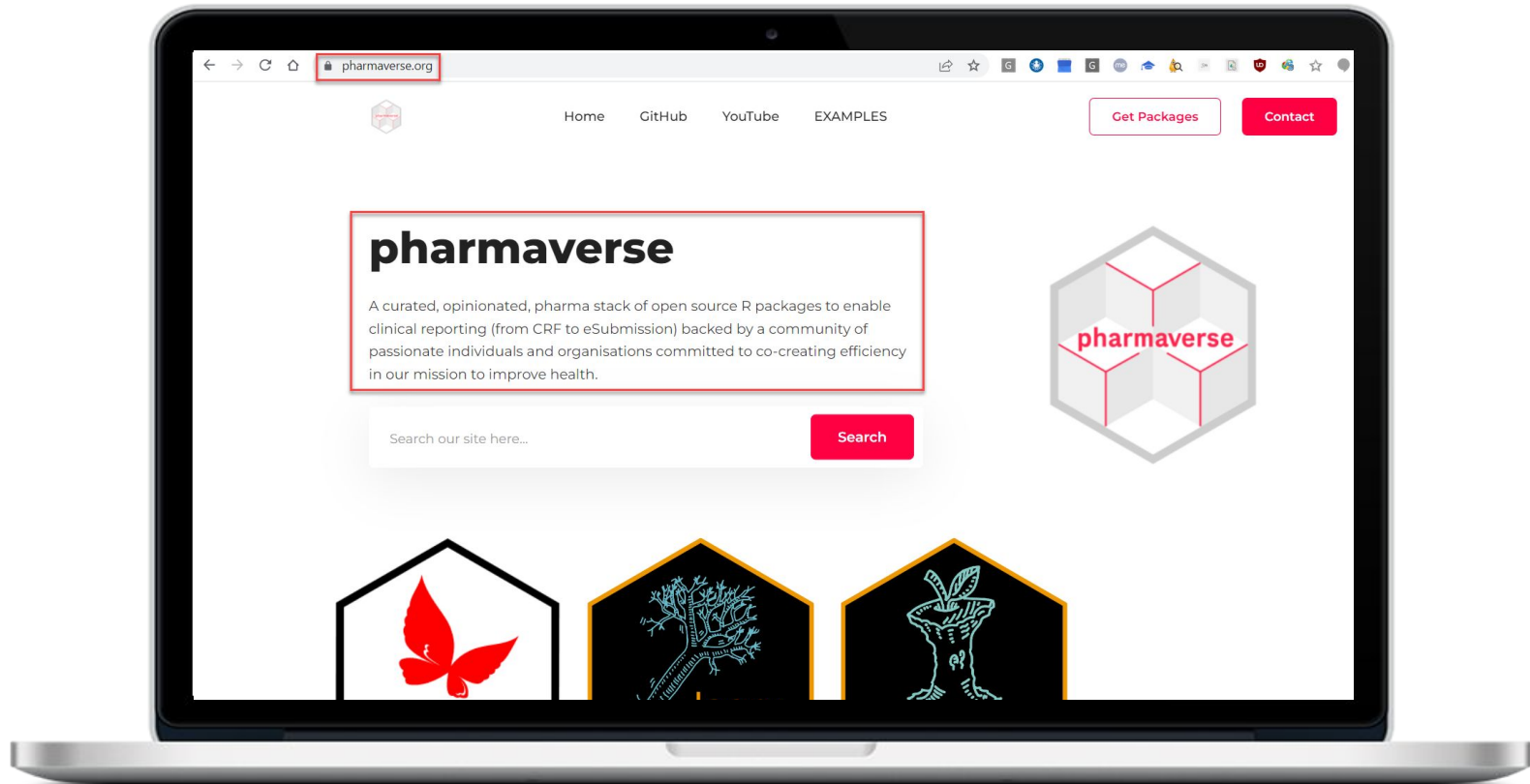
Left: 0 cm Before: 0 pt Right: 0 cm After: 0 pt

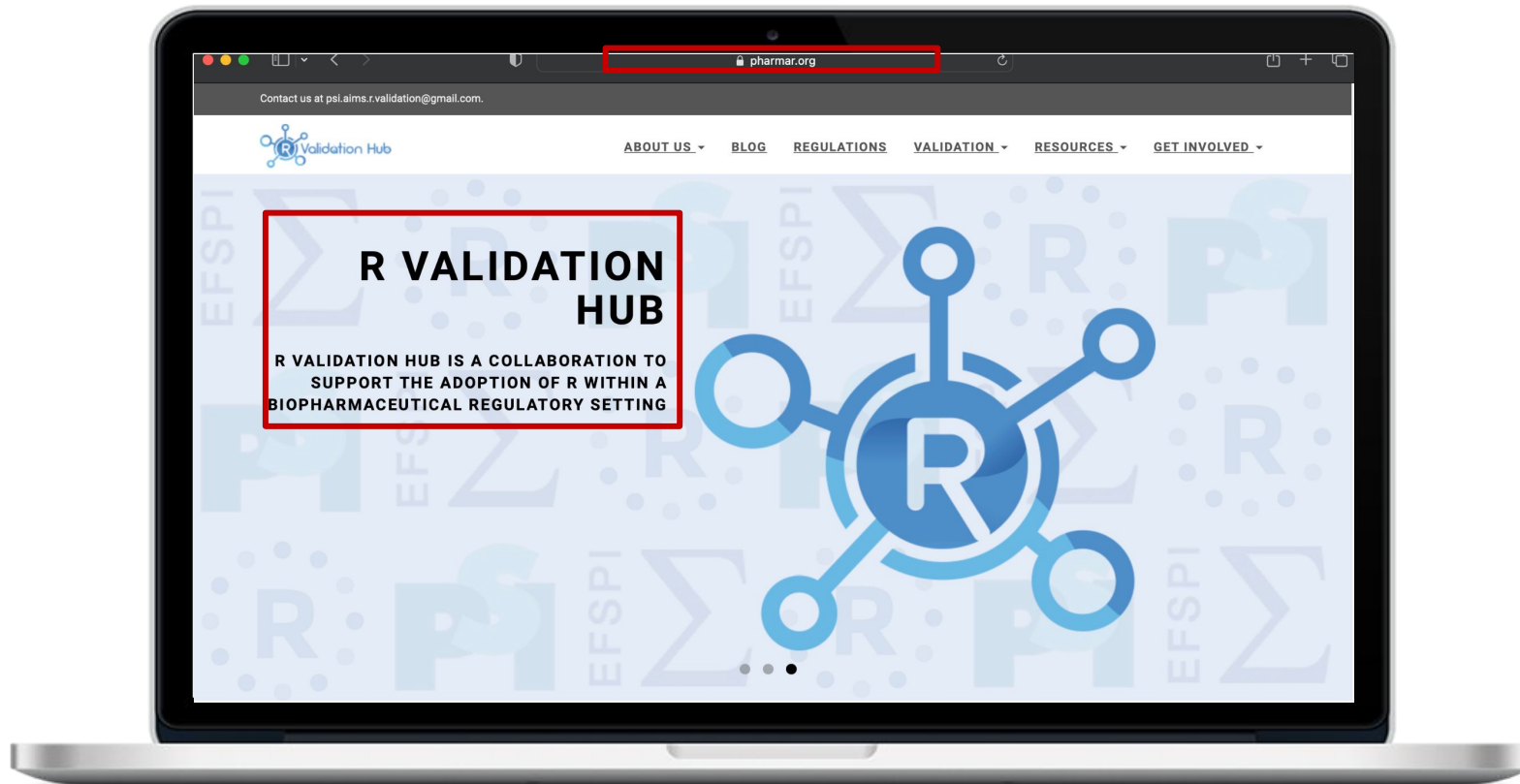
Position Wrap Text Bring Forward Send Backward Selection Pane Align Group Rotate

Baseline Demographics  
Protocol: AB12345

|                                           | A: Drug X |        |        | B: Placebo |        |        | C: Combination |        |         |              |
|-------------------------------------------|-----------|--------|--------|------------|--------|--------|----------------|--------|---------|--------------|
|                                           | A         | B      | C      | A          | B      | C      | A              | B      | C       | All Patients |
|                                           | (N=38)    | (N=47) | (N=49) | (N=44)     | (N=45) | (N=40) | (N=43)         | (N=49) | (N=400) |              |
| Age (yr)                                  |           |        |        |            |        |        |                |        |         |              |
| Mean                                      | 33.08     | 33.85  | 34.22  | 35.11      | 36     | 35.18  | 34.23          | 36.33  | 35.63   | 34.88        |
| Sex                                       |           |        |        |            |        |        |                |        |         |              |
| F                                         | 21        | 25     | 33     | 24         | 27     | 26     | 18             | 21     | 27      | 222          |
| M                                         | 16        | 21     | 14     | 19         | 17     | 19     | 20             | 21     | 19      | 166          |
| U                                         | 1         | 1      | 1      | 1          | 1      | 0      | 1              | 1      | 2       | 9            |
| UNDIFFERENTIATED                          | 0         | 0      | 1      | 0          | 0      | 0      | 1              | 0      | 1       | 3            |
| Race                                      |           |        |        |            |        |        |                |        |         |              |
| ASIAN                                     | 22        | 20     | 26     | 24         | 23     | 20     | 19             | 25     | 28      | 208          |
| BLACK OR AFRICAN AMERICAN                 | 7         | 11     | 13     | 11         | 7      | 10     | 11             | 9      | 12      | 91           |
| WHITE                                     | 8         | 10     | 9      | 7          | 12     | 7      | 8              | 7      | 6       | 74           |
| AMERICAN INDIAN OR ALASKA NATIVE          | 1         | 6      | 1      | 2          | 2      | 7      | 2              | 2      | 2       | 25           |
| MULTIPLE                                  | 0         | 0      | 0      | 0          | 1      | 0      | 0              | 0      | 0       | 1            |
| NATIVE HAWAIIAN OR OTHER PACIFIC ISLANDER | 0         | 0      | 0      | 0          | 0      | 1      | 0              | 0      | 0       | 1            |
| OTHER                                     | 0         | 0      | 0      | 0          | 0      | 0      | 0              | 0      | 0       | 0            |
| UNKNOWN                                   | 0         | 0      | 0      | 0          | 0      | 0      | 0              | 0      | 0       | 0            |

Note: This is just an example.  
Source: Ldm\_R 2021-09-08





# Roche Product Development Data Sciences Strategic Framework



## Disease and Molecule Strategies

We develop, recommend, and influence strategic options that **drive smarter drug development** and patient access.

*We contribute to disease insights, novel endpoints, external collaborations, fast and flexible study designs, efficient and sensitive analytics, rigor in interpretation and decision-making*

## Study Design and Implementation

We promote and apply rigor to study design and drive **quality and efficiency** in implementation to generate compelling evidence.

*We contribute to over 260 molecule projects through innovative talent, flexible staffing, agile execution, growing automation and prospective FAIR*

## Enhanced Data & Insights Sharing

We enable the sharing of **RWD & RCT data** and valuable insights across the enterprise to advance research, development, and medical practice.

*We created 15 data marts with 165 studies and 242k+ patients, which have driven 15 early/late R&D decisions; we host 147 databases for use across Roche*

## Data Science Talent and Culture

Because **Data Science talent** is key for Roche, we attract and develop great talent, and nurture a culture and environment that enables our best work.

*We attract top talent with diverse skill sets globally, develop them with shared sense of purpose and flow to work mindset, build data science communities across Roche*