# Layered Design for R Package Development: Meeting the Needs of Pharmaceutical R&D Stakeholders

**Jean Muller & Ligia Adamska, MSD**

**UseR! | 8-11 July 2024 | Salzburg, Austria**

# Agenda

1. Context
2. Problem Statement
3. Layered Design
   - 3a - Verb Layer
   - 3b - Wrapper Layer
4. Case Study: Health Technology Assessment Analysis
5. Results and Discussion
6. Conclusion

# Context: Introduction to Health Technology Assessment (HTA)

| | | |
|---|---|---|
| Non-clinical | • Drug discovery<br>• Identification of candidate molecule<br>• Non-clinical testing<br>• Safety and efficacy tests | **Average: 3.5 years** (could be 4–5.5 years) |
| Clinical | Phase I (dosing and safety)<br>Phase II/III (safety and efficacy) | **Average: 1 year**<br>Average Phase 2: 2 years<br>Average Phase 3: 3 years |
| Regulatory/ HTA | Marketing authorization request<br>**HTA** | **Average: 2.5 years (USA)**<br>**Average: 1.5 years (UK)** |
| Post-marketing | Phase IV (long-term impacts)<br>Conditional approval requiring additional/long-term data | **Average total time: 12 years** |

## Typical HTA question:

*"What is the added benefit of a newly approved drug compared with conventional drugs?"*


Source: EMA (2017) Lifecycle of a medicinal product. Available at: https://www.ema.europa.eu/en/documents/presentation/presentation-howare-medicines-evaluated-ema-nathalie-bere_en.pdf; Internal MSD 'what is HTA' video.

# Context: Analysis & Reporting (A&R) Workflow in HTA

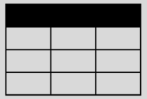
## 1 Data

- Standardized (CDISC)
- Mature (gone through extensive pre-processing steps to make it analysis ready)

## 2 Project setup (R)

- R Dependency management
- Version control

## 3 Analysis planning



- List of unique Tables, Figures and Listings (TFL) to produce
- PICOs of interest

## 4 Analysis generation


- TFL generation
- Scalability to support outputs per PICO
- Traceability and reproducibility
- Validation

## 5 Reporting

- Assemble outputs into a report
- Customer review
- Submission

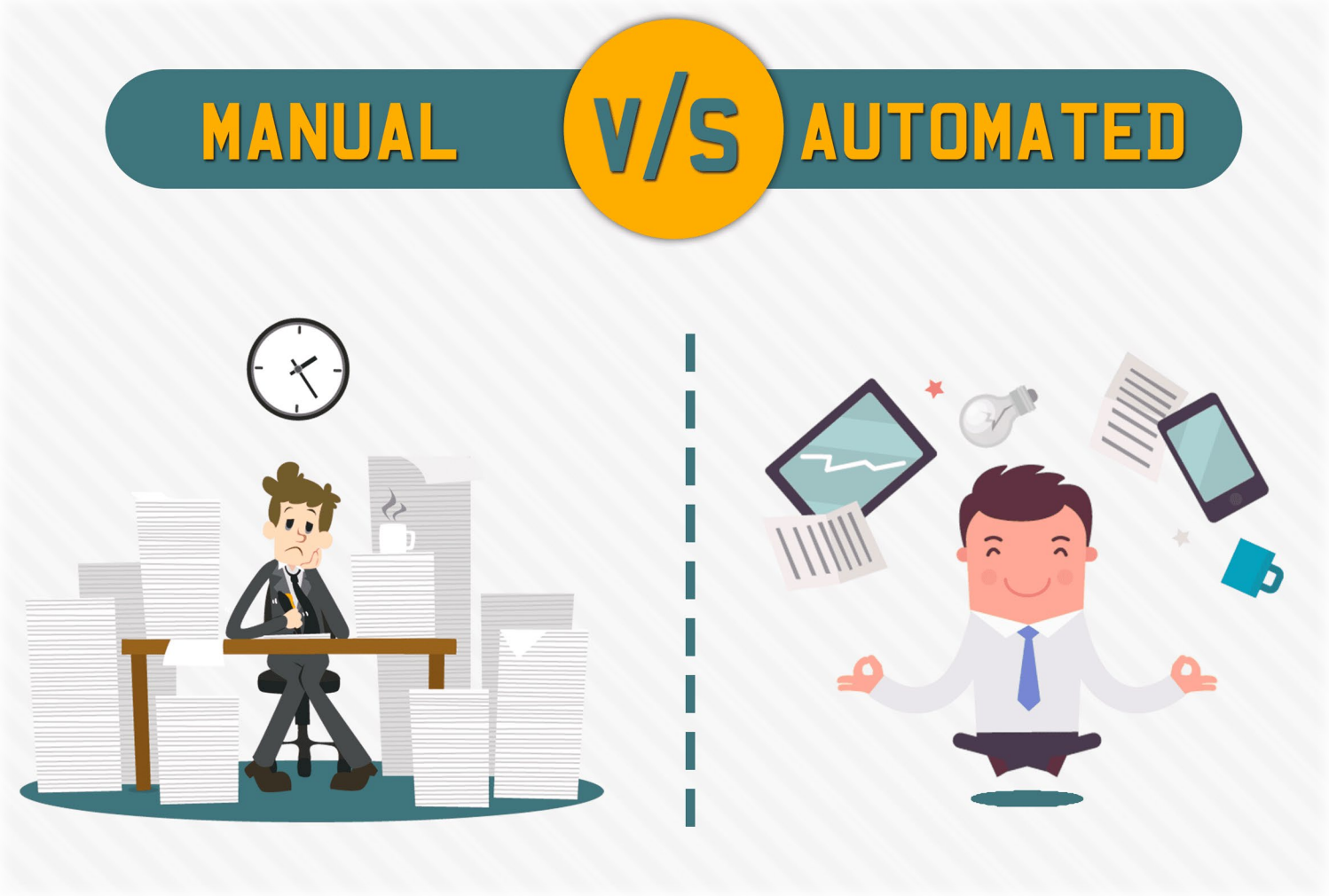
*PICO: Population Intervention Comparator Outcome*

# Context: Focus on step 4

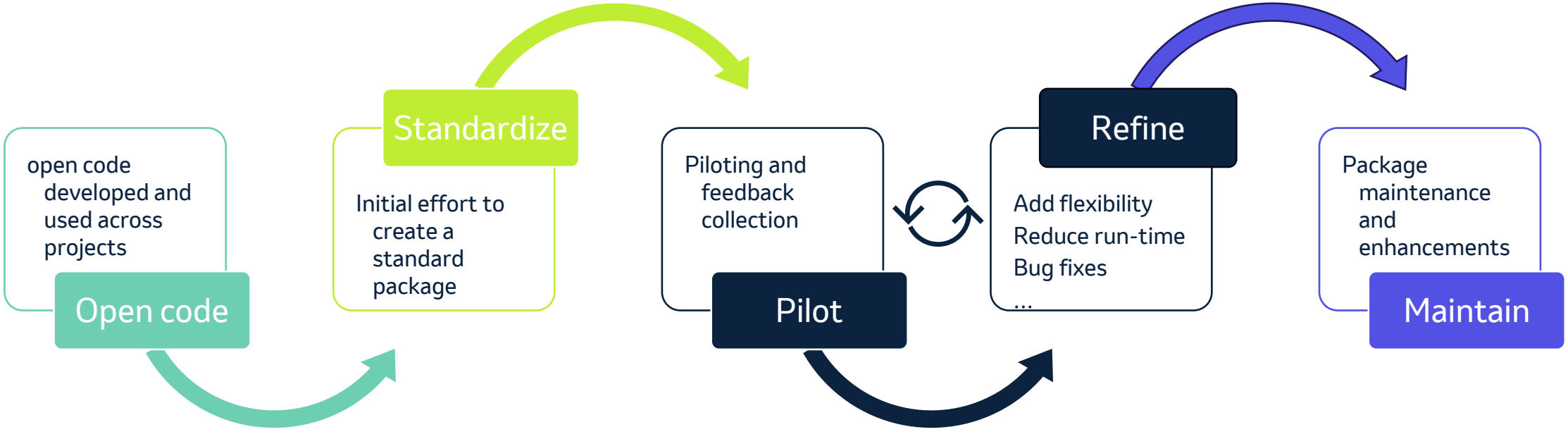
| 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|
| Data | Project setup (R) | Analysis planning | Analysis generation | Reporting |
|  |  |  |  |  |
| - Standardized (CDISC)<br>- Mature (gone through extensive pre-processing steps to make it analysis ready) | - R Dependency management<br>- Version control | - List of unique Tables, Figures and Listings (TFL) to produce<br>- PICOs of interest | - TFL generation<br>- Scalability to support outputs per PICO<br>- Traceability and reproducibility<br>- Validation | - Assemble outputs into a report<br>- Customer review<br>- Submission |

*PICO: Population Intervention Comparator Outcome*

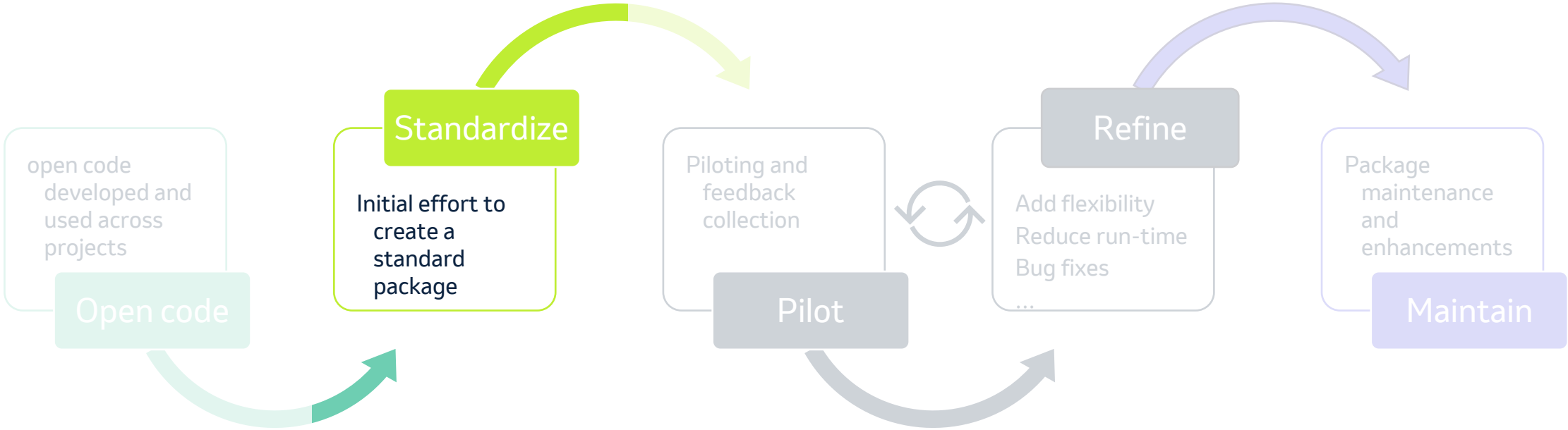
# Context: Opportunity



# Context: Focus on standardization



# Context: Focus on standardization



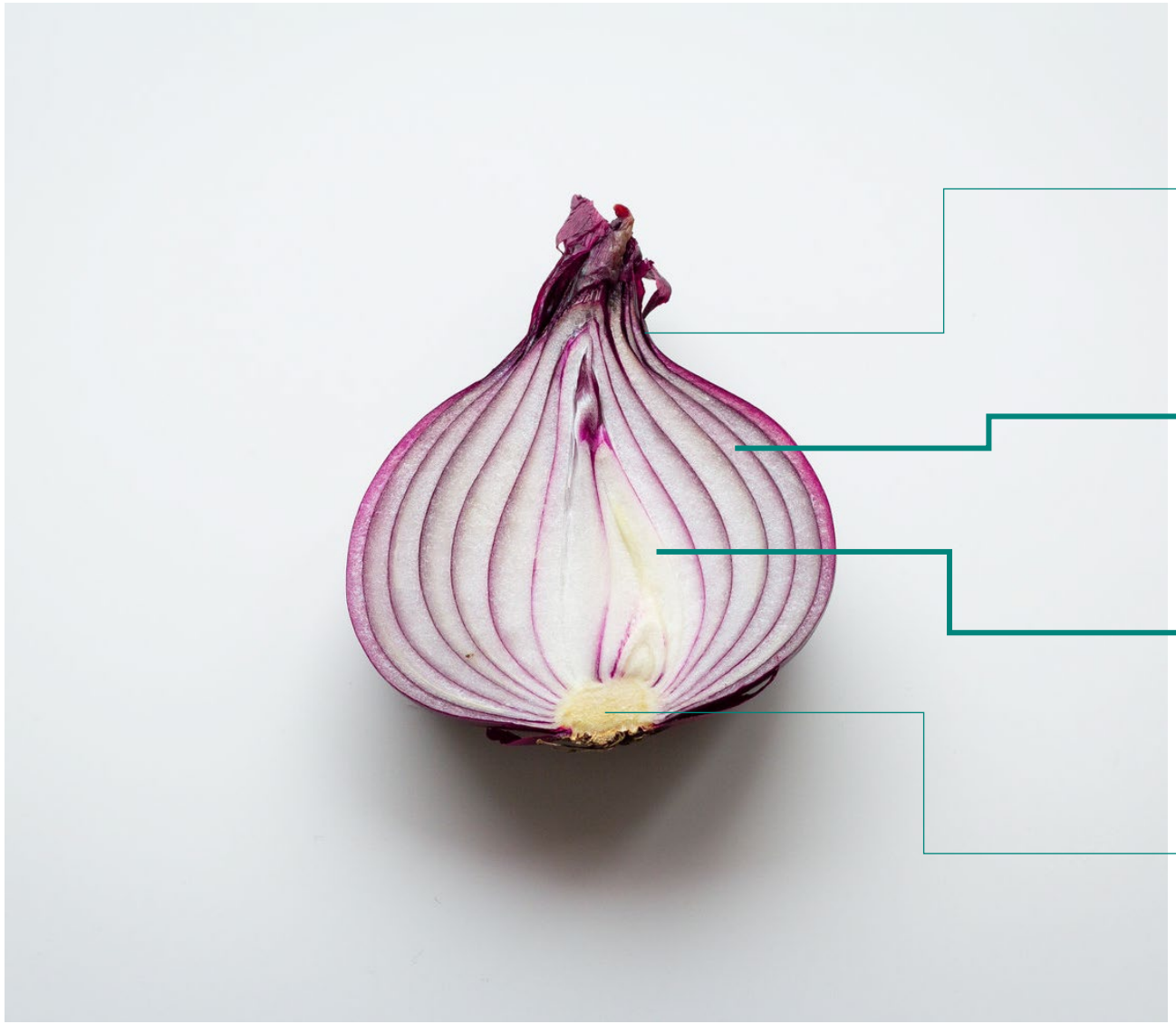
# Problem statement

*How to design an A&R R package to minimize manual tasks for everyone involved?*

The challenge is to have **both** **flexible functions** and a systematic **structure to support automatic reporting** systems.

Without duplication of work.

# Layered Design



Automated Reporting System / UIs

**Wrapper Layer**

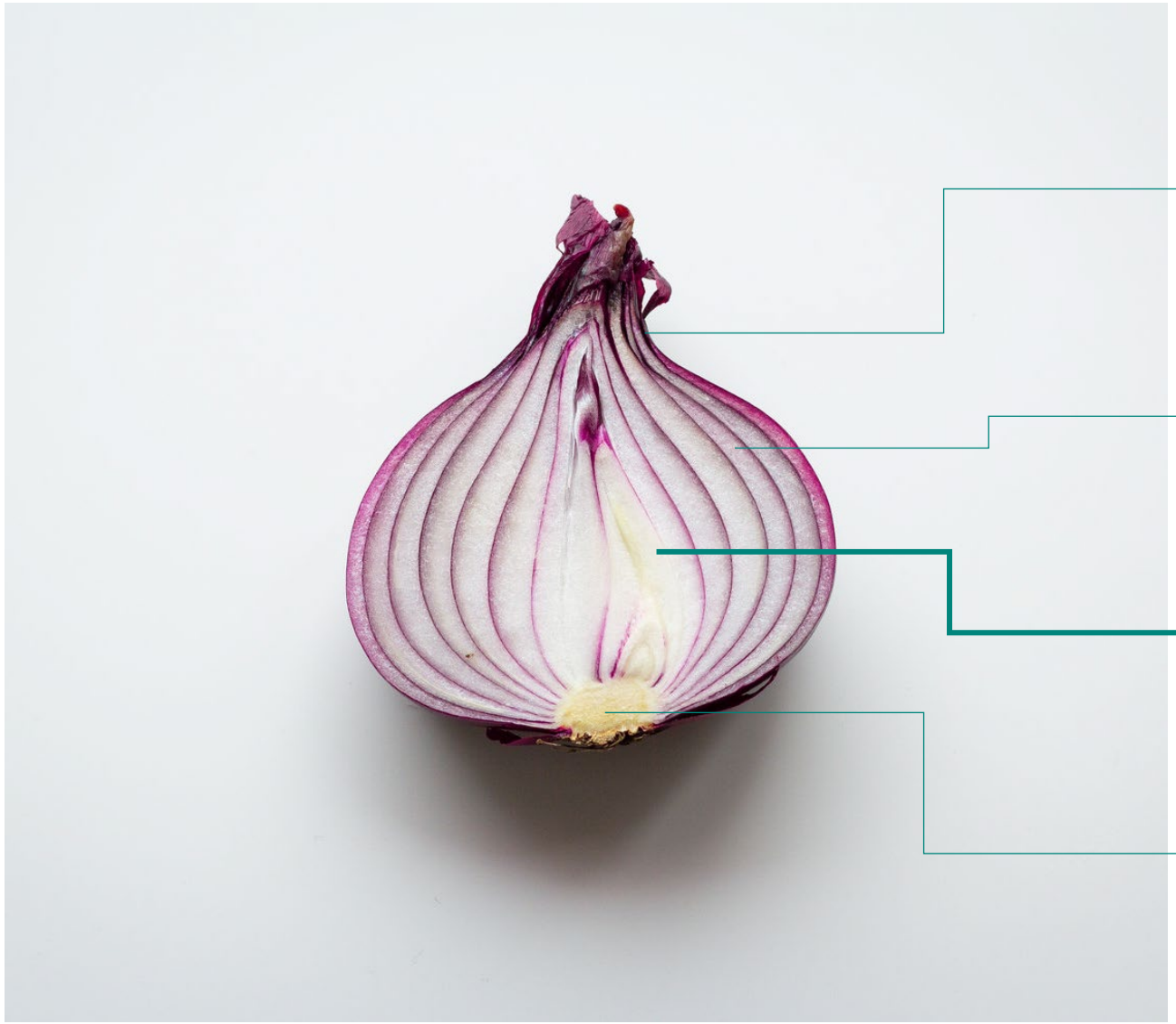
Reporting functions that leverages on verb layer

**Verb Layer**

Individual functions that correspond to each part of the workflow

Core External Packages

# Layered Design



Automated Reporting System / UIs

**Wrapper Layer**

Reporting functions that leverages on verb layer

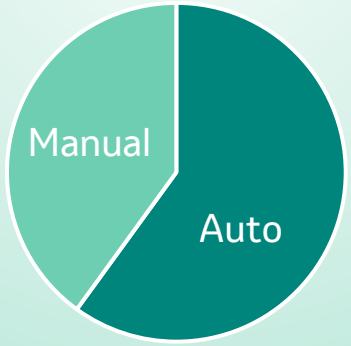
**Verb Layer**

Individual functions that correspond to each part of the workflow

Core External Packages

# Verb Layer: Concept

- Follows tidy design principles
- Embraces functional programing
- Enables pipeable functions
- Provides flexibility

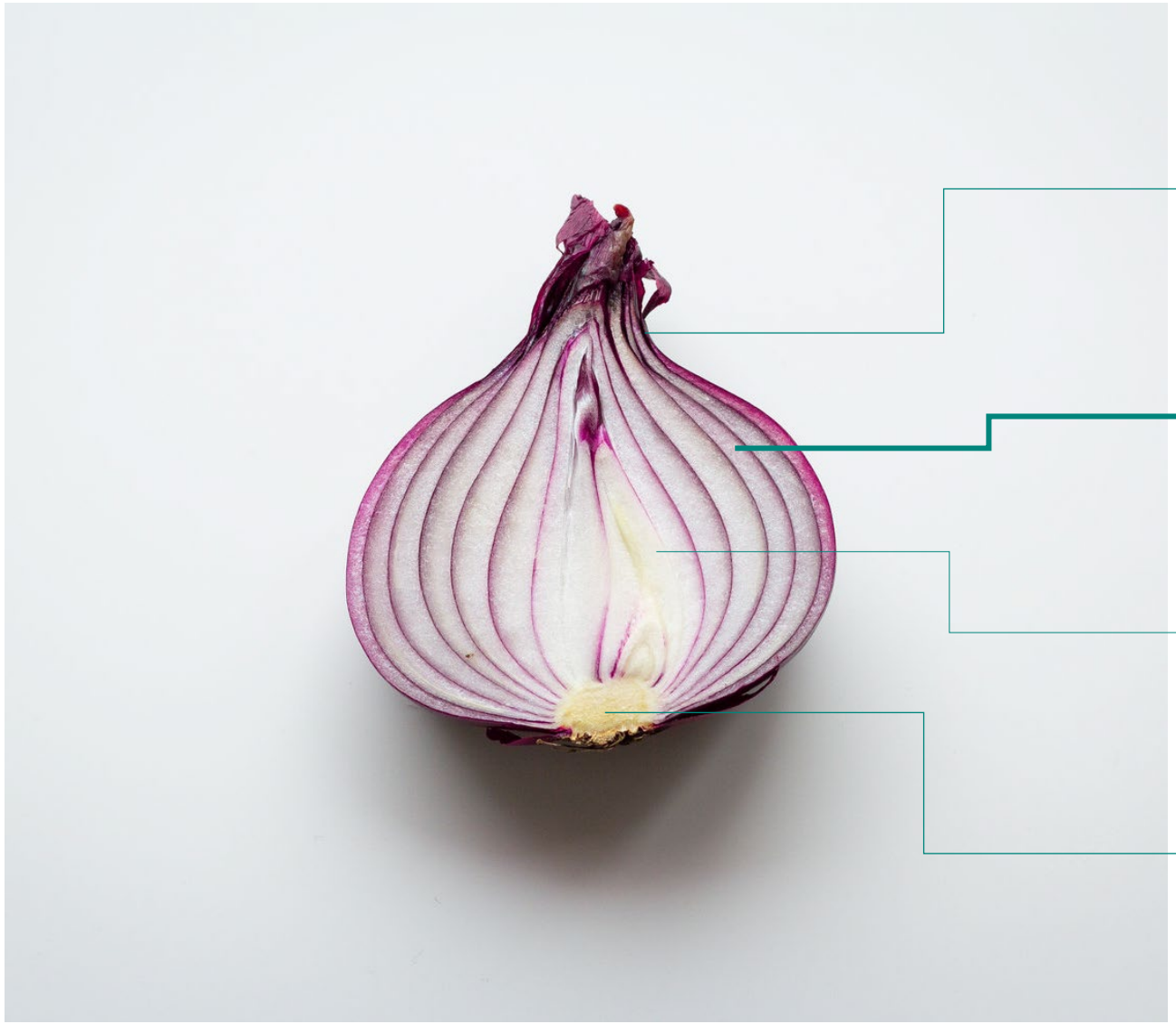



```
verb_function_1(…) |>
      verb_function_2(…) |>
      verb_function_3(…) |>
        …
```

**Example:**

```
1  # Application of pipe-able functions:
2  data |>
3    run_models_flexsurvreg() |>
4    extract_gof() |>
5    format_gof()
6
7
8  data |>
9    run_models_flexsurvreg() |>
10   extract_est() |>
11   format_est()
```

# Layered Design



Automated Reporting System / UIs

**Wrapper Layer**

Reporting functions that leverages on verb layer

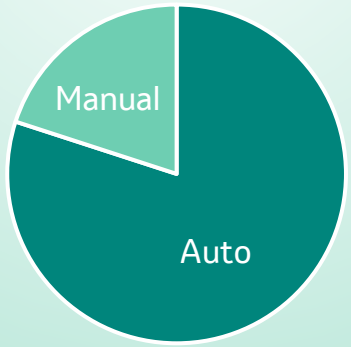
**Verb Layer**

Individual functions that correspond to each part of the workflow

Core External Packages

# Wrapper Layer: Concept

- Integrates verb layer into wrapper layer
- Generates standard analyses with one call
- Adds agreed upon metadata format
- Analysis repeatability for different PICOs




```
wrapper_function(…){
  verb_function_1(…) |>
       verb_function_2(…) |>
       verb_function_3(…) |>
         …
}
```

# Case Study: Health Technology Assessment Analysis

## *Context*

- *In Health Technology Assessment (HTA), to inform economic models, we are interested to extrapolate survival beyond trial follow-up*
- *Several R packages exist to perform the analysis, but all require several steps to arrive at the standard output and data specific issues can occur along the way*

## *Opportunity*

*Develop a tool to support generation of the standard AND be used for exploration and project specific needs*

## *Solution*

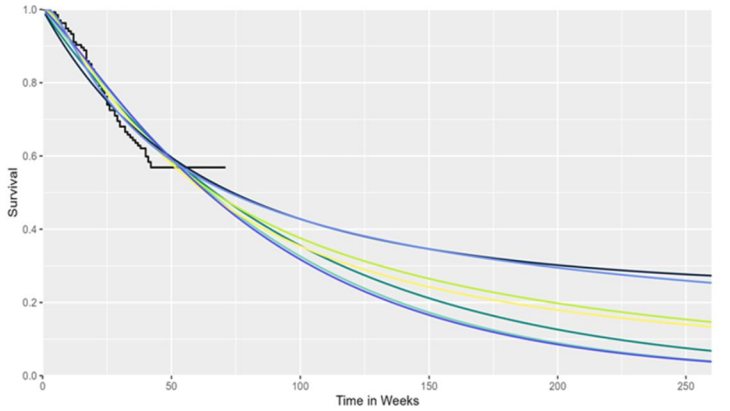
*Leverage on the Layered Design*

Goodness of Fit Statistics Resulting from Independent Standard
for <Endpoint> in A: Drug X
Subtitle 1
(Population)

| Survival Model | AIC | BIC | Mean AIC BIC | Rank AIC | Rank BIC | Rank Mean AIC BIC |
|---|---|---|---|---|---|---|
| Exponential | XXX.X | XXX.X | XXX.X | X | X | X |
| Weibull | XXX.X | XXX.X | XXX.X | X | X | X |
| Log-normal | XXX.X | XXX.X | XXX.X | X | X | X |
| Log-logistic | XXX.X | XXX.X | XXX.X | X | X | X |
| Gompertz | XXX.X | XXX.X | XXX.X | X | X | X |
| Gamma | XXX | | | | | |
| Generalized Gamma | XXX | | | | | |

Footnote 1
Footnote 2
Database Cutoff Date: XXX

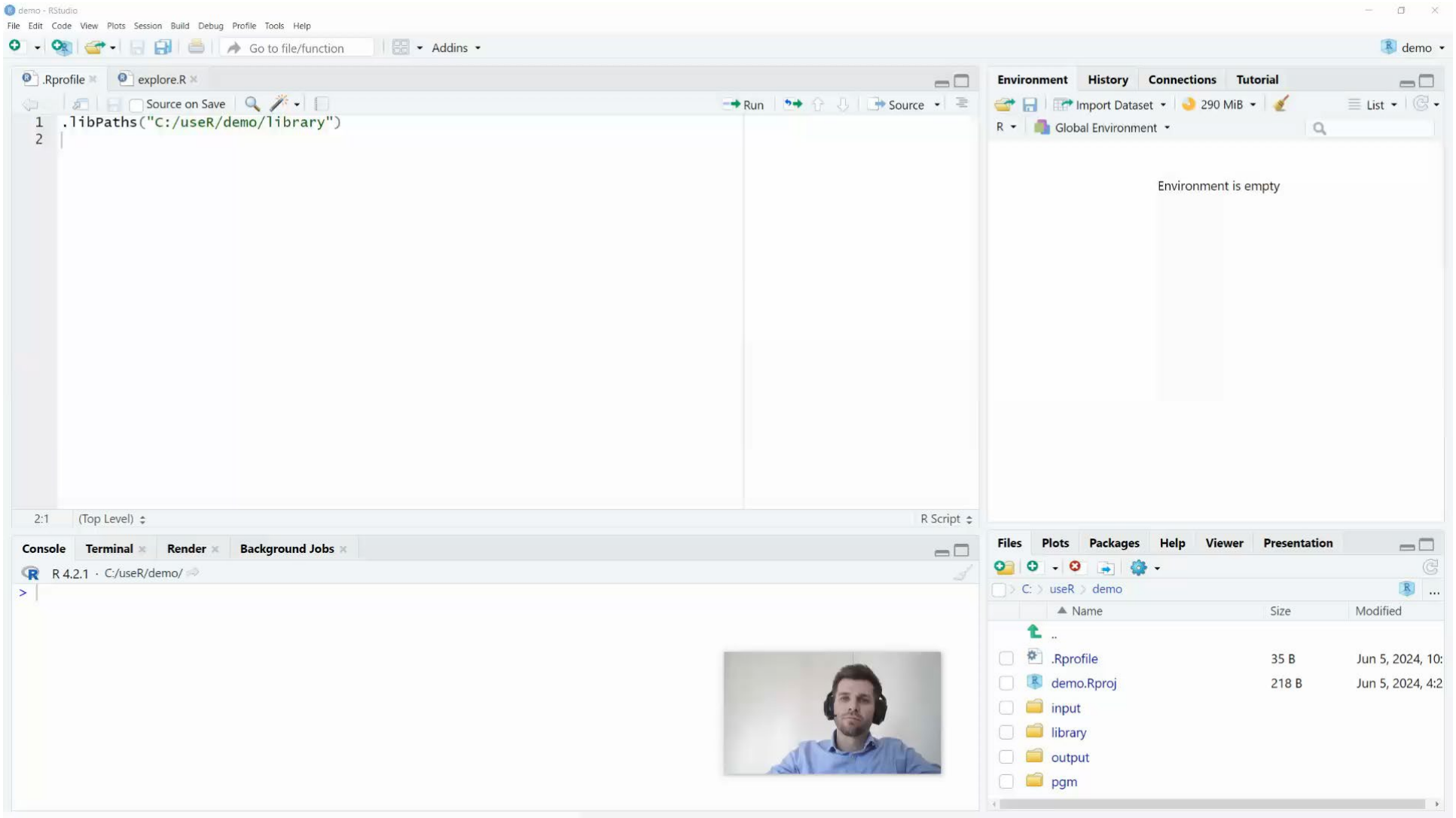
Estimated Long Term Survival Probability (<X> Weeks) for <Endpoint> from Independent Standard Parametric Models Against Empirical Survival Probability from Kaplan-Meier Analysis During the Study Period in A: Drug X
Subtitle 1
(Population)



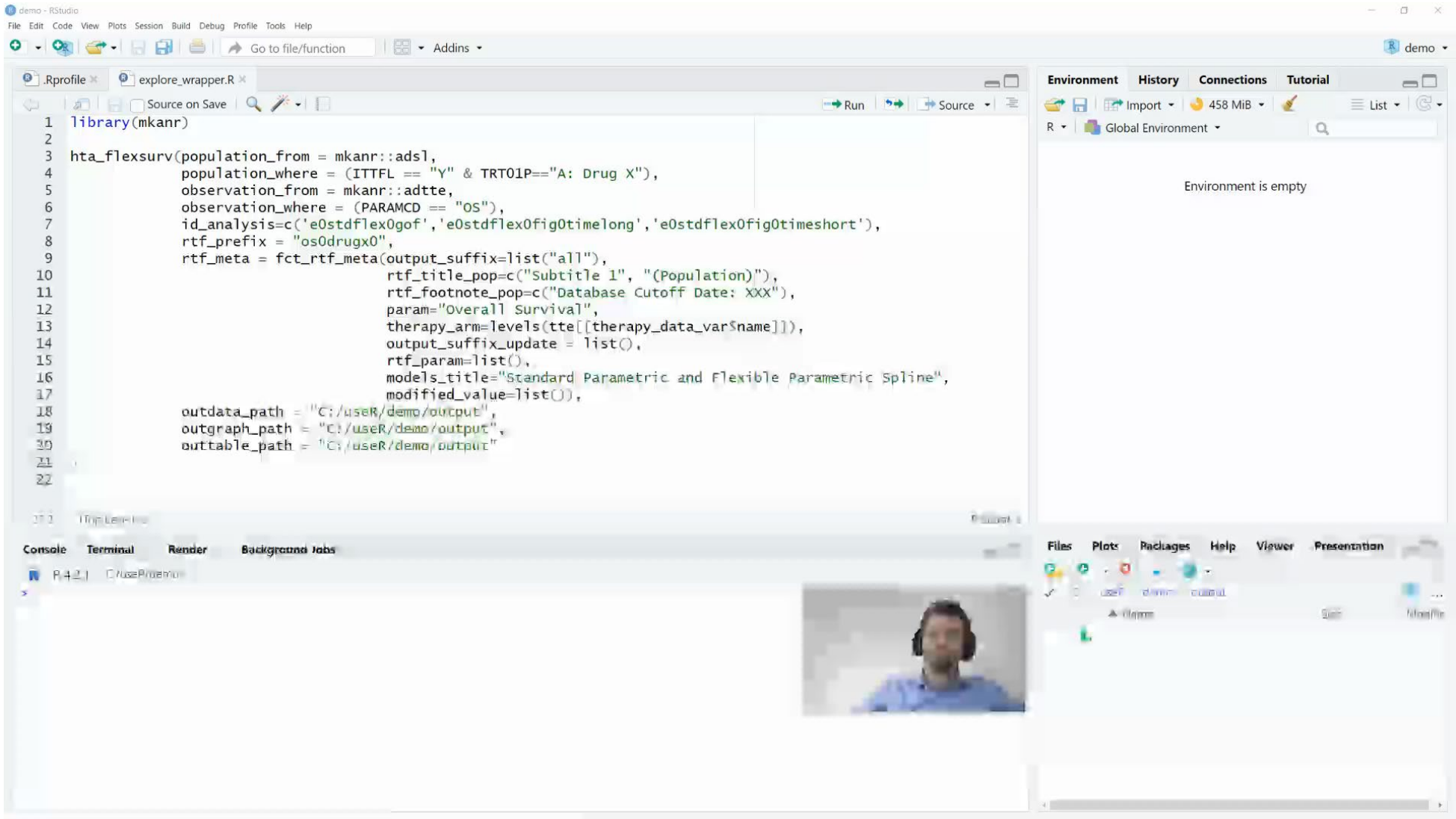

Footnote 1
Footnote 2
Database Cutoff Date: XXX

Check out demo package on Github!

# Analysis outputs with a single call: Verb Layer



# Analysis outputs with a single call: Wrapper Layer



# Results and Discussion

## Benefits ✓

- Flexible and customizable
- Automated, saving time and improving efficiency
- Easy integration with reporting systems (e.g. UIs)
- Streamlined and optimized
- Allows easy replacement/updating of modules
- Caters for expert R programmers (verb layer) and R novices (wrapper layer)

## Trade-offs ✗

- Maintenance and within package dependencies
- Monitoring of core external packages (e.g. new options)
- Wrapper flexibility limited to the agreed upon standard

# Conclusion

Layered R package involves organizing the code and functionality into distinct layers, each serving a specific purpose.

This design allows for modularity, reusability, and scalability

Looking forward:

- Add simple user interface layer
- Leverage more on Object Oriented Programming

https://www.msd.com/

# Thank you

Jean Muller
Senior Statistical Programmer
jean.muller@msd.com

Ligia Adamska
Associate Director, Statistical Programming
ligia.adamska@msd.com