A Risk-based approach for assessing R package accuracy within a validated infrastructure

Paulo Bargo, Janssen R&D; Juliane Manitz, EMD Serono; Andy Nicholls, GSK; Doug Kelkhoff, Roche; Yilong Zhang, Merck & Co., Inc.; Lyn Taylor , Phastar; Joe Rickert, R Consortium; Marly Gotti, Biogen; Keaven M Andersen, Merck & Co. Inc.

~~How do I validate R?~~

Is my software reliable?

Reliable Software

5.8 Integrity of Data and Computer Software Validity

The credibility of the numerical results of the analysis depends on the **quality and validity of the methods and software** (both internally and externally written) used both for data management (data entry, storage, verification, correction and retrieval) and also for processing the data statistically. Data management activities should therefore be based on thorough and effective standard operating procedures. **The computer software used for data management and statistical analysis should be reliable, and documentation of appropriate software testing procedures should be available.**

The credibility of the numerical results of the analysis depends on the quality and validity of the methods and software (both internally and externally written) used both for data management (data entry, storage, verification, correction and retrieval) and also for processing the data statistically. Data management activities should therefore be based on thorough and effective standard operating procedures. The computer software used for data management and statistical analysis should be reliable, and documentation of appropriate software testing procedures should be available.

VI EVALUATION OF SAFETY AND TOLERABILITY

6.1 Scope of Evaluation

In all clinical trials evaluation of safety and tolerability (see Glossary) constitutes an

ICH E9: https://www.ema.europa.eu/en/documents/scientific-guideline/ich-e-9-statistical-principles-clinical-trials-step-5_en.pdf

U.S. FOOD & DRUG ADMINISTRATION

Statistical Software Clarifying Statement

FDA does not require use of any specific software for statistical analyses, and statistical software is not explicitly discussed in Title 21 of the Code of Federal Regulations [e.g., in 21CFR part 11]. However, the software package(s) used for statistical analyses should be fully documented in the submission, including version and build identification.

As noted in the FDA guidance, *E9 Statistical Principles for Clinical Trials* (available at http://www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/default.htm), "The computer software used for data management and statistical analysis should be reliable, and documentation of appropriate software testing procedures should be available." Sponsors are encouraged to consult with FDA review teams and especially with FDA statisticians regarding the choice and suitability of statistical software packages at an early stage in the product development process.

May 6, 2015

Source: https://www.fda.gov/downloads/ForIndustry/DataStandards/StudyDataStandards/UCM587506.pdf

Example

CODE LOG RESULTS OUTPUT DATA

```
1  data simple;
2  input values;
3  datalines;
4  1
5  2
6  3
7  4
8  5
9  ;
10 run;
11
12 proc means data=simple;
13 var values;
14 run;
15
```

The MEANS Procedure

Analysis Variable : values				
N	Mean	Std Dev	Minimum	Maximum
5	3.0000000	1.5811388	1.0000000	5.0000000

Is this reliable?

Reliable Software

Why do we trust the summary?

- *"I know the actual answer"*
- *"It's in the ballpark of what I might expect to see"*
- *"I've used the software before and it did what I expected"*
- *"Many others use the software and it does what they expect"*
- *"When I learnt statistics, I was taught using the software"*
- *"The software is used/cited in statistical literature"*
- *"I trust that the software owner develops it using best practice"*
- *"The software owner provides tests that I can use to verify that it is working"*

Intuition

Community Exposure

Developer SDLC

Example

```
library(dplyr)
simple <- tibble(values=1:5)

simple %>%
  summarise(N = n(),
            Mean = mean(values),
            Std_Dev = sd(values),
            Min = min(values),
            Max = max(values))
```

```
# A tibble: 1 x 5
      N  Mean Std_Dev   Min   Max
  <int> <dbl>   <dbl> <int> <int>
1     5     3    1.58     1     5
```

Is this reliable?

Reliable Software

Why do we trust the summary?

- ✓ *"I know the actual answer"*
- ✓ *"It's in the ballpark of what I might expect to see"*
- ✓ *"I've used the software before and it did what I expected"*
- ✓ *"Many others use the software and it does what they expect"*
- ✓ *"When I learnt statistics, I was taught using the software"*
- *"The software is used/cited in statistical literature"*
- *"I trust that the software owner develops it using best practice"*
- *"The software owner provides tests that I can use to verify that it is working"*

Intuition

Community Exposure

Developer SDLC

Reliable Software

Why do we trust the summary?

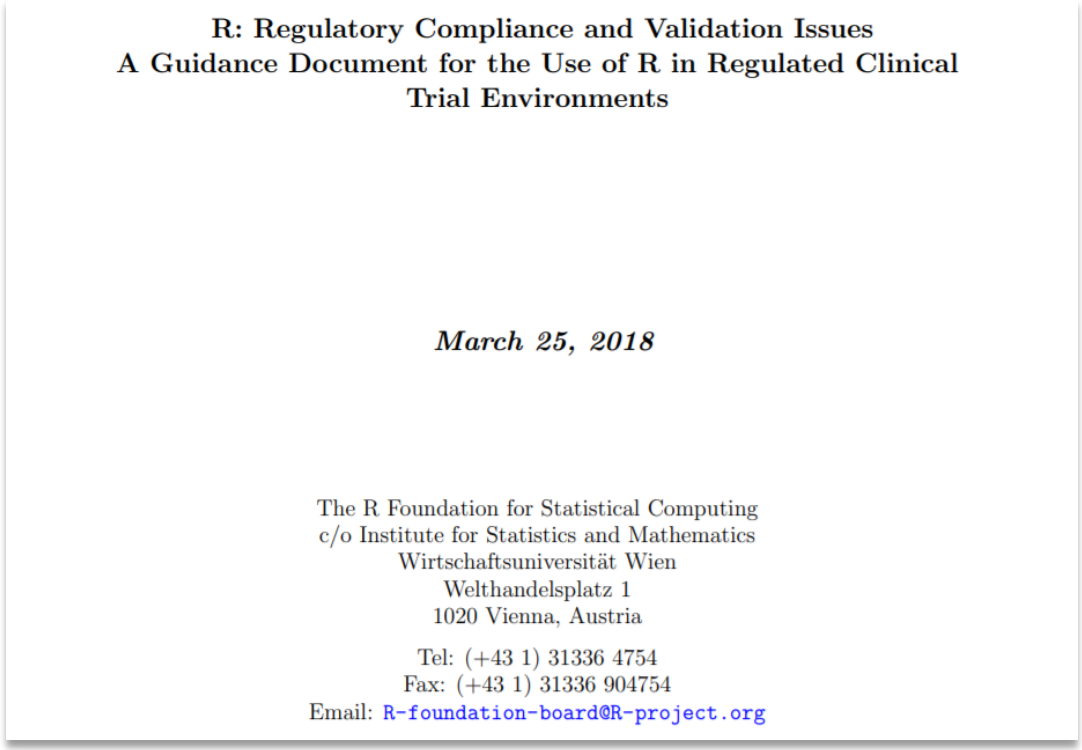
- ✓ *"I know the actual answer"*
- ✓ *"It's in the ballpark of what I might expect to see"*
- ✓ *"I've used the software before and it did what I expected"*
- ✓ *"When I learnt statistics, I was taught using the software"*
- ✓ *"Many others use the software and it does what they expect"*
- ✓ *"The software is used/cited in statistical literature"*
- *"I trust that the software owner develops it using best practice"*
- *"The software owner provides tests that I can use to verify that it is working"*

Intuition

Community Exposure

Developer SDLC

I trust that the software owner develops it using best practice



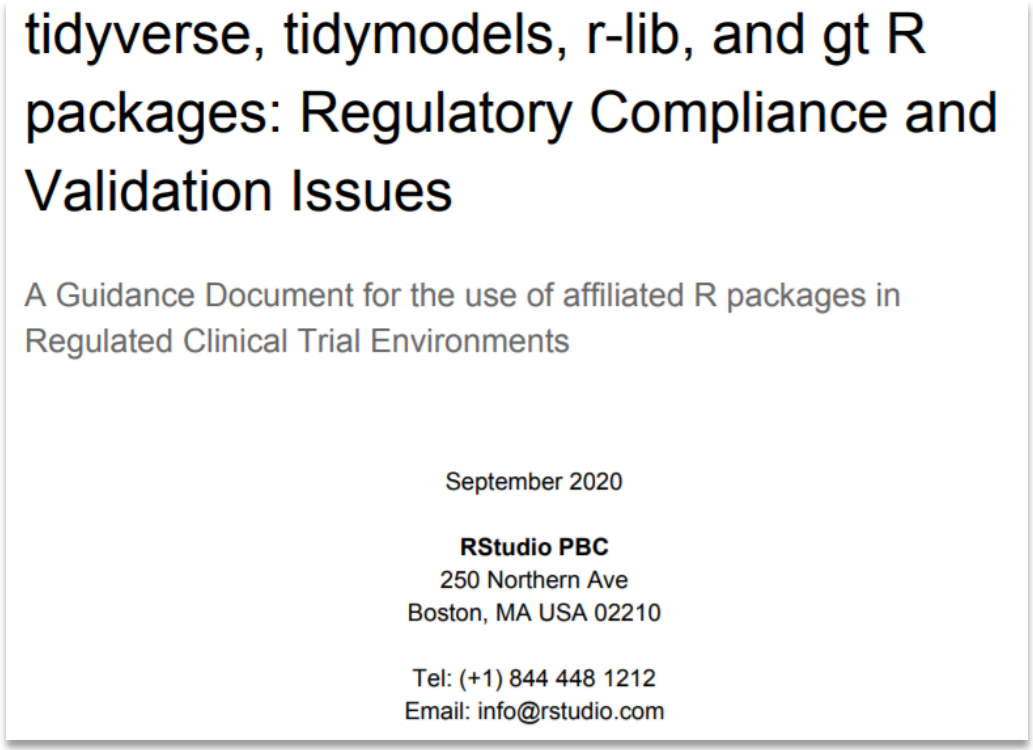
R: Regulatory Compliance and Validation Issues
A Guidance Document for the Use of R in Regulated Clinical Trial Environments

March 25, 2018

The R Foundation for Statistical Computing
c/o Institute for Statistics and Mathematics
Wirtschaftsuniversität Wien
Welthandelsplatz 1
1020 Vienna, Austria

Tel: (+43 1) 31336 4754
Fax: (+43 1) 31336 904754
Email: R-foundation-board@R-project.org

https://www.r-project.org/doc/R-FDA.pdf



tidyverse, tidymodels, r-lib, and gt R packages: Regulatory Compliance and Validation Issues

A Guidance Document for the use of affiliated R packages in Regulated Clinical Trial Environments

September 2020

RStudio PBC
250 Northern Ave
Boston, MA USA 02210

Tel: (+1) 844 448 1212
Email: info@rstudio.com

https://resources.rstudio.com/assets/img/validation-tidy.pdf

Reliable Software

Why do we trust the summary?

- ✓ *"I know the actual answer"*
- ✓ *"It's in the ballpark of what I might expect to see"*
- ✓ *"I've used the software before and it did what I expected"*
- ✓ *"When I learnt statistics, I was taught using the software"*
- ✓ *"Many others use the software and it does what they expect"*
- ✓ *"The software is used/cited in statistical literature"*
- ✓ *"I trust that the software owner develops it using best practice"*
- *"The software owner provides tests that I can use to verify that it is working"*

Intuition

Community Exposure

Developer SDLC

dplyr

CRAN 1.0.5 | R-CMD-check passing | codecov 85%

Overview

master ▾ dplyr / tests / testthat /

Go to file | Add file ▾ | ···

romainfrancois using pillar::format_glimpse() (#5845) ... ✓ 8b036bc 6 days ago History

..		
_snaps	using pillar::format_glimpse() (#5845)	6 days ago
.gitignore	Upgrade to modern testthat dir structure.	7 years ago
helper-dplyr.R	Supply proper caller environments (#5824)	21 days ago
helper-encoding.R	Refactor encoding test helpers and use withr::local_locale() (#5236)	11 months ago
helper-s3.R	Fix `arrange()` issue with unruly class	11 months ago
helper-torture.R	Fix slice() error with gctorture() (#2794)	4 years ago
test-DBI.R	using test that 3rd edition (#5731)	2 months ago
test-across.R	Fix quosure propagation in `across()`	13 days ago
test-all-equal.r	using test that 3rd edition (#5731)	2 months ago
test-arrange.r	using test that 3rd edition (#5731)	2 months ago
test-bind.R	using test that 3rd edition (#5731)	2 months ago
test-case-when.R	using test that 3rd edition (#5731)	2 months ago

Reliable Software

Why do we trust the summary?

- ✓ *"I know the actual answer"*
- ✓ *"It's in the ballpark of what I might expect to see"*
- ✓ *"I've used the software before and it did what I expected"*
- ✓ *"When I learnt statistics, I was taught using the software"*
- ✓ *"Many others use the software and it does what they expect"*
- ✓ *"The software is used/cited in statistical literature"*
- ✓ *"I trust that the software owner develops it using best practice"*
- ✓ *"The software owner provides tests that I can use to verify that it is working"*

Intuition

Community Exposure

Developer SDLC

So why is it so difficult to use Open Source languages for GxP analyses?!

Challenge 1: The R Ecosystem

- **Core R** (Base+Recommended) - Low risk

- **Contributed** - Variable risk
  - Many different authors
  - Varying SDLCs
  - Varying levels of popularity
  - Potentially lots of unknowns

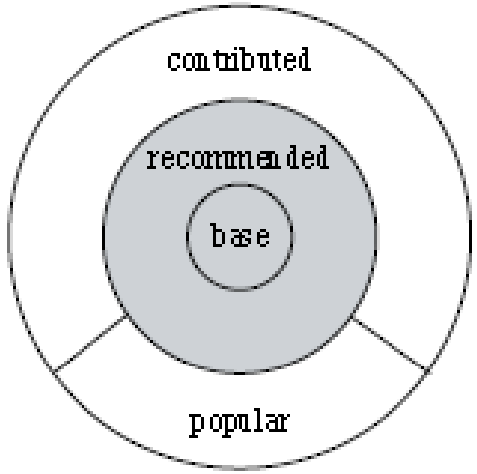


Image source: German, D.M. & Adams, Bram & Hassan, Ahmed E.. (2013). The Evolution of the R Software Ecosystem. Proceedings of the Euromicro Conference on Software Maintenance and Reengineering, CSMR. 243-252. 10.1109/CSMR.2013.33.

A RISK-BASED APPROACH FOR ASSESSING R PACKAGE ACCURACY WITHIN A VALIDATED INFRASTRUCTURE

Andy Nicholls, *Statistics Director, Head of Statistical Data Sciences, GSK*

Paulo R. Bargo, *Director Scientific Computing, Statistics & Decision Sciences, Janssen R&D*

John Sims, *Director, Analytical Systems Architect & Data Science - Pfizer Vaccine Research*

*On behalf of the **R Validation Hub**, an R Consortium-funded ISC Working Group*

January 23, 2020

Download the PDF version of this white paper here.

1. Scope and Background

This white paper addresses concerns raised by statisticians, statistical programmers, informatics teams, executive leadership, quality assurance teams and others within the pharmaceutical industry about the use of R and selected R packages as a primary tool for statistical analysis for regulatory submission work. When discussing validation of software systems two areas should be considered:

1. Infrastructure validation
2. Software validation

Infrastructure includes the server, OS, necessary infrastructure software, etc... For example, a system may use a server running Redhat Enterprise Linux (RHEL) version 6 and several

R Package Risk Assessment App

Package Control Panel

Select Package:

dplyr

Select Version:

1.0.5

Status: **Under Review**

Score: **0.21**

Leave Your Overall Comment:

current comment:

Submit Comment

Overall Risk:

Low High

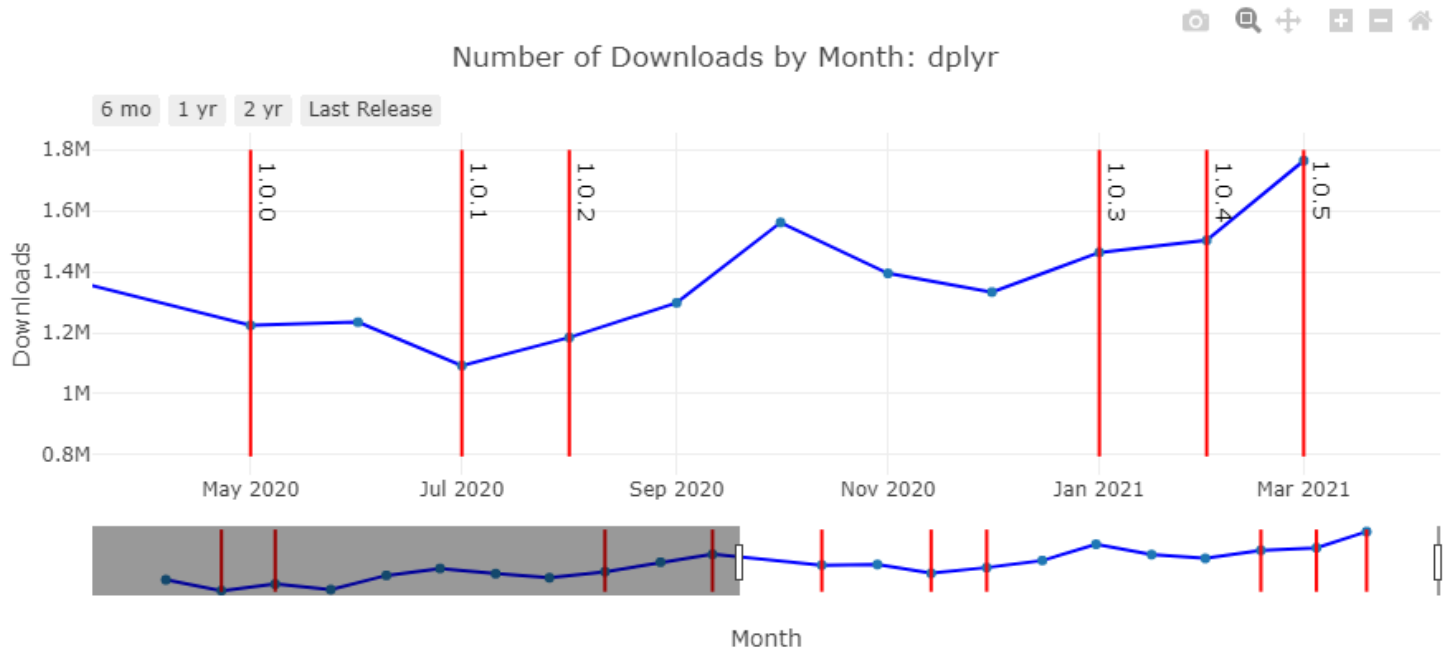
Low Medium High

Upload Package | Report Preview | Maintenance Metrics | Community Usage Metrics

PACKAGE MATURITY
88
Months since first release.

VERSION MATURITY
1
Months since version release.

DOWNLOAD COUNT
16,449,337
Downloads in Last Year



Leave Your Comment for Community Usage Metrics:

Commenting as Andy (Statistician)

Challenge 2: Responding to Risk

ICH: "*Appropriate* software testing procedures"

Challenge 2a: An Appropriate Comparison

vs

Summary

Definitely possible

*"The computer software used for data management and statistical analysis should be **reliable**, and documentation of **appropriate** software testing procedures should be available"*

ICH E9

More of a challenge!

Thank You

Acknowledgements

- R Validation Hub
- The PSI AIMS SIG
- Shutterstock

Further Reading

- **R**
  - R Validation Hub
  - R: Regulatory Compliance and Validation Issues A Guidance Document for the Use of R in Regulated Clinical Trial Environments
  - tidyverse, tidymodels, r-lib, and gt R packages: Regulatory Compliance and Validation Issues
- **ICH**
  - E9
- **FDA**
  - FDA Statistical Software Clarifying Statement
  - 21 CFR Part 11
  - Guidance for Industry Part 11, Electronic Records; Electronic Signatures — Scope and Application
  - Glossary of Computer System Software Development Terminology
  - General Principles of Software Validation; Final Guidance for Industry and FDA Staff
- **EMA**
  - Notice to sponsors on validation and qualification of computerised systems used in clinical trials
  - Q&A: Good clinical practice (GCP)