
Introduce R in Pharma
PhUSE Day Basel 2018-06-28
Reinhold Koch





is advancing in pharma



- New hires typically come with R know-how from university
- Statisticians love Rmarkdown or Sweave for reporting
- Management loves shiny apps
- Analysis techniques of new kind of data, -omics, biomarker, images, likely to be pioneered in R
- Promise of more efficient programming thanks to integration with modern software tools like source code management, automated testing, packaging

First steps: technical setup for

The R logo, a blue capital letter "R" inside a grey circle.

- *RStudio server pro*
 - with access to clinical study data
 - low entry barrier to try things out
- *Shiny server pro* (to be superseded by *RStudio Connect*)
 - Rmarkdown for web-based reporting, integrating text, figures, and code
 - shiny for interactive apps in the intranet
- Guidelines for installing R & RStudio locally

Build and leverage internal community

- Business value can immediately be generated by
 - visualizing, exploring data
 - checking data consistency
- Users are willing to help each other
 - good chance some R know-how already exists in the company
 - new hires
- Good collaboration facilities support R adoption

Develop community

- Local meetings
- Local presentations
- Local hands-on workshops
- Encourage self-learning
 - from books or internet
 - online courses
- Global electronic forum for
 - discussions
 - questions & answers

Develop collaboration

- **Source code management** with
 - documentation capabilities
 - issue tracking
- **Code packaging** with
 - automated testing
 - delivery to package server
(continuous integration / continuous deployment)

Develop reproducibility

- Source code management
 - tags
- Packaging
 - version numbers
 - integrated tests and documentation
 - static reports as vignettes
- Best: container, typically docker;
good stable base provided by <https://github.com/rocker-org/rocker-versioned/>
- Is this the future for **collaboration with FDA**?



Pharma & HA collaboration around containerized

- Reproducibility by containers is high
 - much better than any proprietary server installation
- Quality of container with R installation and R packages needs only be shown once
- Opportunities for **pre-competitive collaboration**
 - open source validation of a set of R installations
 - general sharing of reusable packages,
one example <https://github.com/roche/rtables>
 - maybe open source platform for electronic submissions?

Doing now what patients need next