Embrace R in Pharma -

Building an Internal R Community and Establishing Fit-for-Purpose Pilots

Ning Leng
Principal Statistical Scientist
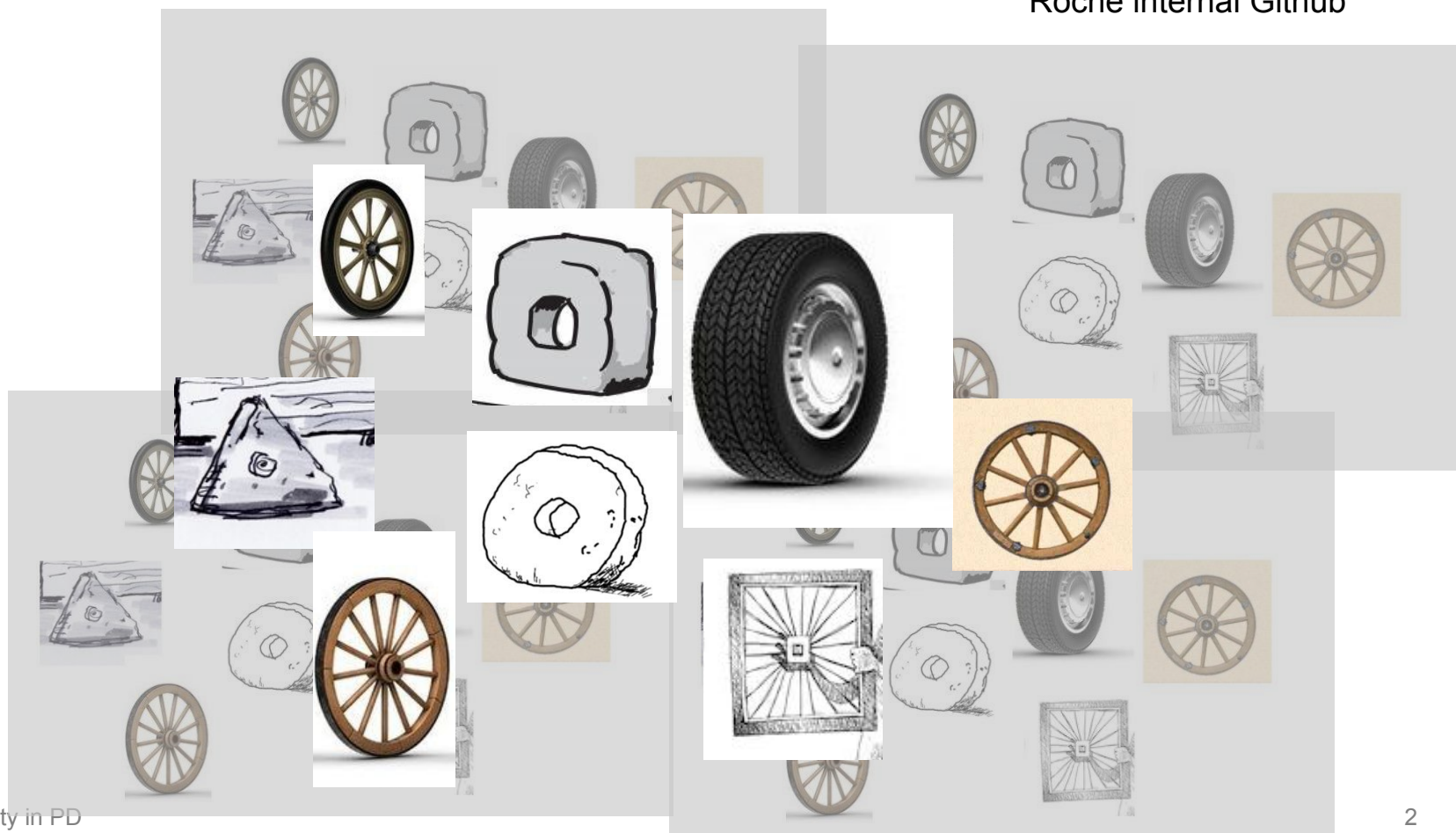*Roche-Genentech*
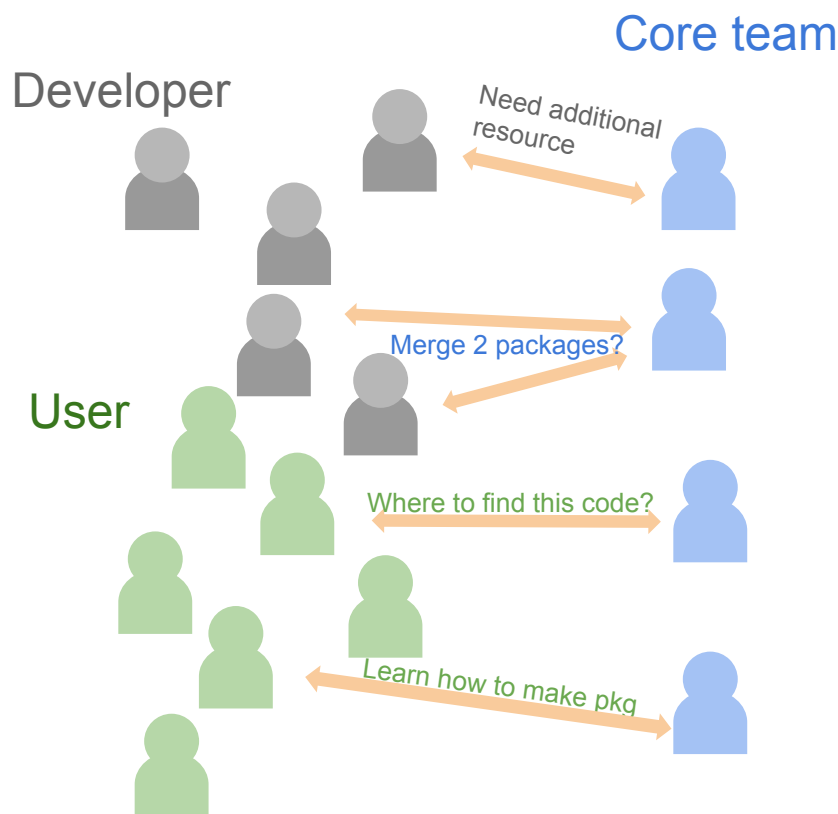
is becoming increasingly popular in pharma

Roche internal Github

Build an Internal R Community
to Connect the Dots & Improve Reusability

Core team

Developer

Need additional resource

Merge 2 packages?

User

Where to find this code?

Learn how to make pkg

Core team: volunteers representing different functions (statisticians, programmers, data scientists, data managers, etc.)

What We do

Connect R users/developers and advocate inner-source

R Community Website

Newsletters

Catalog existing tools

Electronic forum

Further R expertise

Trainings

R user share and learn

Good practice guidelines

Connect with external stakeholders

Roche/Genentech as the first Pharma company joining R Consortium

Improve R adoption

Fit-for-purpose R pilots

Example Pilot
Project Sautoir: Semi-AUTOmated Interpretation and Report

On behalf of Sautoir Team: Heng Wang, Shengchun Kong, Paul Paczuski, Dominik Heinzmann, Antoine Francois

The Need

- Biomarker analysis of >10 studies within one molecule (Post-Marketing Commitment)
  - Multiple outputs for each combination of:
    - o >10 studies
    - o 3 endpoints (OS, PFS, Response)
    - o 2 modeling methods
    - o >1 missing data imputation methods
  - Plus a comprehensive report

The Solution

- "One-click" workflow with automated report writing

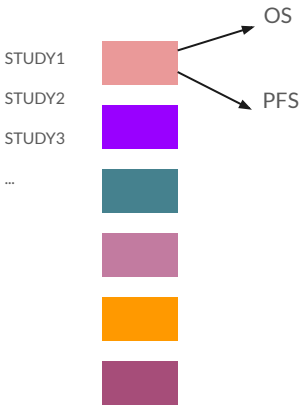# Project Sautoir: "One-click" Workflow Implementation

metadata

name	study	endpoint	...
STUDY1	GO12345	OS, PFSINV	...
STUDY2	GO12346	OS, PFSINV	...

Data

Sautoir::
auto_process()

STUDY1
STUDY2
STUDY3
...

→ OS

→ PFS

- Meta data driven pipeline
- Automated writing
- Github workflow/project board
- Regression test to ensure backward compatibility (to "ground truth")
- Make file & CI (ongoing)
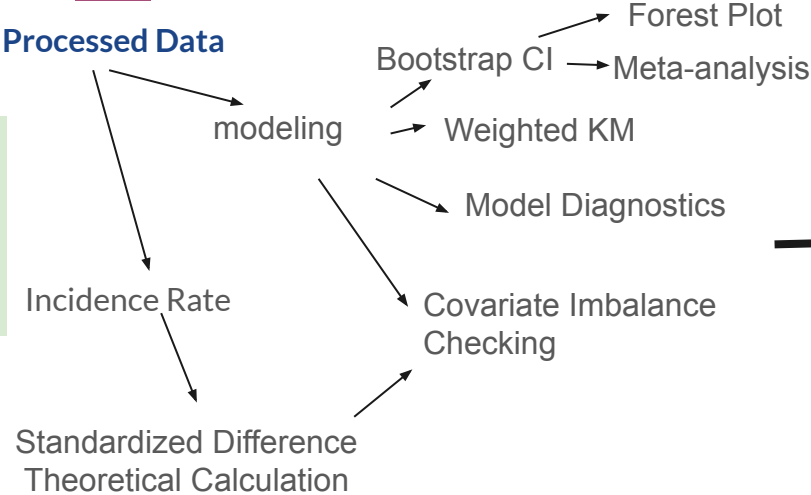
Developed 3 R packages
- PropensitySub: analysis modules
- gofaiWrites: auto-writing modules
- Sautoir: Workflow package

Processed Data

modeling → Bootstrap CI → Forest Plot

Bootstrap CI → Meta-analysis

→ Weighted KM

→ Model Diagnostics

Incidence Rate

Covariate Imbalance Checking

Standardized Difference Theoretical Calculation

Auto Writing

OS and PFS landmark analysis template
02 March, 2019

1 POPLAR

1.1 POPLAR OS

1.1.1 option1
1.1.1.1 POPLAR OS option1 : Sample Size

	Control	ADA-	ADA+
unadjusted	129	85	50
adjusted	122	81	48

1.1.1.2 POPLAR OS option1 : Variables in model
ECOGGR + SEX + TOBHX + METSITES + BALBUM + BLDH + BASLD + RACE1 + AGE65 + BNLR + BWT + HIST + TC1IC1 + LIVER + PRIORTX + BCRP

1.1.1.3 POPLAR OS option1 : Summary Statistics and CI

Sautoir::auto_analysis()
Backend: PropensitySub, tern

Sautoir::report.Rmd
Backend: gofaiWrites

Project Sautoir: Current Stage

- Submitted multiple reports to FDA
  - Auto-generated languages incorporated
  - Submitted R packages for esub

- R package esub challenges
  - All program files are required to be in .txt, file names need to be in lower case
  - fun.R -> fun.r.txt; DESCRIPTION -> description.txt

Roche

Doing now what patients need next