Reenvisioning **Clinical Content Delivery** in the **Open Source** World

Doug Kelkhoff
Sr Stats Programmer
Roche/Genentech

Genentech
A Member of the Roche Group

Who am I

- Eager to improve **reproducibility** and **transparency** in Pharma
- Vested in enabling best possible decisions for therapeutics
- Believe in the collective power of groups with shared vision
- Contributing to **containerized analysis** pilots with FDA (PhUSE Data Visualization Workstream*)
- Contributing toward the **R Validation Hub** project under the R Consortium (www.pharmar.org)

*PhUSE Best Practices for Interactive Analyses for Decision Making & Submissions Workstream

Genentech
A Member of the Roche Group

What I Do

1 Scientists discover promising theraputic

2 Viable manufacturing research

3 Clinical trial run

4 Analysis to prove efficacy, safety

Health Authority (HA) reviews analysis 5

Genentech
A Member of the Roche Group

What I Do

4 Analysis to prove efficacy, safety

- Extremely time-sensitive
- High throughput, predominantly heavily standardized analysis
- Emphasis on consistency, reliability
- Increasing requests for novel analysis
  (Most easily performed using open source tooling)
- Constant drive to expedite delivery of analysis to Health Authorities

What We Need

4 Analysis to prove efficacy, safety

- Extremely time-sensitive — ***speed***
- High throughput, predominantly heavily standardized analysis — ***reproducibility***
- Emphasis on consistency, reliability
- Increasing requests for novel analysis — ***flexibility***
  (Most easily performed using open source tooling)
- Constant drive to expedite delivery of analysis to Health Authorities — ***transparency***

Clinical Content Delivery

con·tent /ˈkäntent/ (*noun*)

(1) **figures** & **data** for informing theraputic decisions

(2) **interpretation** of trial results

(3) **documentation** of methods, computation

de·liv·er·y /dəˈliv(ə)rē/ (*noun*)

(1) supply of content **internally** and **externally**

Genentech
A Member of the Roche Group

Clinical Content Delivery

content delivery today

Genentech
A Member of the Roche Group

Content Delivery Today

collect trial data

clean data

analyze data

review analysis

bundle findings

deliver results

designed by Vexels

Content Delivery Today

collect trial data

clean data

analyze data

review analysis

bundle findings

deliver results

study data extracted, **copied to** cleaning location

analysts **emailed** about available study data

analyst **emails** scientists/ statisticians for review

writing group **notified of** final analysis

submission group **notified to** prepare eCTD

designed by Vexels

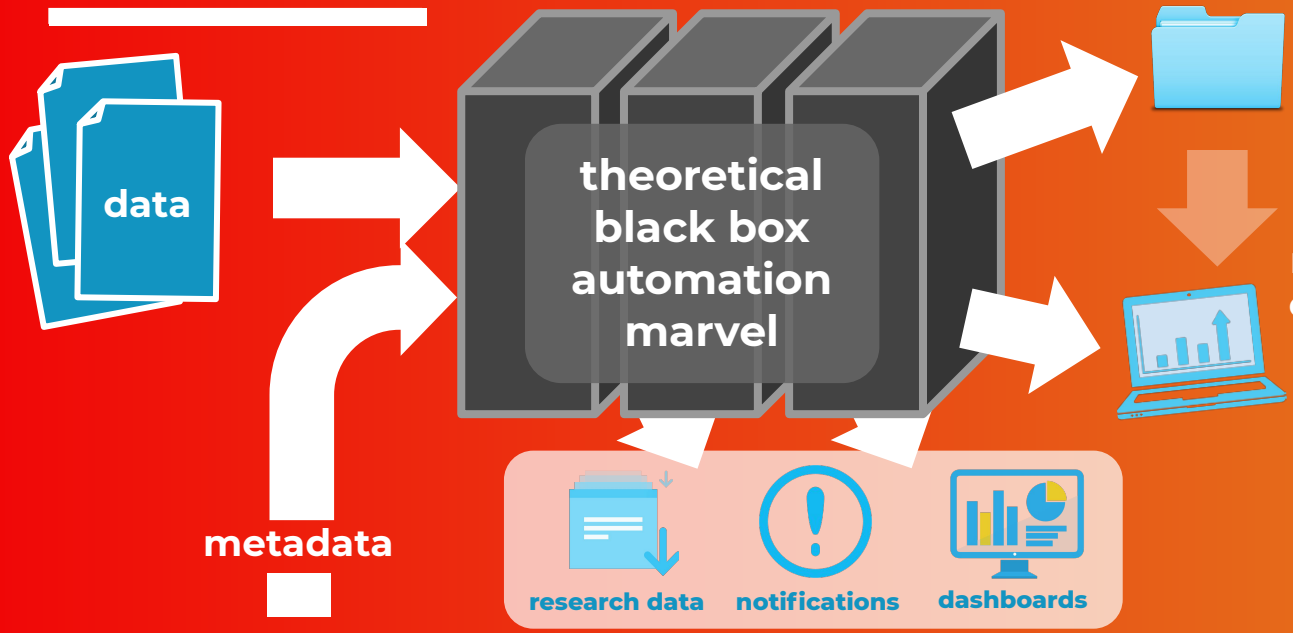
Genentech
A Member of the Roche Group

Clinical Content Delivery

data-first content delivery

Genentech
A Member of the Roche Group

Data-First Content Delivery

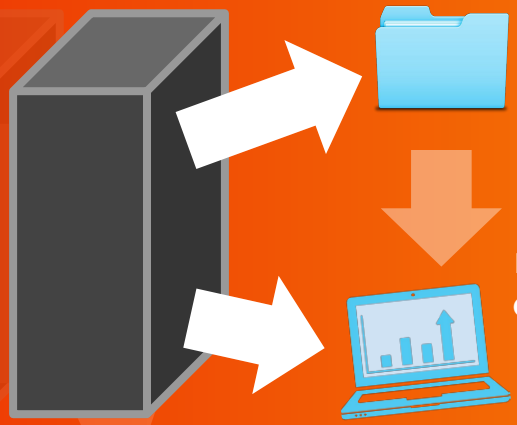


- standardized methods
- document instructions
- stats interpretations
- research questions

- patient status changes
- we have research-ready data
- our assumptions aren't met
- my results change

Harmonizing **Internal & Submitted** Analysis

- The environment we use, is the environment we ship
- The “final mile” as a continuous product
- Treat entirety of dataflow as ETL process
- **Analysis** + **Environment** + **Deliverables** as an ETL artifact



eCTD

reproducible environment

- analysis
- interactive exploration

Harmonizing **Internal & Submitted** Analysis

Requirements

serve analytic environments & views internally

transposable environment for HA delivery

data access easily recreated at HA

scalable content creation & iteration

Harmonizing **Internal & Submitted** Analysis

Containerized Analysis

- ✅ **serve** analytic environments & views internally
- ✅ **transposable** environment for HA delivery
- ❓ **data access** easily recreated at HA
- ❓ **scalable** content creation & iteration

Docker Pilot
for Interactive Analysis

Minimal Goal

- Deliver a Shiny app via docker container to FDA

Target

- Ship app with internal R package dependencies

Extended

- Socialize containers for reproducibility at HAs
- Explore restrictions of shipping R packages
- Ship analytic code to assist in review
- Ship a library of validated R packages

Docker Pilot
only possible with Open Source

- **Amazing prior art**
  rocker, shiny, public package repos
- **Installing R language & packages**
  wherever you want, as often as you want
- **Tools popular in the Open Source expect containers**
  ETL tooling, CI/CD, workflow managers, orchestration
- **Common Open Source foundation**
  makes it easier to collaborate across the industry

R VALIDATION HUB

Validated Analysis

The **riskmetric** package

> A community effort to standardize R package risk assessment for validated analysis

```
pkg_ref(c("riskmetric", "utils", "tools")) %>%
  as_tibble() %>%
  assess() %>%
  score() %>%
  mutate(risk = summarize_risk(.))
```

assess()

package	export_help
<chr>	<list<pkg_>
1 riskmet…	<lgl [15]>
2 utils	<lgl [219]>
3 tools	<lgl [119]>

score()

package	export_help
<chr>	<dbl>
1 riskmetr…	1
2 utils	0.995
3 tools	1

summarize_risk()

package	risk
<chr>	<dbl>
1 riskme…	0
2 utils	0.005
3 tools	0

pharmaR/riskmetric

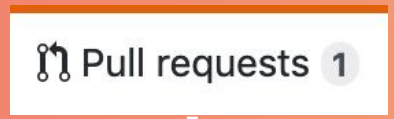
R VALIDATION HUB

Validated Analysis

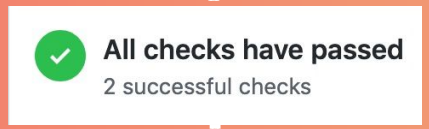
Validation Proof as a Continuous Product

R Package

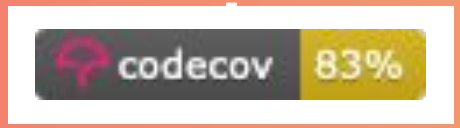
code altered



unit tests run



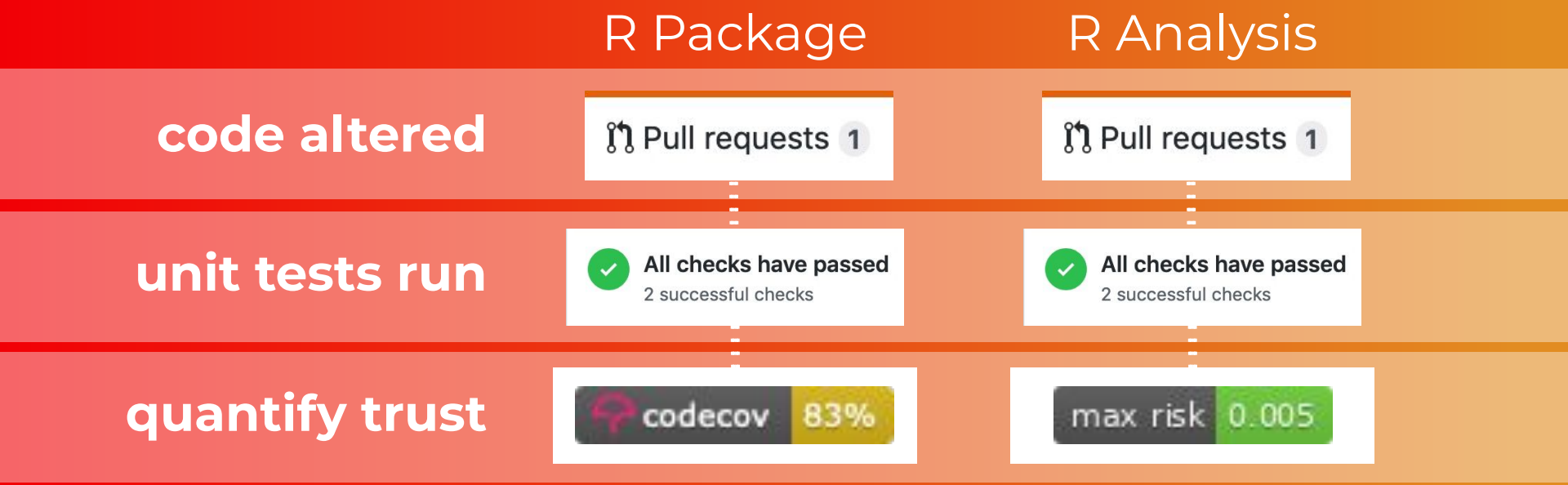
quantify trust



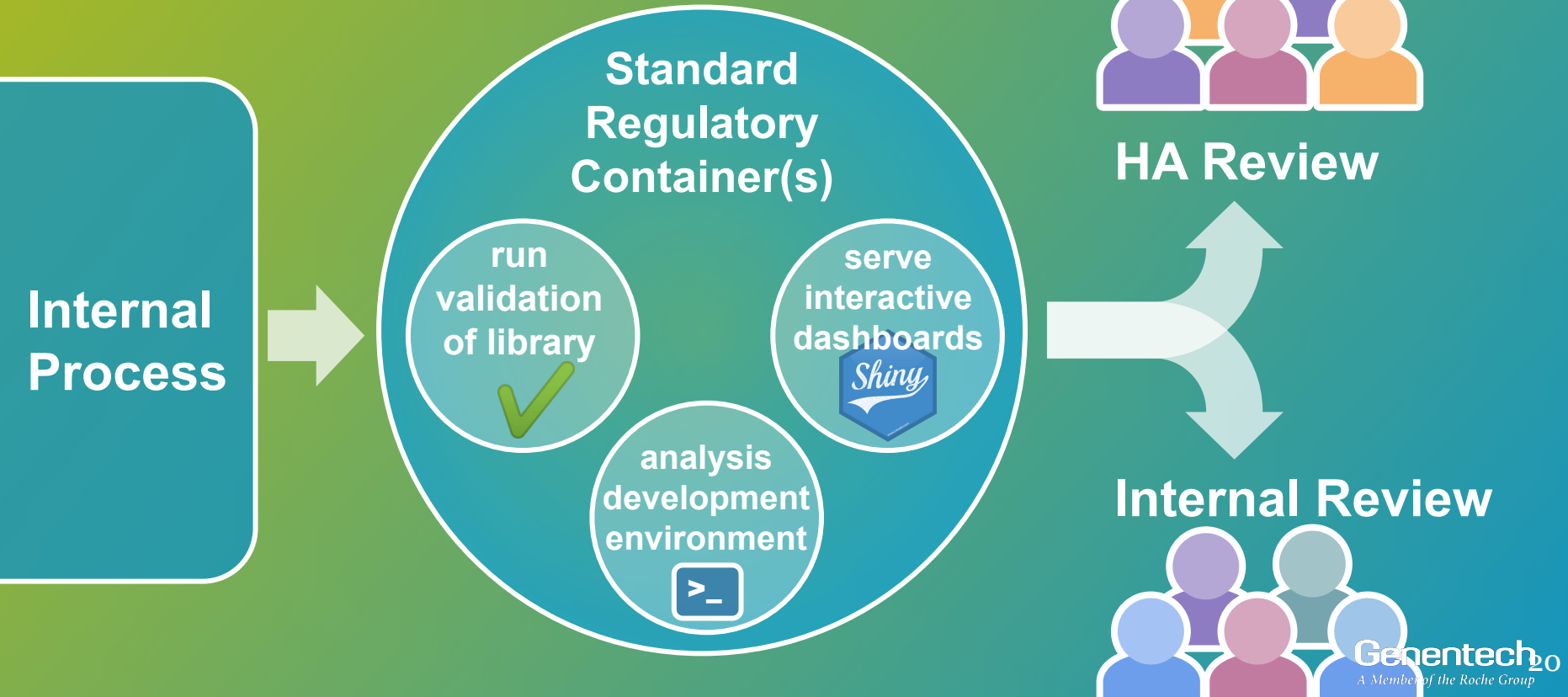
R VALIDATION HUB

Validated Analysis

Validation Proof as a Continuous Product



Putting it all together



A Community Effort

Benefits

- Working in the open source to drive **consistency**, **reproducibility**, **transparency**
- **Lower costs** of development by distributing effort
- **Lower risk** with cross-industry buy-in
- **Align** with practices outside pharma

A Community Effort

What you can do

- Reach out, provide feedback
- Get Involved
  - **R Validation Hub**  `pharmaR/riskmetric`
    - Review documentation, package design
    - Write metrics
  - **PhUSE Docker Pilot**
- Advocate within your organizations
- Share what you're doing

Thank You

Roche / Genentech
A Member of the Roche Group

phuse Data Visualization Workstream

R VALIDATION HUB

Containerization
Reinhold Koch
Adrian Waddell

Analysis as CI Artifact
Craig Gower

Docker Pilot
Xiangyun (Sharon) Wang
Nilesh Narayan
Paul Schuette
Alan Shapiro

`riskmetric`
Andy Nicholls
Yilong Zhang
Keven Kunzmann
Kieran Martin
Becca Krouse
Eric Nantz
Keaven Anderson
Matthias Trampisch