

The Validation of Open Source Tools for Clinical Statistics & Programming



Richard Pugh
Chief Data Scientist
Mango Solutions
rpugh@mango-solutions.com

Ali Dootson
Director, Business Consulting
d-wise
ali.dootson@d-wise.com

PHUSE Virtual EU Connect 2020

Agenda



Validating open source technologies



Case study: Validating R for clinical use



Proposing a Validated R-centric Environment

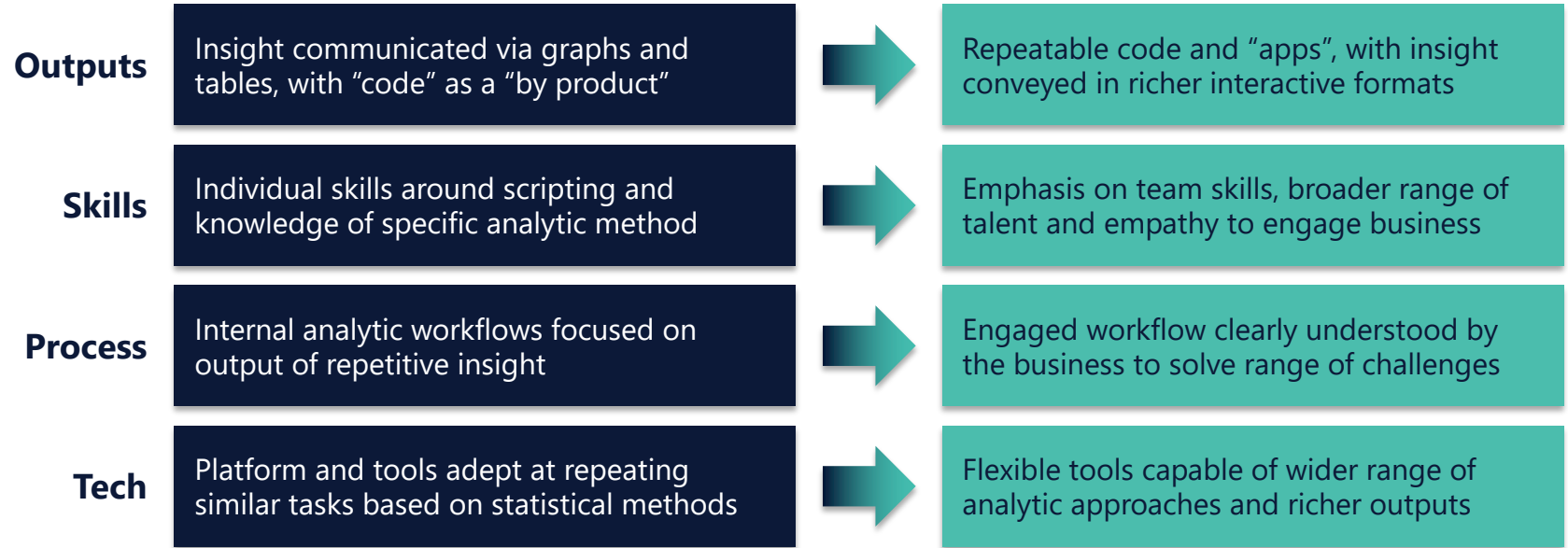


Questions?

Validating Open Source Tech

Data Science Hype and Technology Advances are raising expectations on Data & Analytics in pharma

This is having a significant impact on various aspects of our work



Open Source tools such as R, Python and Spark offer many advantages

However, pharma regulation means we need to work with validated tools in a clinical setting

What do we need to do for open source validation? Why is it non-trivial etc?

“Establishing **documented evidence** which provides a **high degree of assurance** that a specific process will **consistently produce** a product, meeting its **predetermined specifications** and **quality attributes**”

U.S. Food and Drug Administration (2013). 21 CFR Part 11: Electronic Records, Electronic Signatures,
<http://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfcfr/CFRSearch.cfm?CFRPart=11&showFR=1>

Case Study: Validating R for Clinical Use

Mango provides a “validated” build of R to pharma clients (called ValidR).

This section will walk through how we achieve this so you can replicate if desired

When validating R, we focus on providing control and validation



Control

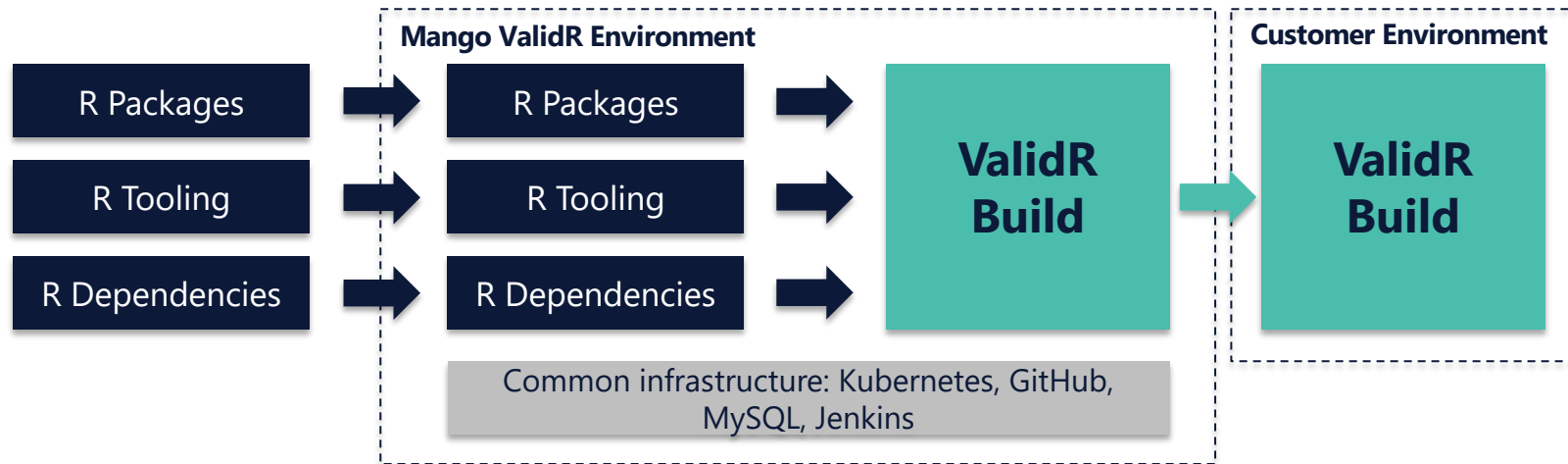
Ensuring we have control over the version of R and all of its dependencies



Validation

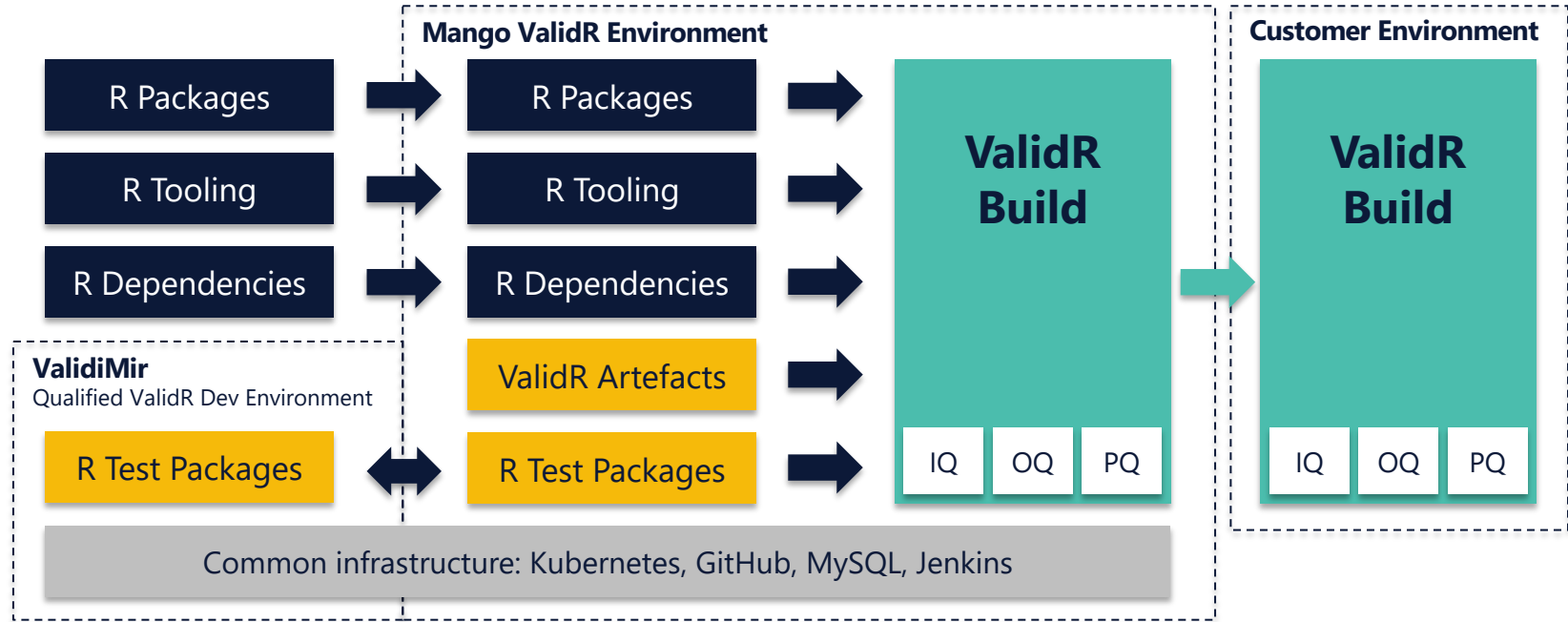
Generating the required documentation to assure an identified build is operating as expected

✖ ✖ We use DevOps approaches to ensure
➤ ➤ control over the R build





We add test packages and reporting capabilities to support IQ, OQ, PQ



We use a 7-step Process to validate packages to be included in a build

A critical (and time consuming) step in the process is the validation of R packages – these could be identified by clients or as part of our ongoing review

This is conducted by R experts working in a dedicated environment (Vladimir)

Individual package validation can range from 1-3 days depending on package complexity, size and maturity



Validating R is time consuming and complex ... but certainly possible!

github.com/mangothecat
mango-solutions.com/products/validr/

Proposing a Validated R-Centric Environment

d-wise provides “validated” platforms
to pharma clients

This section will walk through how we
achieve this so you can replicate if
desired

Building a validated platform



BUSINESS CHALLENGES

Clients need a centralized server for a SAS/R analytics environment and may not have IT resources to support it

May not have internal resources with experience in validating a SAS or R environment



PREFERRED SOLUTION

- Validated SAS environment
- Qualified pre-installation of R, Python and additional tools
- Version controlled storage for files
- User-centered folder structure
- Validated R packages as required

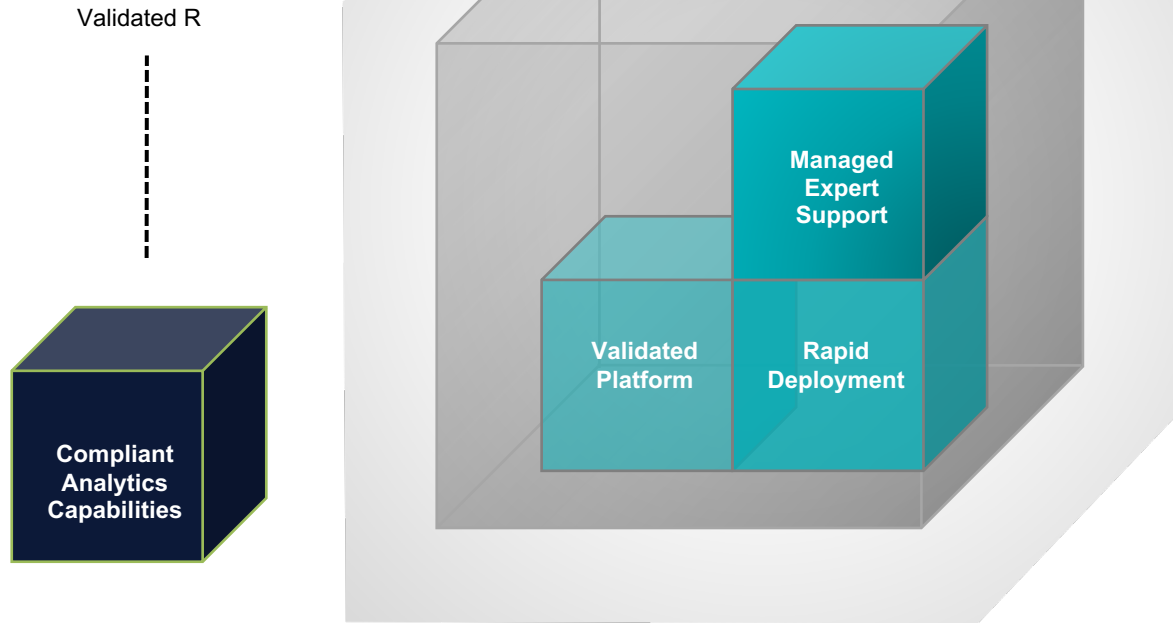


BUSINESS VALUE

Within a week, a solution can be implemented and fully operational that assists IT by providing validation service and user and system support.

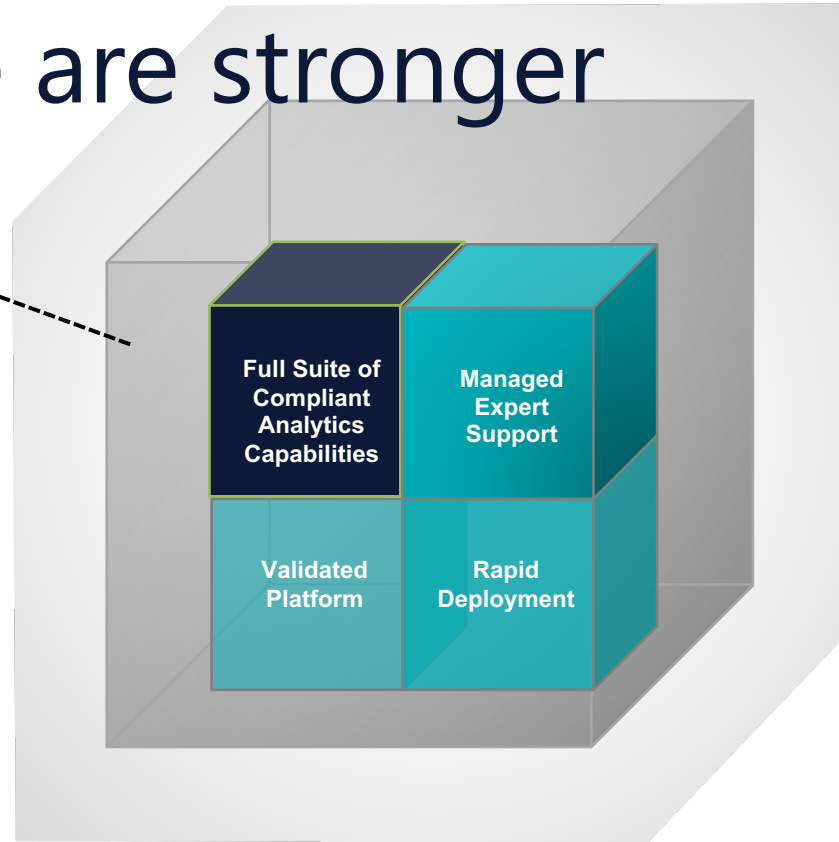
Removing system management hurdles enables clients to focus on analysis

Individually we are strong...



Together we are stronger

Validated R
Validated SAS
...



Get in touch if you have questions



Richard Pugh

Chief Data Scientist & co-Founder
Mango Solutions
rpugh@mango-solutions.com



Alistair Dootson

Director, Business Consulting
d-wise
ali.dootson@d-wise.com

The Validation of Open Source Tools for Clinical Statistics & Programming



Richard Pugh
Chief Data Scientist
Mango Solutions
rpugh@mango-solutions.com

Ali Dootson
Director, Business Consulting
d-wise
ali.dootson@d-wise.com

PHUSE Virtual EU Connect 2020