

OS06 Let's not get lost in translation: open-source adoption at Novartis

Can open-source meet the rigorous demands of regulatory science?

*“In **industry** the drug is a pill in a box, in **academia** the target discovery is already a drug, and all the rest is just technicalities”*

Open-source can meet the rigorous demands of regulatory science





EUROPEAN MEDICINES AGENCY
SCIENCE MEDICINES HEALTH

17 February 2020
EMA/CHMP/ICH/436221/2017
Committee for Medicinal Products for Human Use

ICH E9 (R1) addendum on estimands and sensitivity analysis in clinical trials to the guideline on statistical principles for clinical trials
Step 5

Transmission to CHMP	July 2017
Adoption by CHMP for release for consultation	20 July 2017
Start of consultation	31 August 2017
End of consultation (deadline for comments)	28 February 2018
Final adoption by CHMP	30 January 2020
Date for coming into effect	30 July 2020



Why open source?



Scientific & regulatory advances

1998

INTERNATIONAL CONFERENCE ON HARMONISATION OF TECHNICAL
REQUIREMENTS FOR REGISTRATION OF PHARMACEUTICALS FOR HUMAN
USE

ICH HARMONISED TRIPARTITE GUIDELINE

STATISTICAL PRINCIPLES FOR CLINICAL TRIALS
E9

2015



EUROPEAN MEDICINES AGENCY
SCIENCE MEDICINES HEALTH

26 February 2015
EMA/CHMP/295050/2013
Committee for Medicinal Products for Human Use (CHMP)

Guideline on adjustment for baseline covariates in clinical
trials

2023

Adjusting for
Covariates in
Randomized Clinical
Trials for Drugs and
Biological Products
Guidance for Industry

2020



EUROPEAN MEDICINES AGENCY
SCIENCE MEDICINES HEALTH

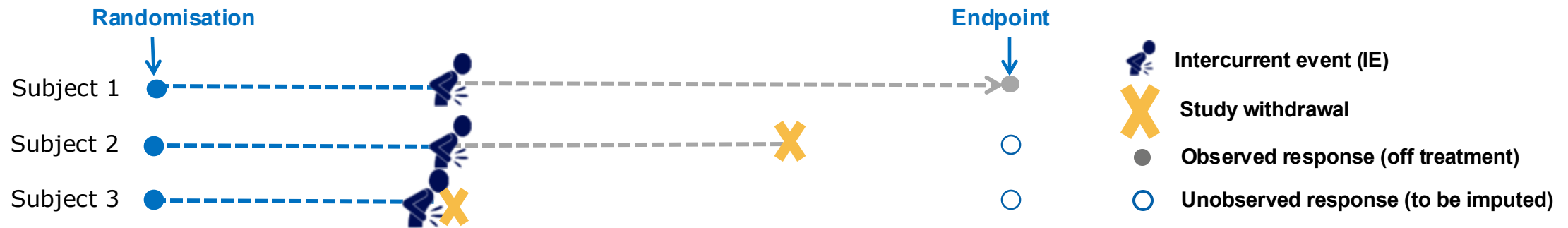
17 February 2020
EMA/CHMP/ICH/436221/2017
Committee for Medicinal Products for Human Use

ICH E9 (R1) addendum on estimands and sensitivity
analysis in clinical trials to the guideline on statistical
principles for clinical trials

These updates are changing how we analyse and report clinical trials



Adequate handling of missing data essential for estimation of treatment policy estimand



- Standard MMRM models often do not adequately handle missing data
- Common models for imputing missing data in treatment policy setting
 - Reference-based multiple imputation (RBMI), e.g., jump-to-reference, copy reference, etc.
 - “Retrieved Dropout”

Only open-source software available for reference-based and RDO imputation

Journal of Biopharmaceutical Statistics, 23: 1352–1371, 2013
Copyright © Taylor & Francis Group, LLC
ISSN: 1054-3406 print/1520-5711 online
DOI: 10.1080/10543406.2013.834911



Methodology goes back to Carpenter et al.

- “We have implemented our proposal in a generic SAS macro, which, together with instructions, is freely available from www.missingdata.org.uk.”
- Website now defunct

Method not implemented in commercial software package, e.g., SAS BASE

- Various open-source solution exists
 - SAS: %5macros, %RMConjPlus, %MISTEP
 - R: {rbmi}

ANALYSIS OF LONGITUDINAL TRIALS WITH PROTOCOL DEVIATION: A FRAMEWORK FOR RELEVANT, ACCESSIBLE ASSUMPTIONS, AND INFERENCE VIA MULTIPLE IMPUTATION

James R. Carpenter, James H. Roger, Michael G. Kenward
Medical Statistics Department, London School of Hygiene & Tropical Medicine, London, UK

Protocol deviations, for example, due to early withdrawal and noncompliance, are unavoidable in clinical trials. Such deviations often result in missing data. Additional assumptions are then needed for the analysis, and these cannot be definitively verified from the data at hand. Thus, as recognized by recent regulatory guidelines and reports, clarity about these assumptions and their implications is vital for both the primary analysis and framing relevant sensitivity analysis. This article focuses on clinical trials with longitudinal quantitative outcome data. For the target population, we define two estimands, the de jure estimand, “does the treatment work under the best case scenario,” and the de facto estimand, “what would be the effect seen in practice.” We then carefully define the concept of a deviation from the protocol relevant to the estimand, or for short a deviation. Each patient’s postrandomization data can then be divided into predeviation data and postdeviation data. We set out an accessible framework for contextually appropriate assumptions relevant to de facto and de jure estimands, that is, assumptions about the joint distribution of pre- and postdeviation data relevant to the clinical question at hand. We then show how, under these assumptions, multiple imputation provides a practical approach to estimation and inference. We illustrate with data from a longitudinal clinical trial in patients with chronic asthma.

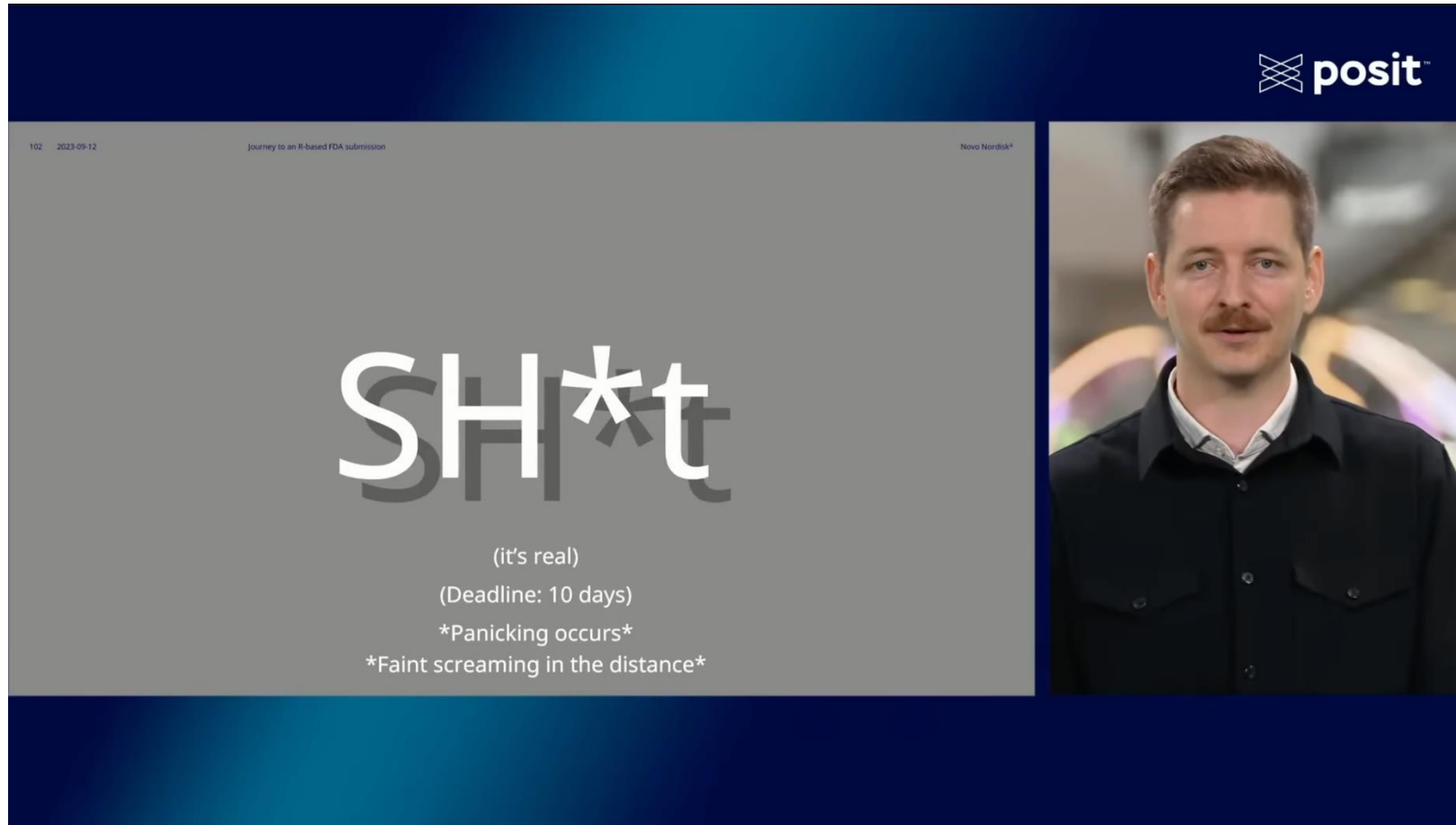
{rbmi} recommended after internal review: based on software engineering best practice

	rbmi	%RMConjPlus
Design principles		
Version control (i.e., Git, SVN)		
Issue tracking		
Consistent and readable code style		
Documentation		
Unit and integration tests		
Publication on website and repositories		

Adopting Open-Source challenges traditional ways of working

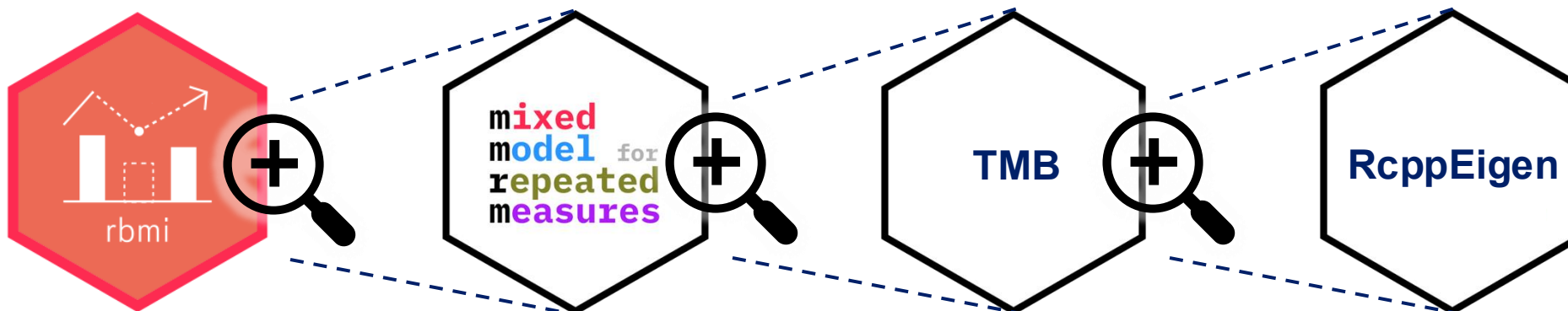
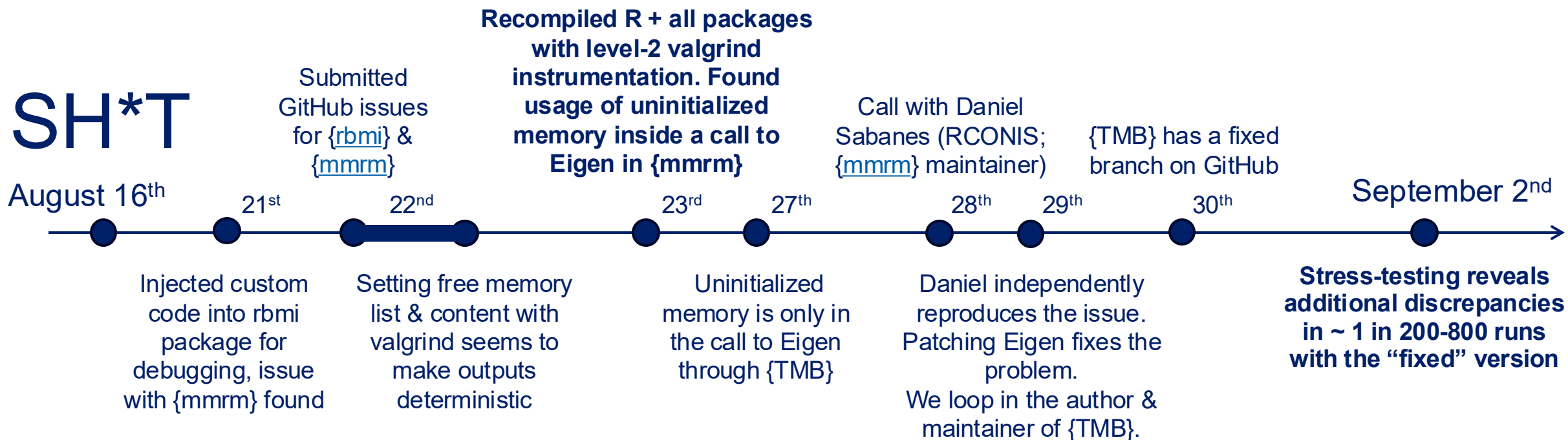


When the rubber hits the road



Collaborating across groups to find the root cause

SH*T



Transparency fostered trust

Confidence was restored through **fast, open & transparent* collaboration** with the developers and maintainers:
→ From the identification of bug to its fix, it took only two weeks!

Changes between {TMB} & {mrmr} releases **documented** on GitHub / CRAN:

- <https://github.com/kaskr/adcomp/compare/v1.9.14...v1.9.15>
- <https://cran.r-project.org/web/packages/TMB/NEWS>
- <https://github.com/openpharma/mrmr/compare/v0.3.12...v0.3.14>
- <https://cran.r-project.org/web/packages/mrmr/news/news.html>

The problem is now fixed for **everyone!**

See the RCONIS Blog: <https://www.rconis.com/blog/mrmr-tmb-debug/>

* {rbmi} issue: <https://github.com/insightsengineering/rbmi/issues/420>

{mrmr} issue: <https://github.com/openpharma/mrmr/issues/462>

{TMB} issue: <https://github.com/kaskr/adcomp/issues/398>



**The pilot also lead
to further
innovations**



{beeca}: Binary Endpoint Estimation with Covariate Adjustment

```
library(beeca)
```

```
## Set treatment to a factor
trial01$trtp <- factor(trial01$trtp)
```

```
## Fit a logistic regression working model and pass it to beeca
fit1 <- glm(aval ~ trtp + bl_cov, family="binomial", data=trial01) |>
  get_marginal_effect(trt="trtp", method="Ye", contrast="diff")
```

```
## View the results in Analysis Results Data (ARD) structure
fit1$marginal_results
```

- Structured ARD output supports clinical reporting
- Facilitates results re-use and reproducibility



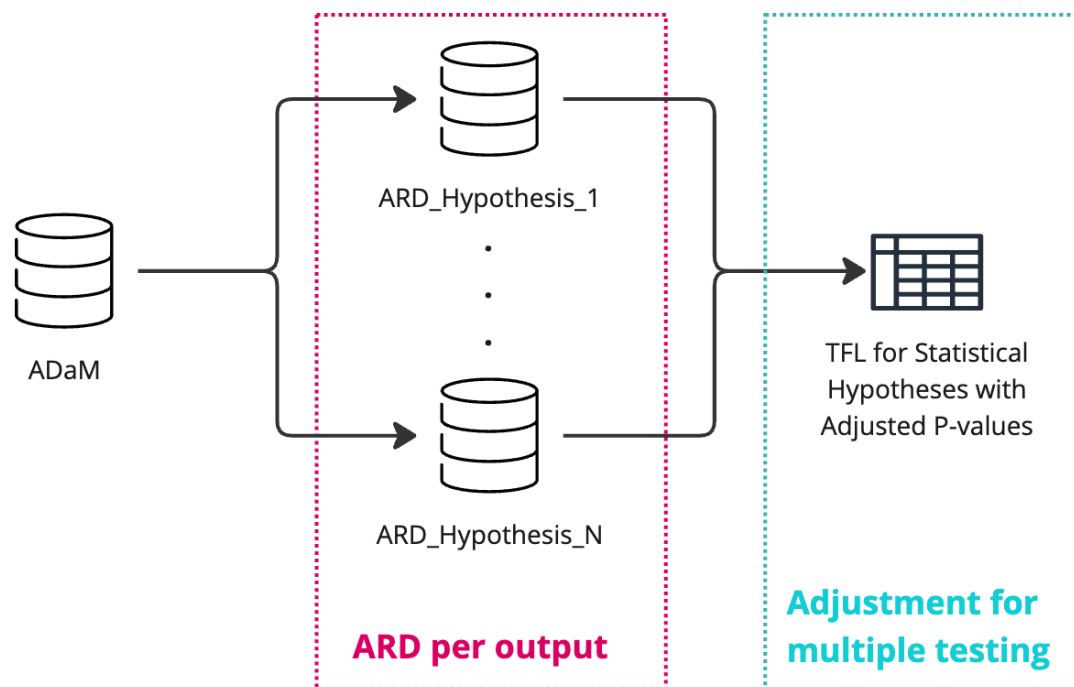
```
## # A tibble: 19 x 8
##   TRTVAR TRTVAL      PARAM ANALTP1 STAT  STATVAL ANALMETH ANALDESC
##   <chr>  <chr>      <chr>  <chr>    <chr>    <dbl>  <chr>    <chr>
## 1 TRTP   Placebo     AVAL  DESCRIP... N      86      count  Compute...
## 2 TRTP   Placebo     AVAL  DESCRIP... n      29      count  Compute...
## 3 TRTP   Placebo     AVAL  DESCRIP... %     33.7    percent... Compute...
## 4 TRTP   Placebo     AVAL  INFEREN... risk    0.343    g-compu... Compute...
## 5 TRTP   Placebo     AVAL  INFEREN... risk... 0.0518   Ge - HC0  Compute...
## 6 TRTP   Xanomeline High Dose  AVAL  DESCRIP... N      84      count  Compute...
## 7 TRTP   Xanomeline High Dose  AVAL  DESCRIP... n      61      count  Compute...
## 8 TRTP   Xanomeline High Dose  AVAL  DESCRIP... %     72.6    percent... Compute...
## 9 TRTP   Xanomeline High Dose  AVAL  INFEREN... risk    0.716    g-compu... Compute...
## 10 TRTP  Xanomeline High Dose  AVAL  INFEREN... risk... 0.0494   Ge - HC0  Compute...
## 11 TRTP  Xanomeline Low Dose  AVAL  DESCRIP... N      84      count  Compute...
## 12 TRTP  Xanomeline Low Dose  AVAL  DESCRIP... n      62      count  Compute...
## 13 TRTP  Xanomeline Low Dose  AVAL  DESCRIP... %     73.8    percent... Compute...
## 14 TRTP  Xanomeline Low Dose  AVAL  INFEREN... risk    0.743    g-compu... Compute...
## 15 TRTP  Xanomeline Low Dose  AVAL  INFEREN... risk... 0.0460   Ge - HC0  Compute...
## 16 TRTP  diff: Xanomeline High ... AVAL  INFEREN... diff    0.374    g-compu... Compute...
## 17 TRTP  diff: Xanomeline High ... AVAL  INFEREN... diff... 0.0717   Ge - HC0  Compute...
## 18 TRTP  diff: Xanomeline Low D... AVAL  INFEREN... diff    0.400    g-compu... Compute...
## 19 TRTP  diff: Xanomeline Low D... AVAL  INFEREN... diff... 0.0694   Ge - HC0  Compute...
```

We actively exploring ARDs in our studies



1. Study Pilot: learnings

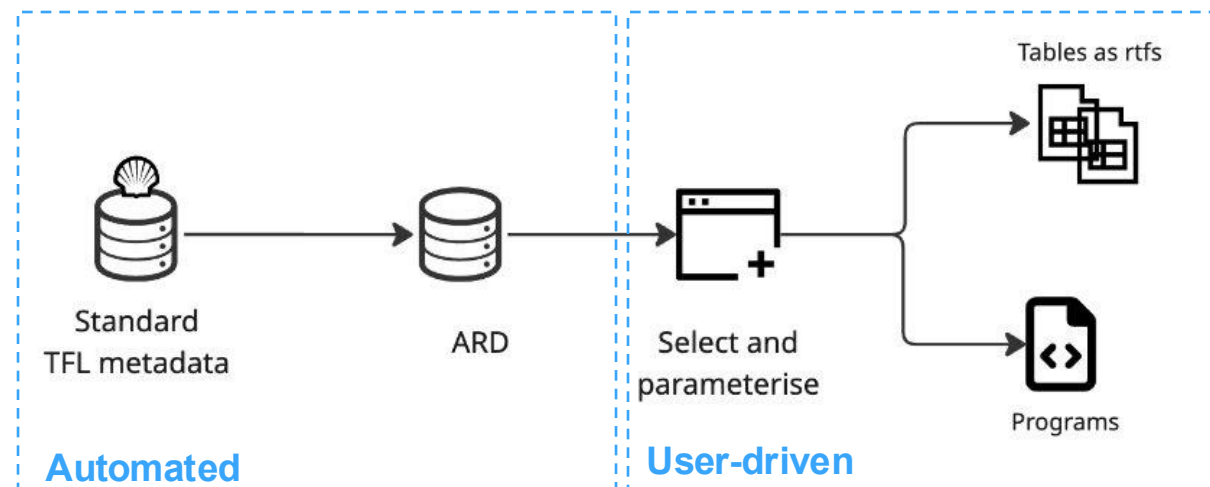
Value: management of a large set of results prior for downstream use in a complex hierarchical testing strategy



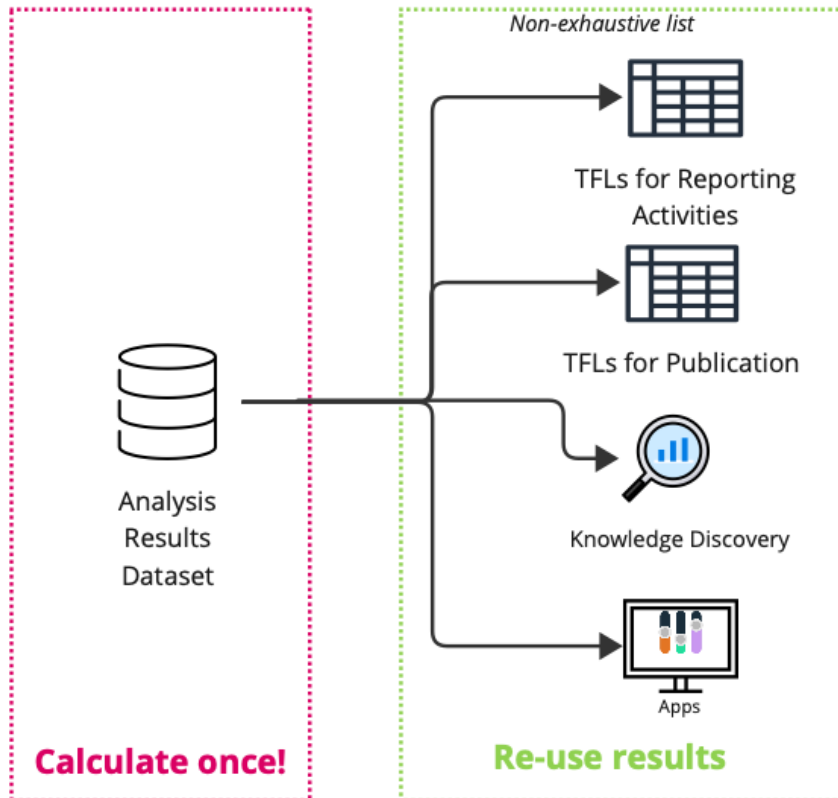
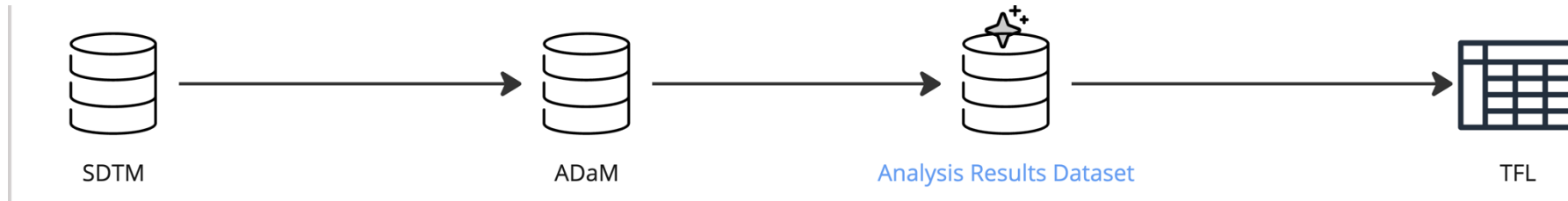
2. Refine and Scale: mosaic

Value: automate our **standard outputs** through domain-level ARDs.

Mosaic will eliminate double programming for these outputs.



Mosaic: From results as TFLs to results as datasets



Primary Objectives

-  **Automate** the generation and QC of our standard TFLs
-  **Standardise** the Analysis Results Dataset (same format for all outputs)
-  **Rapid iteration** for rendering TFLs
-  **Validated, standardized, machine readable results** for further use (e.g. apps, genAI)

Conclusions and key messages



Open-source can meet the rigorous demands of regulatory science





EUROPEAN MEDICINES AGENCY
SCIENCE MEDICINES HEALTH

17 February 2020
EMA/CHMP/ICH/436221/2017
Committee for Medicinal Products for Human Use

ICH E9 (R1) addendum on estimands and sensitivity analysis in clinical trials to the guideline on statistical principles for clinical trials

Step 5

Transmission to CHMP	July 2017
Adoption by CHMP for release for consultation	20 July 2017
Start of consultation	31 August 2017
End of consultation (deadline for comments)	28 February 2018
Final adoption by CHMP	30 January 2020
Date for coming into effect	30 July 2020



Thank you

Contact: mark.baillie@novartis.com

