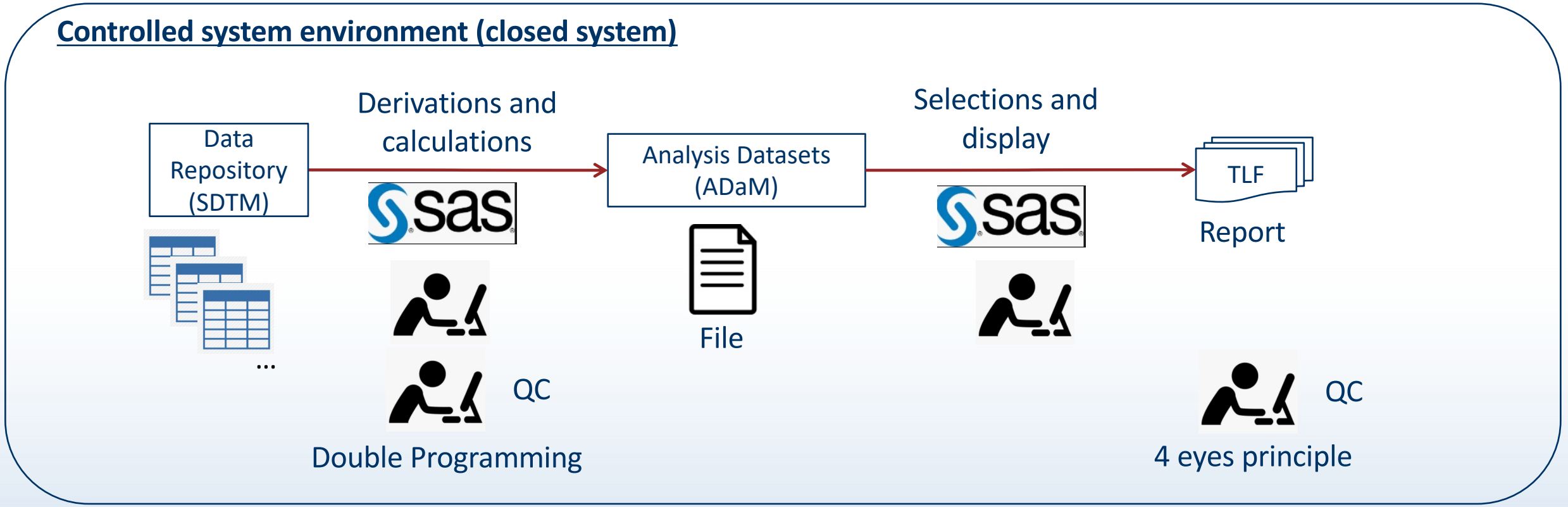
Implementation of a highly automated dataflow connected to a visualization tool to achieve an efficient interaction with scientific interfaces

PHUSE Spring SDE, 04MAY2021, Igor Varvodic, Magnus Nastoll

Boehringer Ingelheim

Clinical Data Analysis in a regulated environment



- The procedure is in place for the final reporting of Clinical Data
- The procedure has been applied to Clinical Data Analysis during trial conduct

Understanding the current clinical data concept @ BI

- **No Database:**
  Clinical data is stored in SDTM format as SAS datafiles within various folders and subfolders on a file server.

- **Processing of data needed:**
  Before the data can be analysed, the datafiles have to be combined, transformed and new columns/values derived.

- **Additional sources:**
  In addition to the SDTM data files, other sources need to be considered
  (e.g. statistical analysis datasets).

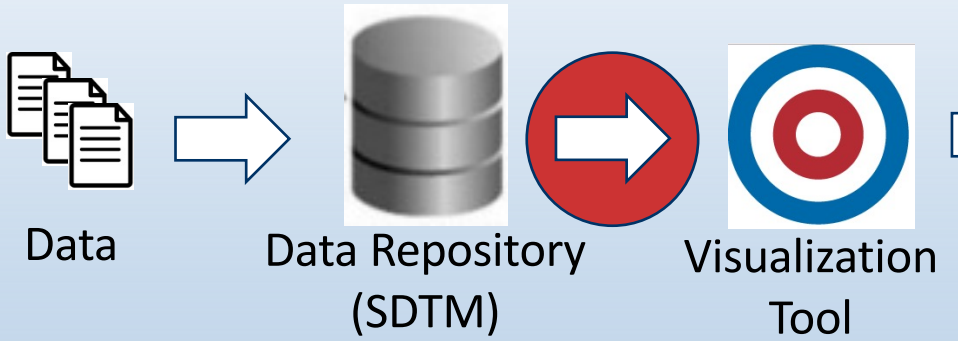
- **Data sharing:**
  Before data is shared, data protection and data security need to be ensured.

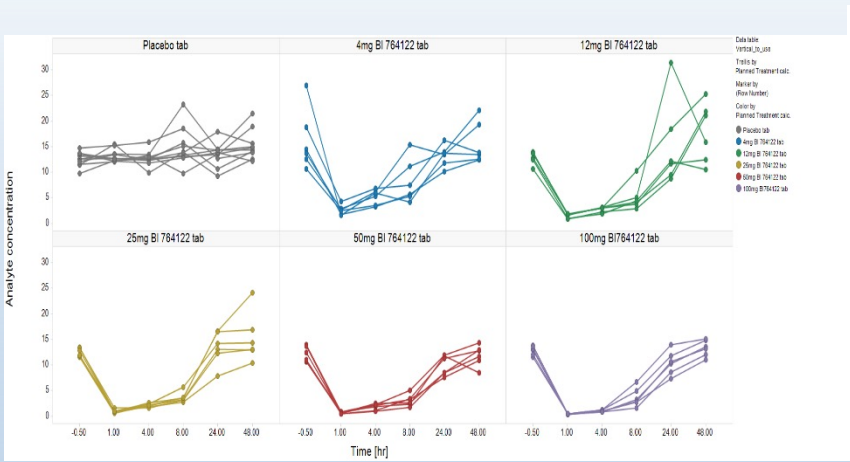
Challenge: Regulatory vs. exploratory process

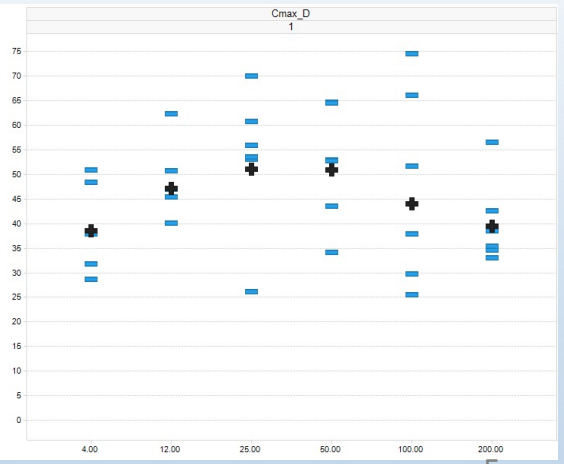
Regulatory analysis process	**Exploratory** analysis process
...aims for a static clinical report **at the end** of a trial	...aims for a flexible, interactive, and visual analysis at **any time during trial conduct**
...requires many interfaces & manual steering steps	...directly enables the customer to access clinical data
...is designed for the documentation of results	...allows agile guidance of trials and projects

VISTA – Visual & Integrated Support of Translational Analysis

- Provide **one** automated procedure to access, visualize & analyze clinical data…
  … interactively & flexible… for any open label trial… at any point in time
- Support trial teams with interactive trial analysis on PK & biomarkers (including covariates)
- Apply CDISC compliant standard data structure
- Use powerful but user friendly analysis technologies (Tibco Spotfire ®)







Components of the process

User interface for trial specific setup:
User friendly, machine readable specification files, to be filled by scientific function.

Vertical datafile structure:
Vertical datafile structure allowing to combine and pool the data needed in one table.

SAS programs:
Automatically running, robust and flexibel SAS routines which are controlled by user defined parameters.

Data Security:
Permission to data is limited to users allowed to work with the data during trial conduct.

Database:
Simple and fast database connection instead of processing separate files to speed up the processes and to reduce manual interactions (e.g file exchange for new versions)

Automated VISTA – Dataflow

Source data:
Available as SAS datasets in SDTM format on clinical data repository

Data harmonization and creation:
SAS environment to harmonize and cure data for addition to database

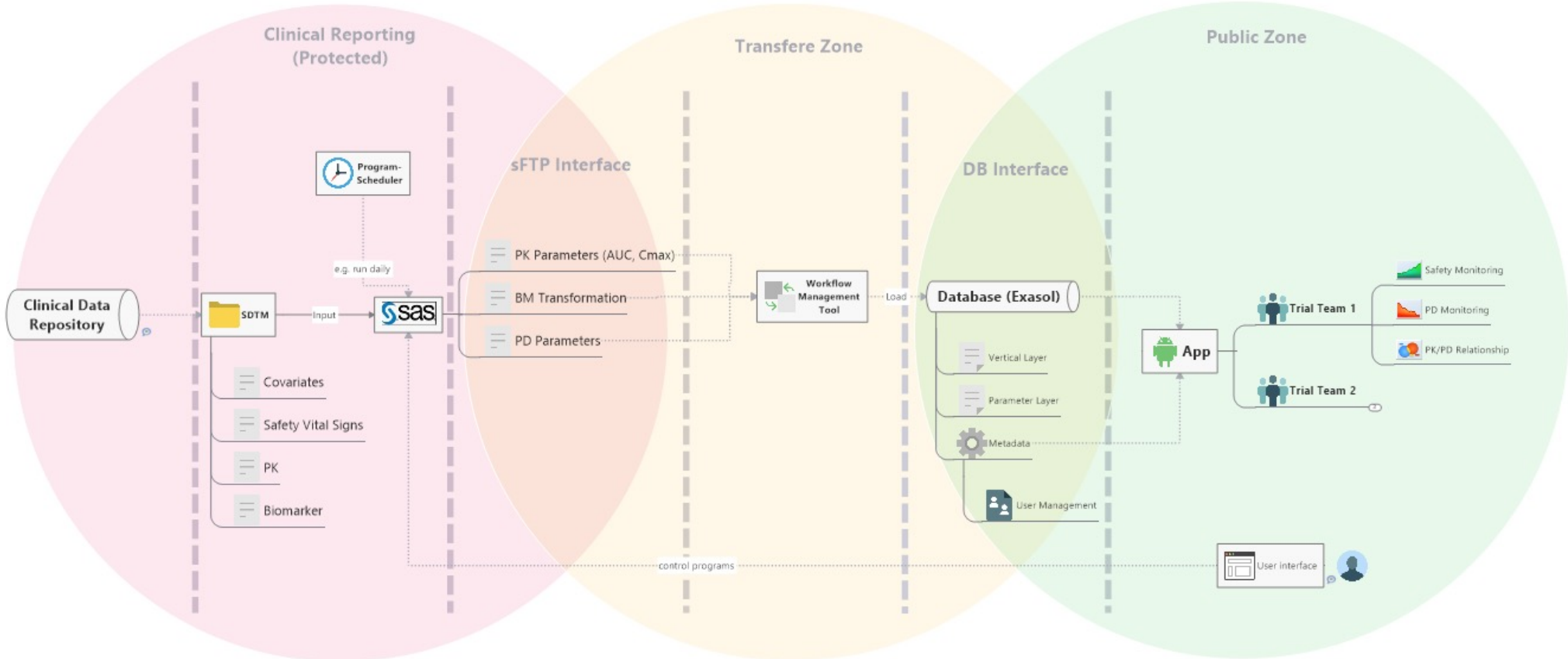
Scheduler:
Used for automated run of SAS programs at a defined regular timepoint (e.g. daily)

Workflow management tool:
New or updated data get loaded and replaced automatically in database

Database:
All trial data will be stored in a harmonized structure, and can be used for analysis and visualization

Visualization:
The database is connected to Spotfire and users are able accessing data according to their permission

VISTA Dataflow



Clinical Environment

...preparing datafiles which should be included in the Database.

sFTP Folder

...new or updated files are stored on a sFTP folder.

Data Conditioning & Processing

...detect new or updated files.

...identified by the filename -> study number, table name (e.g. 1234-5678_vertical.csv).

...update database tables

Database

...store datafiles as vertical and parameter layer

...store metadata e.g. data for user management/access control

Web based Apps

..connect Apps to Database

...appropriate script (R) to compare datafiles on sFTP against files used to validate the transfer.

Demonstration of VISTA Workflow & Spotfire