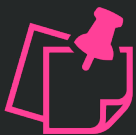


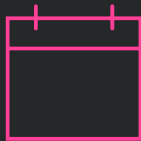
Beyond P21 Data Checks

ABDULKADIR LOKHANDVALA
SENIOR PRINCIPAL PROGRAMMER, ICON
PHUSE EU Connect
15-Nov-2021

Agenda



- Overview
- Components
- Workflow
- Overview of Rule-book and Macro Framework
- Output Reports
- Merits and Tradeoffs
- Future Enhancement
- Q&A



- Pinnacle21 is considered industry standard tool to conform/validate CDISC data, but what if there is a need to perform additional Study/Company Specific checks?
- To achieve this in efficient manner, *Beyond P21 Data Checks* has been developed
- It is an infrastructure level Excel- and SAS[®]- based tool allowing user to define their own custom checks in form of Macro or Open SAS[®] snippets.
- It executes codes specified in excel and generates final reports in an excel format
- The presentation focuses different elements of tool and real-life example
- It does **not** encourage to eliminate use of Pinnacle21

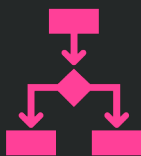


Beyond P²¹
Checks



Better Quality Data

Components



➤ ***Rule-book.xlsm***: Used for adding/defining Study Specific Checks



➤ ***Global-rules.xlsm***: A globally accessible workbook which holds list of Global Checks

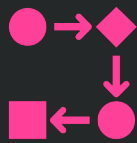


➤ ***NP21CHK.sas***: An Engine macro program that reads both the excels and generates reports



➤ ***Global Macro Catalog***: A globally accessible macro catalog to store Global Checks

Workflow

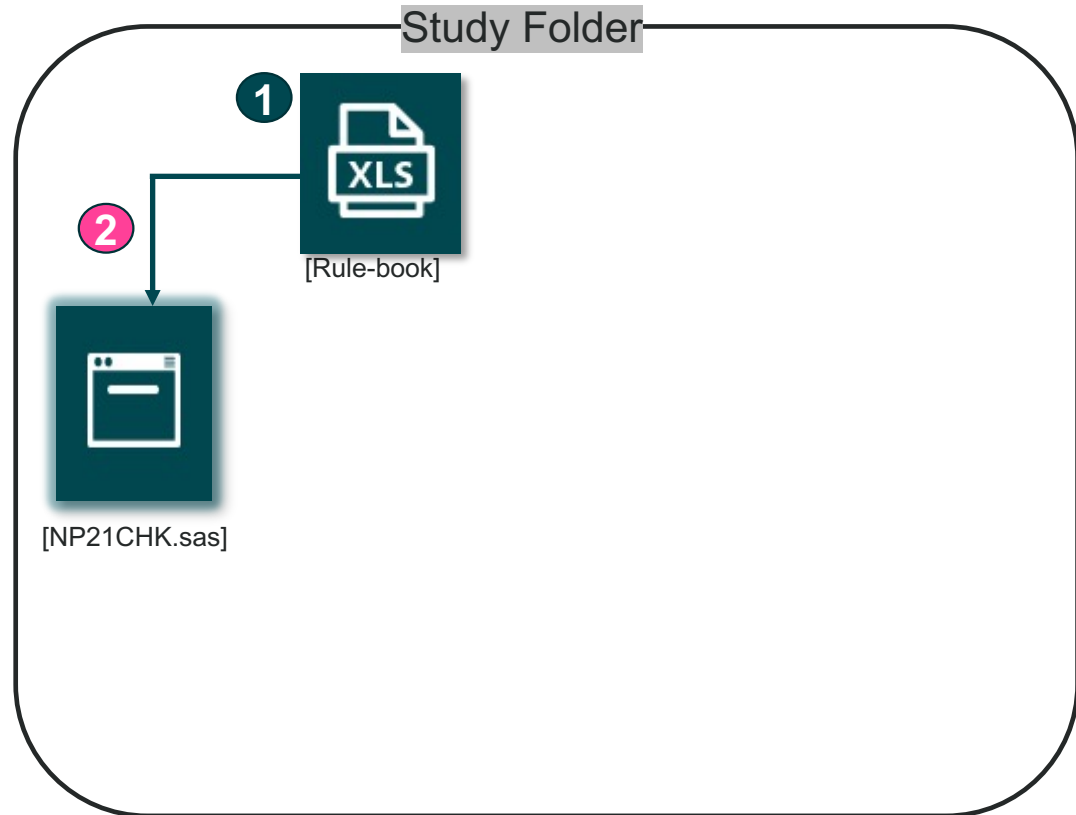
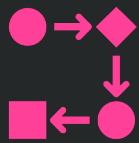


Study Folder

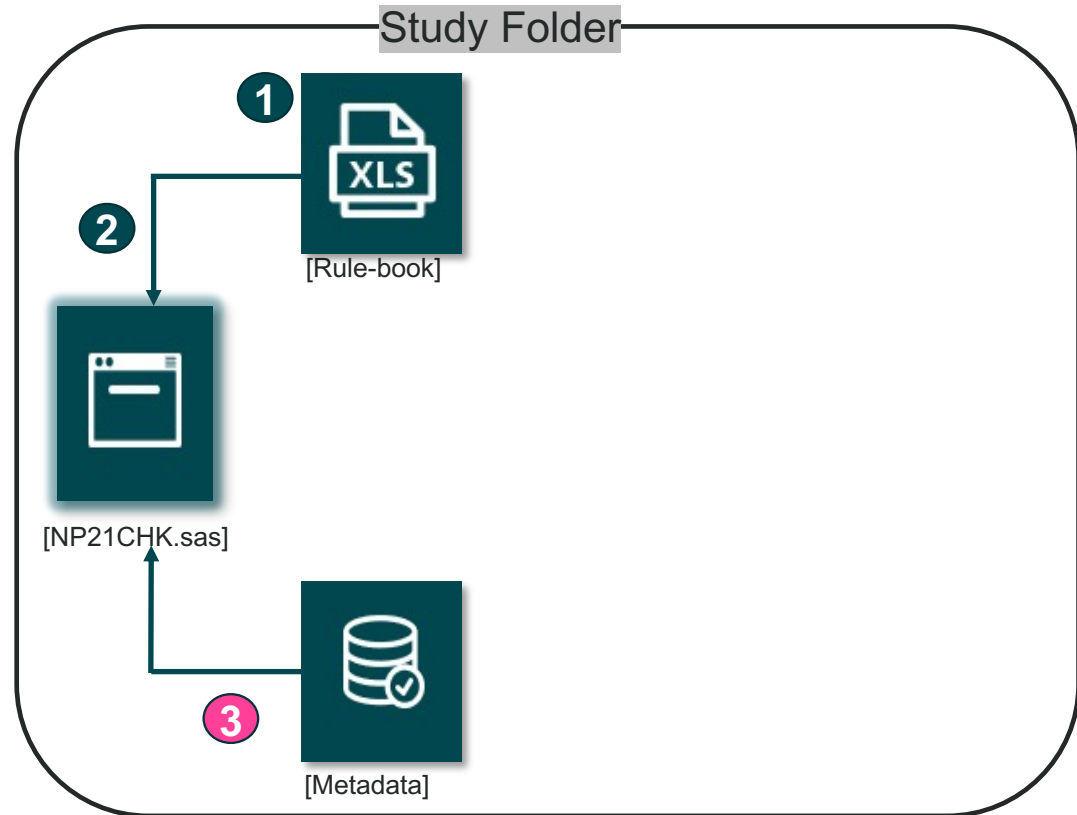
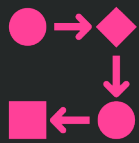
1



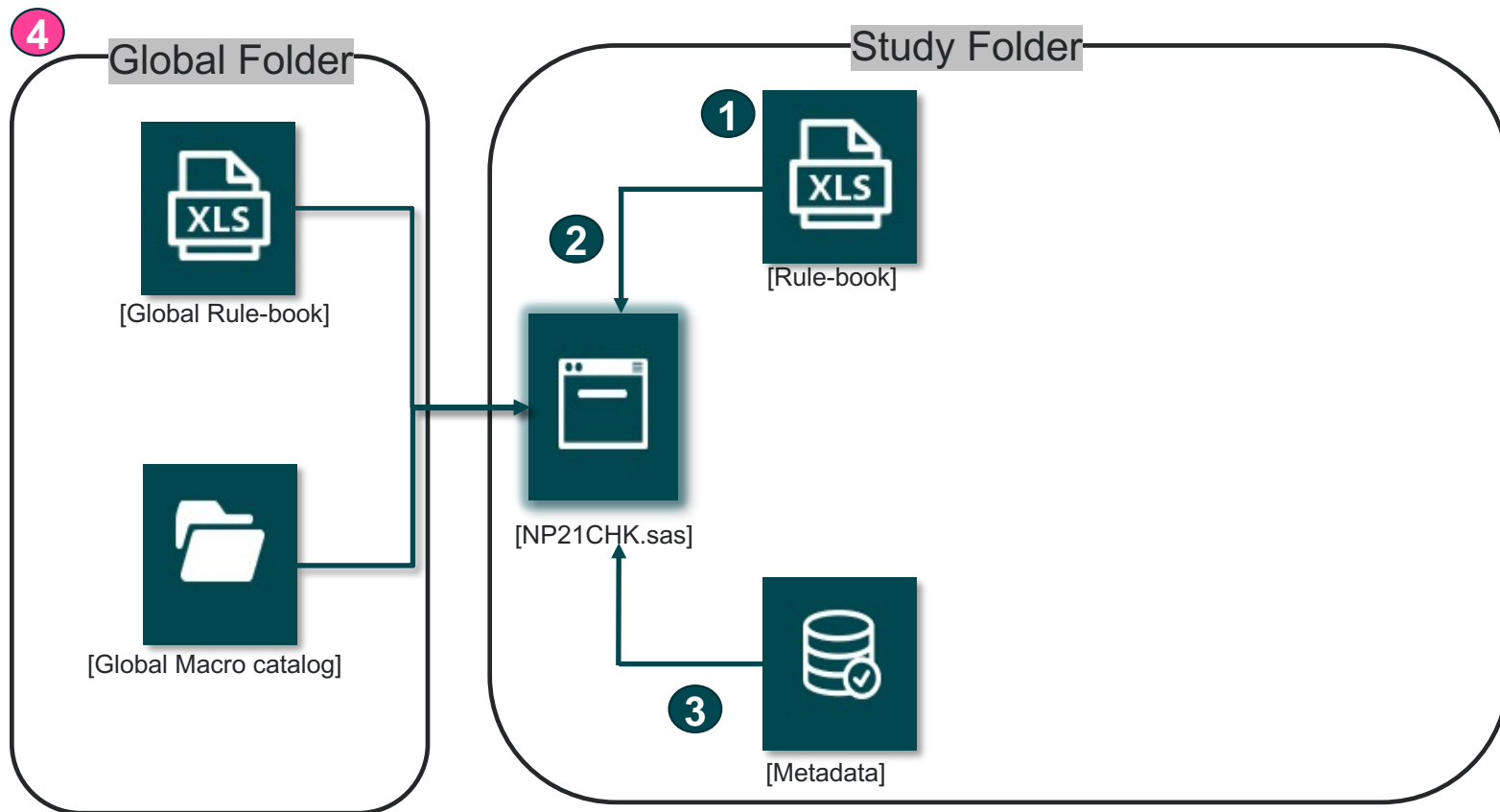
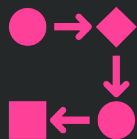
Workflow



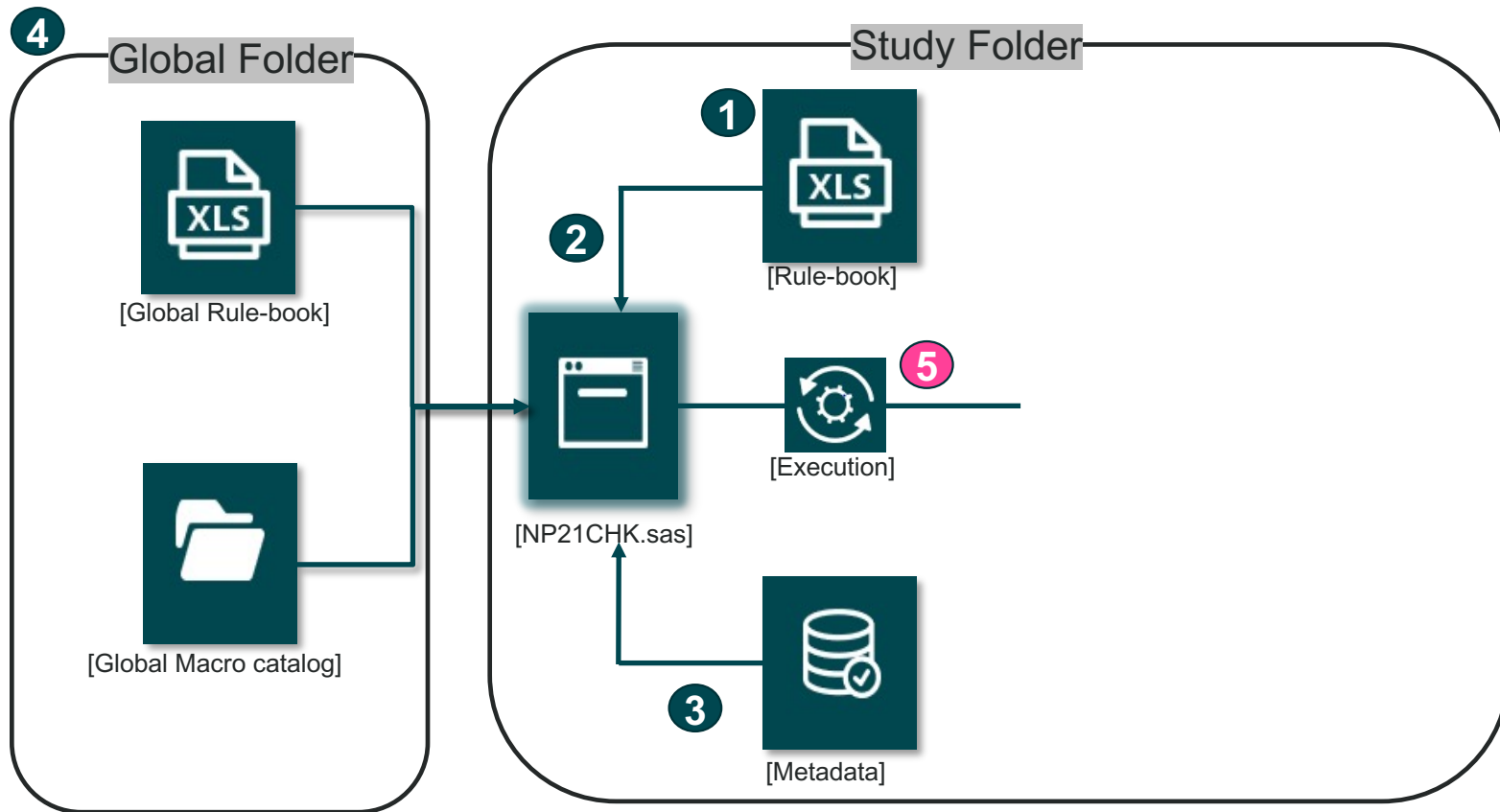
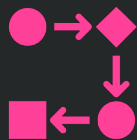
Workflow



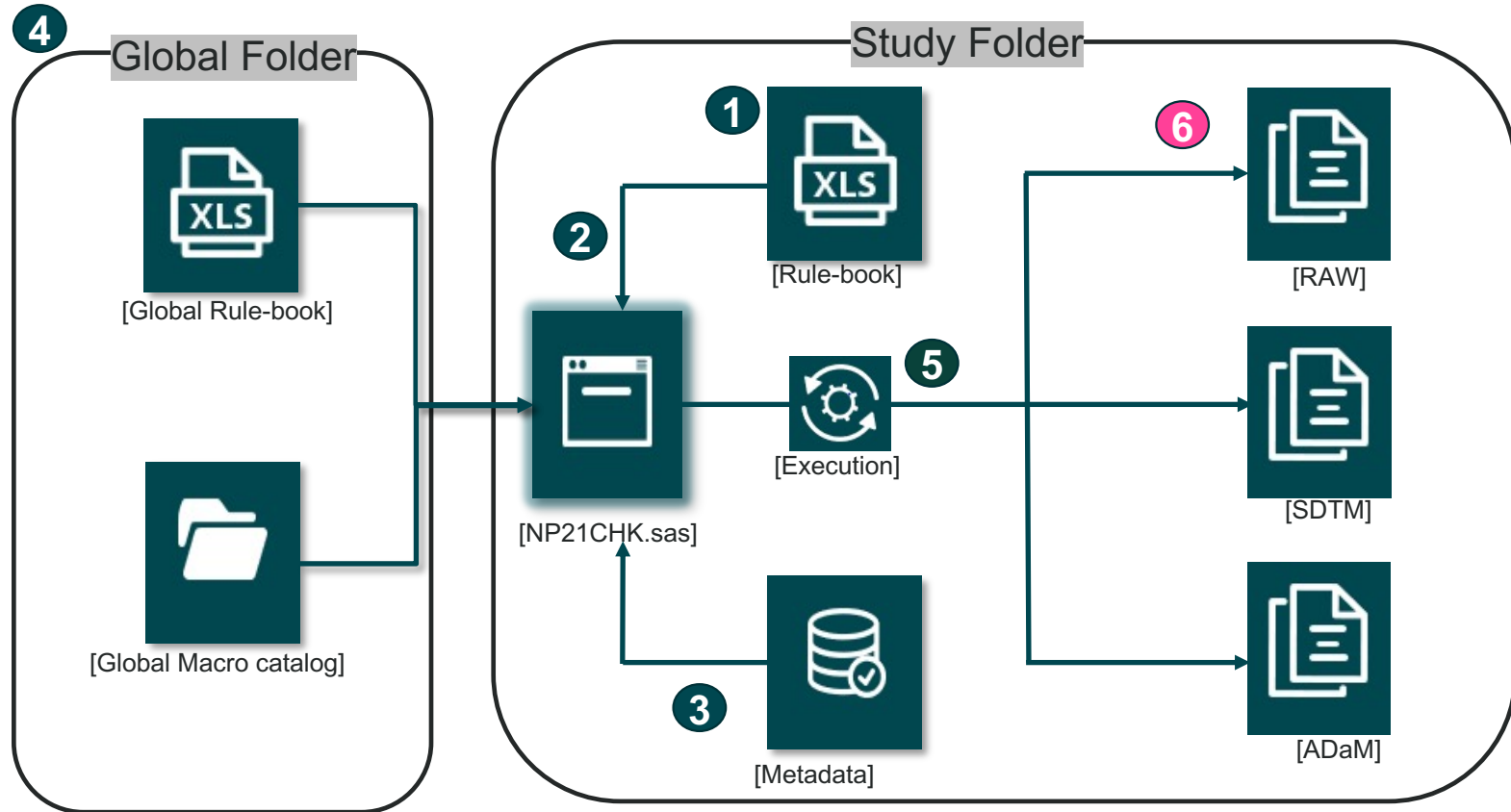
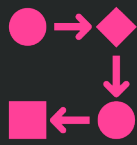
Workflow



Workflow



Workflow



Overview of Rule-book Config sheet



Following are different sheets within Study Rule-book.xlsm

Instructions	Config	Study_Raw	Study_Tabulate	Study_Analysis
CONFIG_PARM	CONFIG_VAL	DESCRIPTION		MANDATORY
STUDY ID		Click on individual cells for more information		Y
FOLDER ID				Y
CT VERSION				N
RAW SUBJECT VARIABLE				N
RAW DATA				N
RAW METADATA				N
TABULATE DATA				N
TABULATE METADATA				N
ANALYSIS DATA				N
ANALYSIS METADATA				N

Config: Users are first required to provide input in CONFIG_VAL column.

Overview of Rule-book

Study_Tabulate sheet



Instructions	Config	Study_Raw	Study_Tabulate	Study_Analysis
--------------	--------	-----------	----------------	----------------

RULE_ID	DATA_TYPE	RULE_NATURE	STUDY_PROJECT	SR_NO	RULE_DESCRIPTION	VALIDATION_NOTE
TABULATE_CUSTOM_ICON1	TABULATE	CUSTOM	ICON	1	Partial dates should not be reported in study as per protocol	Assessment/collection dates are partial
TABULATE_CUSTOM_ICON2	TABULATE	CUSTOM	ICON	2	Actual dose should be consumed at least 5ml per day	EXDOSE is less than 5 ml per day in EX domain

SAS code pasted by user in SASCODE column. It can be either snippet or macro call with % sign

(Continue...)

DOMAIN	VARIABLE	VARVAL_LIST	EXECUTE	CODE_TYPE	SASCODE
ALL	/DTC/	USUBJID DOMAIN	Y	FULL CODE	data &RULE_ID; set tabulate.&DOMAIN; where lengthn(strip(&VARIABLE)) lt 10; run;
EX	EXDOSE	USUBJID DOMAIN EXDOSE EXSEQ	Y	FULL CODE	data &RULE_ID; set tabulate.&DOMAIN; where 0 lt EXDOSE lt 5; run;

- Users shall ensure that code/macro has been tested and works without issues.
- Note: Other two tabs follow the same structure and method of adding checks as Study_Tabulate

Overview of Rule-book

Example of code



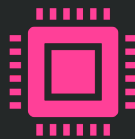
DOMAIN	VARIABLE	VARVAL_LIST	EXECUTE	CODE_TYPE	SASCODE
ALL	/DTC/	USUBJID DOMAIN	Y	FULL CODE	data &RULE_ID; set tabulate.&DOMAIN; where lengthn(strip(&VARIABLE)) lt 10; run;
EX	EXDOSE	USUBJID DOMAIN EXDOSE EXSEQ	Y	FULL CODE	data &RULE_ID; set tabulate.&DOMAIN; where 0 le EXDOSE lt 5; run;

TABULATE_CUSTOM_ICON1

```
data &Rule_id.;  
  set tabulate.&domain.;  
  where 0 lt lengthn(strip(&variable)) lt 10;  
run;
```

TABULATE_CUSTOM_ICON2

```
data &Rule_id.;  
  set tabulate.&domain.;  
  where 0 lt EXDOSE lt 5;  
run;
```

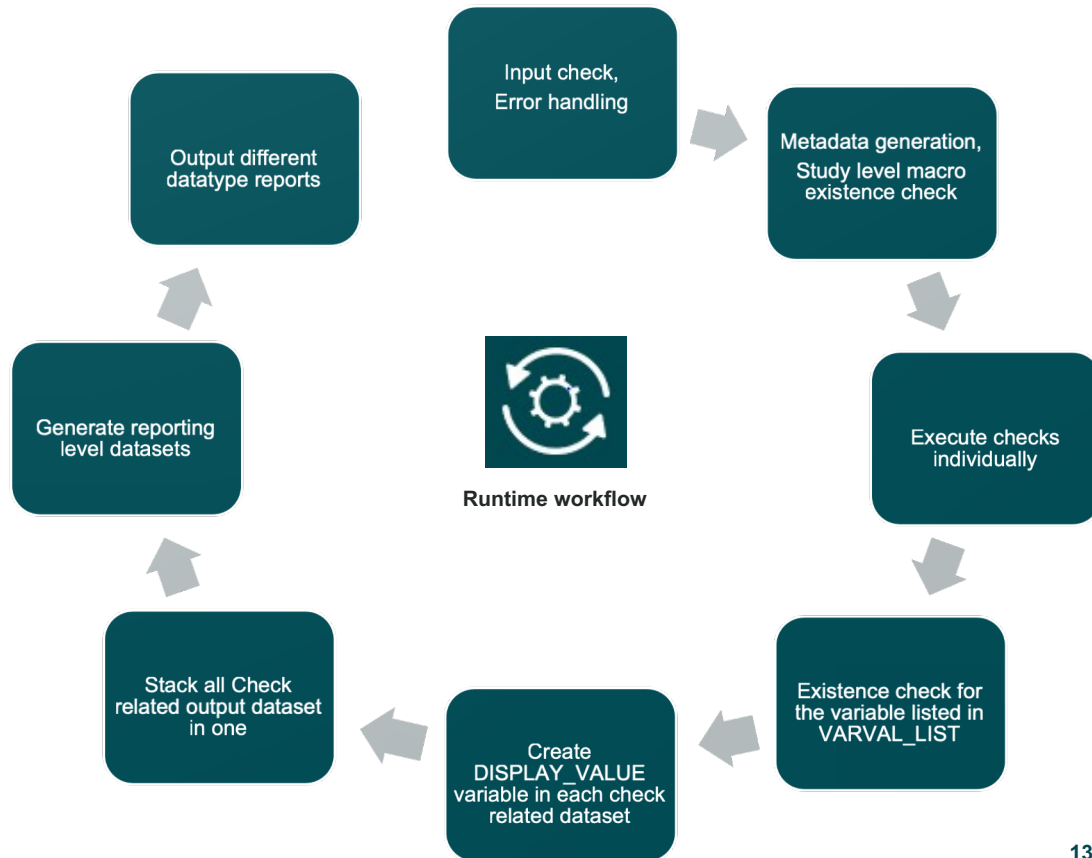


Sample macro call:

```
%NP21CHK(RULEBOOK = \\ICON_01\docs\tabulation\Rule-book.xlsm,  
          DATATYPE = _ALL_ , /*or TABULATE or ANALYSIS or RAW */  
          CLEARUP = N )|
```

Important message

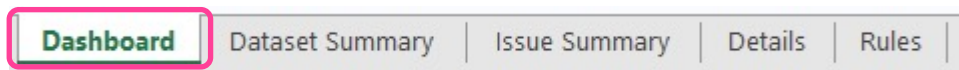
NP21CHK is only an infrastructure macro, it does not create/defines checks on its own. The checks must be pre-programmed in the form of macro or open sas codes in study Rule-book.xlsm or globally accessible location in case of global checks



Output Reports by Macro



- Total 3 reports are generated which are easy to understand as they mimic layout generated by P21
- **Dashboard** tab provides configurational overview



Beyond P21 Data Check Validator Report

Configuration Details

PROJECT ID: ICON_01

STUDY ID: ICON_01

CT VERSION: 5/28/2020

RAW SUBJECT VARIABLE: SUBJECT vendor.PATIENTID

RAW DATA: \\ICON_01\Data\Datasets\raw

RAW METADATA

TABULATE DATA: \\ICON_01\Data\Datasets\tabulation

TABULATE METADATA: \\ICON_01\Data\tabulation\Metadata

ANALYSIS DATA: \\ICON_01\Data\Datasets\analysis

ANALYSIS METADATA: \\ICON_01\Data\Analysis\Metadata

Generated by: LokhandvalaAbdul

Generated on: 2021-08-26T07-29-57

Output Reports by Macro



- **Data set summary** tab provides counts of affected rows of a dataset.
- Zero value indicates there were no issues in datasets on which check(s) executed

Dashboard	Dataset Summary	Issue Summary	Details	Rules
-----------	-----------------	---------------	---------	-------

Beyond P21 Data Check Validator Report			
Dataset	Dataset Label	Affected Records	Total Records
AE	Adverse Events	0	26
DM	Demographics	4	97
DS	Disposition	82	247
MH	Medical History	0	332
SUPPAE	Supplemental Qualifiers for AE	0	52
SUPPCM	Supplemental Qualifiers for CM	0	2049
SUPPDM	Supplemental Qualifiers for DM	0	150

Output Reports by Macro



- **Issue summary** tab provides counts of each checks per dataset.
- **Description** column has an overall description about what each check did

Dashboard	Dataset Summary	Issue Summary	Details	Rules
-----------	-----------------	---------------	---------	-------

Beyond P21 Data Check Validator Report

Dataset	Rule ID	Description	Fou
DM	TABULATE_CUSTOM_ICON01	Domain date is greater than RFPENDTC	4
DS	TABULATE_GLOBAL_ALL6	EXSTDTC or EXENDTC is not less than or equal to the latest treatment disposition date	82
SUPPDM	TABULATE_GLOBAL_ALL3	QNAM is either absent in Specification VLM sheet or in SUPP dataset.	1
SUPPMB	TABULATE_GLOBAL_ALL3	QNAM is either absent in Specification VLM sheet or in SUPP dataset.	1

Output Reports by Macro



- **Details** tab provides granulated information on affected row(s) of each check
- Users can filter rows to precisely identify values within variables
- *Validation Note* column depicts exact issue

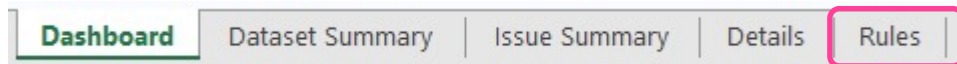
Dashboard	Dataset Summary	Issue Summary	Details	Rules
-----------	-----------------	---------------	---------	-------

Dataset	Variable	Record Number	Variable List	Variable Value Display	Rule ID	Validation note
SUPPDM	QNAM	N/A	DMN QNAM	SUPPDM, RACEOTH	TABULATE_GLOBAL_ALL3	QNAM is either absent in Specification VLM sheet or in SUPP dataset.
SUPPMB	QNAM	N/A	DMN QNAM	SUPPMB, MBVAR1	TABULATE_GLOBAL_ALL3	QNAM is either absent in Specification VLM sheet or in SUPP dataset.
DM	RFPENDTC	N/A	USUBJID DTC DT_NAME	ICON_01_001, 2018-11-22T12:26, LBDTC	TABULATE_CUSTOM_ICON1	Domain date is greater than RFPENDTC

Output Reports by Macro



- **Rules** tab provides a list of checks that were applicable and executed successfully
- For affected row Users can refer to either **Issue Summary** or **Details**

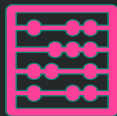


Dataset	Variable	Rule ID	Description	Rule Applicable
SUPPAE	QNAM	TABULATE_CUSTOM_ICON2	QNAM=AETRTEM is missing from SUPPAE	Y
DM	RFPENDTC	TABULATE_CUSTOM_ICON1	Domain date is greater than RFPENDTC	Y
AE	DOMAIN	TABULATE_GLOBAL_ALL1	MedDRA coded variables are not included in SDTM AE domain	Y
SUPPDM	QNAM	TABULATE_GLOBAL_ALL11	DM.RACE has values as 'Other' but SUPPDM dataset does not have QNAM='RACEOTH'	N
SUPPCM	DOMAIN	TABULATE_GLOBAL_ALL2	ATC coded variables are not included in SDTM SUPPCM domain	Y
SUPPALL	QNAM	TABULATE_GLOBAL_ALL3	QNAM is either absent in Specification VLM sheet or in SUPP dataset.	Y
AE	AE_VAR	TABULATE_GLOBAL_ALL4	AE Start/End is before RFCIDTC date	Y
MH	MHSTDTC	TABULATE_GLOBAL_ALL5	MHSTDTC date is on or after RFXSTDTC	Y
DS	DSSTDTC	TABULATE_GLOBAL_ALL6	EXSTDTC or EXENDTC is not less than or equal to the latest treatment disposition date	Y
AE	DMN_VAR	TABULATE_GLOBAL_ALL8	AE flag variables: AESCONG, AESDISAB, AESDTH, AESHOSP, AESLIFE, AECONTRT, AESCAN, AESMIE, AESOD, AESER, AESDISA do not follow NY codelist	Y
SUPPDM	RACE	TABULATE_GLOBAL_ALL9	Subjects with RACE='MULTIPLE' value must have at least two records in SUPPDM with QNAM='RACE1' or 'RACE2'	Y



- Easy to Understand/Interpret produced reports
- Raw dataset report can be directly fed back to DM team
- Flexibility to define checks on any Data Type
- Mixed type checks can be executed
- Repository of General Global checks can be established
- Repository of Customer/Project/Molecule specific checks can be established

Tradeoffs



- *Execution time* depends on amount of Study data and checks defined
- Output dataset names must start with <RULE_ID/DOMAIN> or &RULE_ID/DOMAIN
- Requires users' proper attention while filling Rule-book
- Providing flexibility of adding code as Open SAS code

Future Enhancements

- Adding modular macros which have different functionality
- Adding a functionality to import rules from different study Rule-books to save time
- Developing UI to add checks and have a proper database to store different elements



Ideas



Q&A

THANK YOU

ABDULKADIR LOKHANDVALA
SENIOR PRINCIPAL PROGRAMMER
[Lokhandvalaabdul@prahs.com]

iconplc.com

