



doLoop Technologies

Business Analytics. Delivered.

ADaM - Where Do I Start?

By :- Krupali Ladani & Dr.Sangram Parbhane

- ❑ The opinions in this presentation are those of the presenter and may not necessarily reflect the views of doLoopTech, PhUSE or CDISC.
- ❑ All the examples displayed as 'Table no. x' are taken from ADaMIG_v1.1

❑ ADaM Introduction

❑ ADaM Specification Development

- Where Do I Start ?
- Considerations
- Importance
- Challenges
- Recommendation

❑ ADaM Dataset Programming

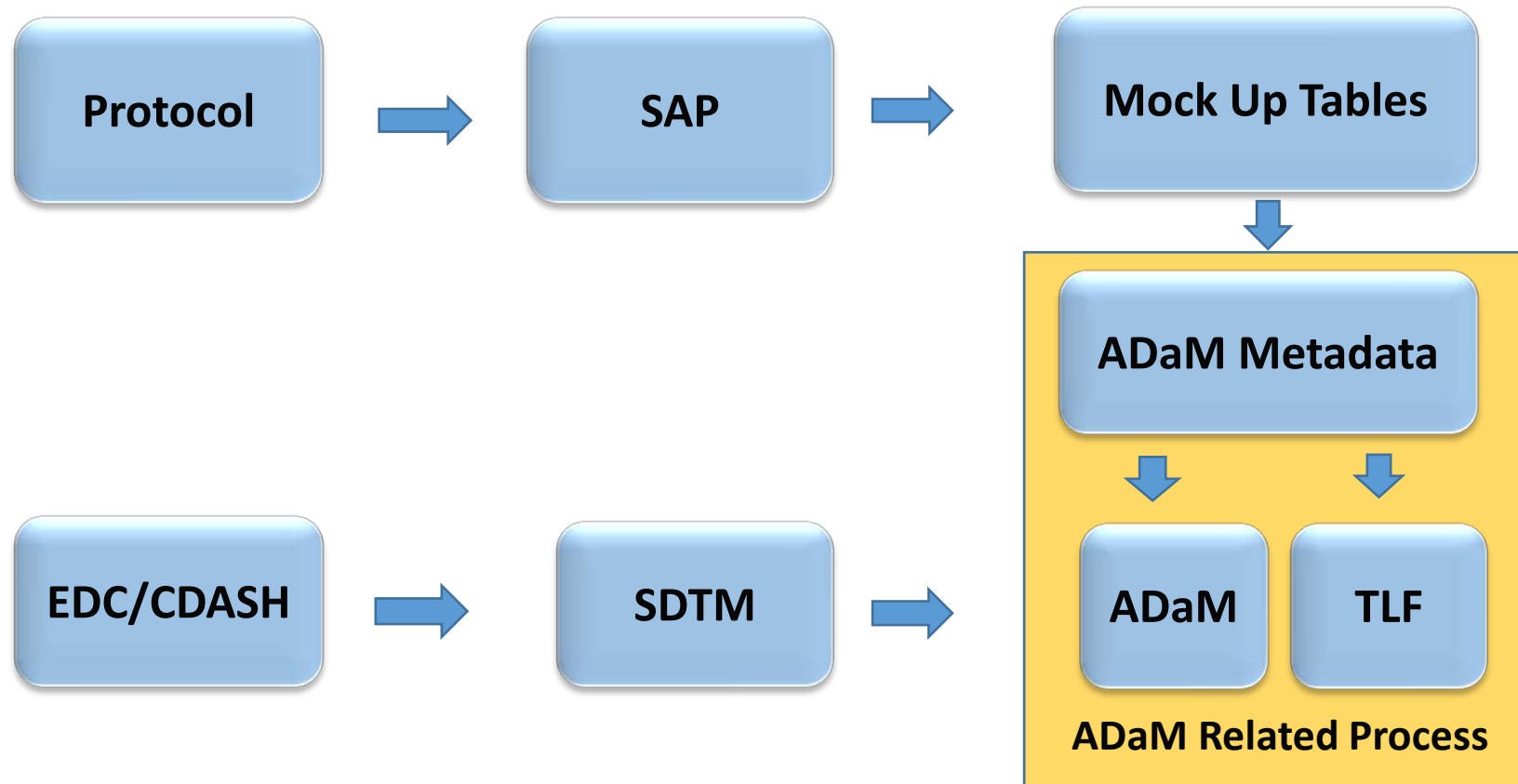
- Where Do I Start ?
- Considerations
- Importance
- Challenges
- Recommendation

❑ ADaM QC

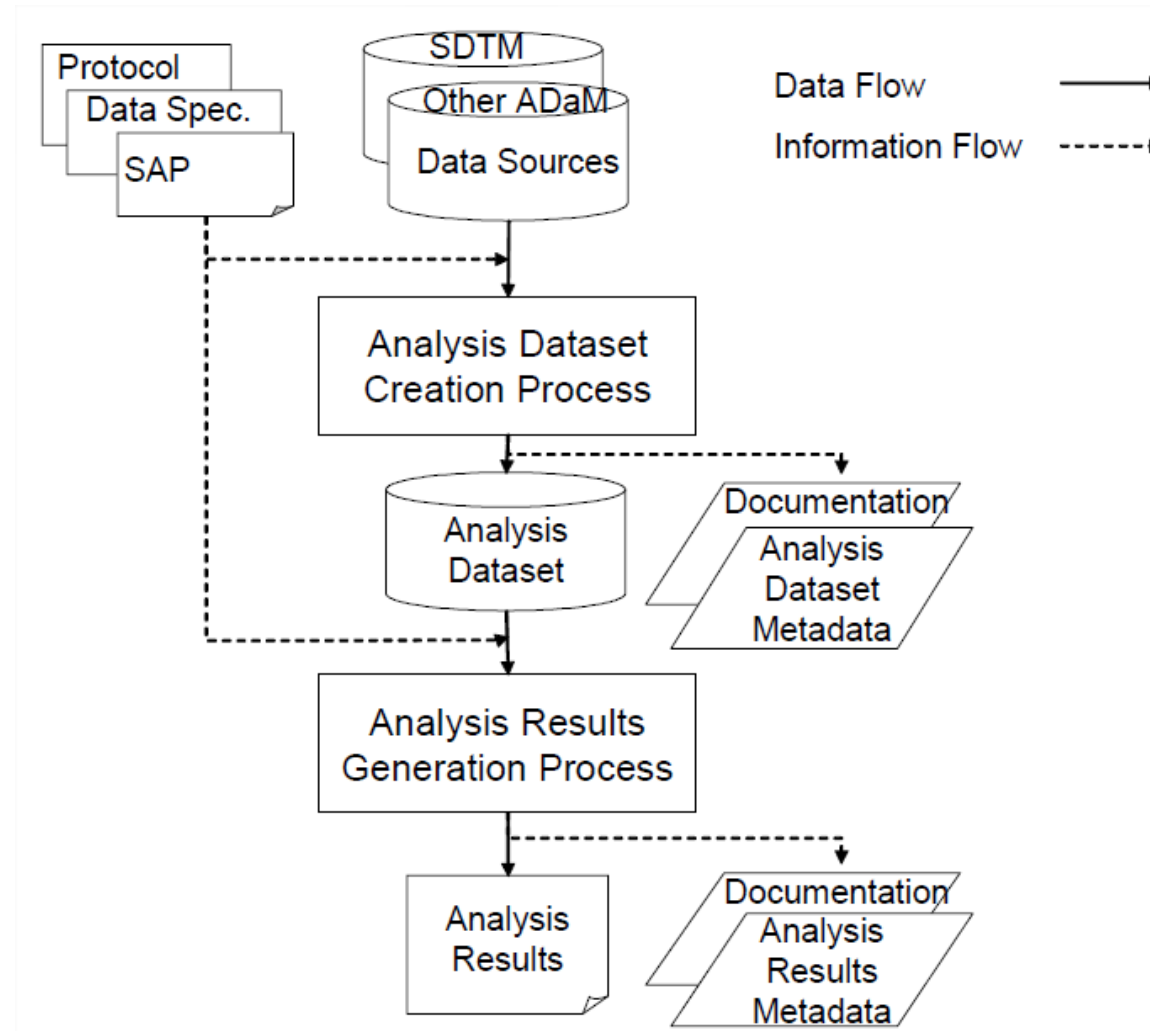
- Where Do I Start ?
- Considerations
- Importance
- Challenges
- Recommendation

Analysis Data Model (ADaM)

“Framework that enables analysis of the data, while at the same time allowing reviewers and other recipients of the data to have a clear understanding of the data’s lineage from collection to analysis to results”

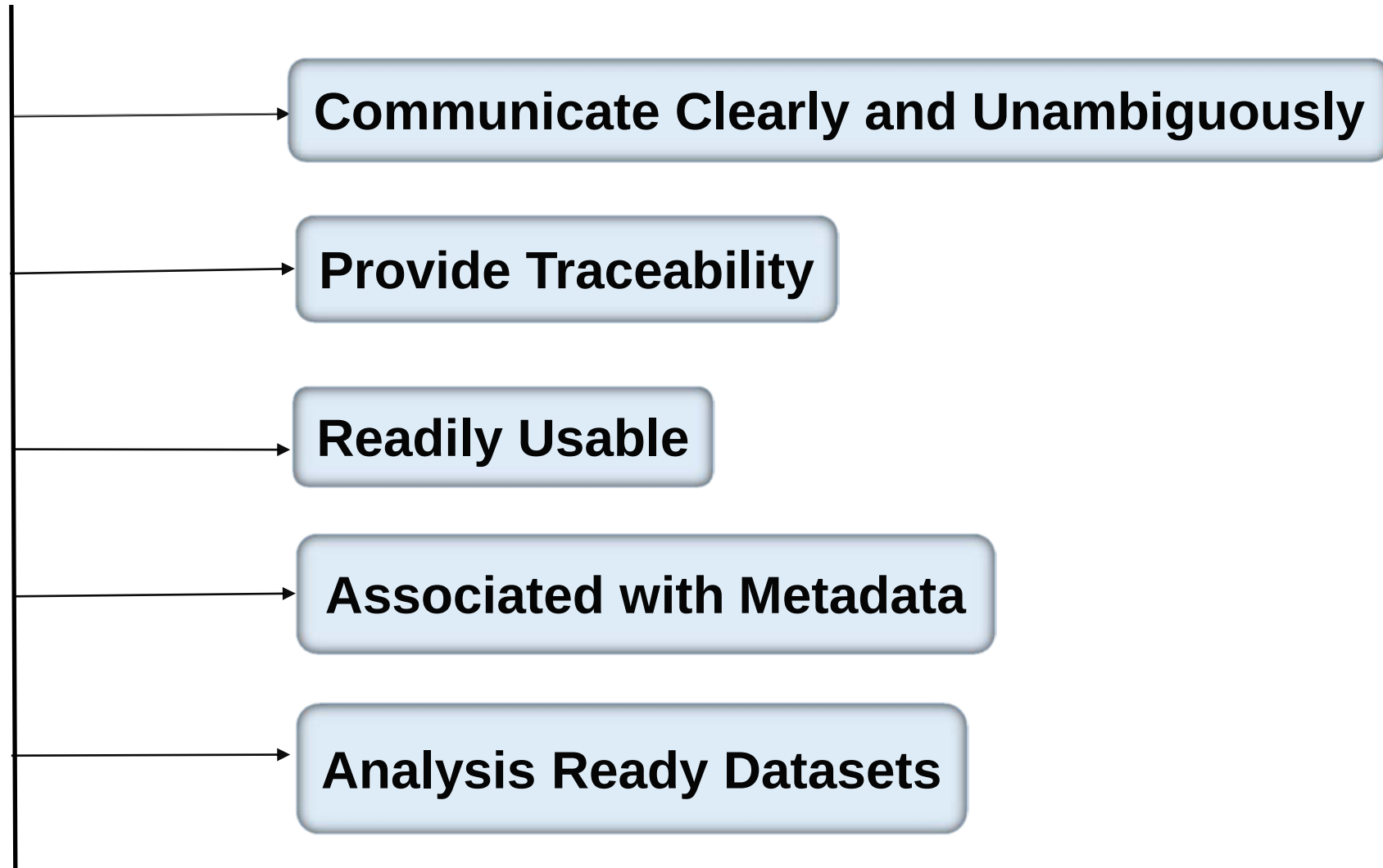


CDISC ADaM V2.1 - Analysis data flow



Analysis Data Flow Diagram Showing One Scenario for the Flow of Data and Information

Fundamental Principles of ADaM



The ADaM Data Structure

ADSL	BDS	OCCDS
1) Known as “ Subject level Analysis Data ”	1) Known as “ Basic Data Structure ”	1) Known as “ Occurrence Data Structure ”
2) Structure: One record per subject	2) Structure: Contains one or more records per subject, per analysis parameter, per analysis time point	2) Structure: One record per record in SDTM domain (optional: per coding path, per Analysis Period and/or Phase)
3) Contains subject-level information	3) Contains PARAM, AVAL and AVALC and related variables	3) Support Occurrence Data Models such as Medical History, Concomitant Medications, and Lab Events
4) Example: ADSL	4) Example: ADLB, ADVS	4) Example: ADAE, ADMH

ADSL Variables

Variables	Examples
Subject Identifier Variables	STUDYID, USUBJID, SUBJID, SITEID, SITEGRy, SITEGRyN, REGIONy, REGIONyN
Subject Demographics Variables	AGE, AGEU, AGEGRy, AGEGRyN, AAGE, SEX, RACE, RACEGRy, RACEGRyN
Population Indicator Variables	FASFL, SAFFL, ITTFL, PPROTFL, COMPLFL, RANDFL, ENRLFL
Treatment Variables	ARM, ACTARM, TRTxxP, TRTxxPN, TRTxxA, TRTxxAN
Dose Variables	DOSExxP, DOSExxA, DOSExxU
Subject-Level Period, Subperiod, and Phase Timing Variables	APxxSDT, APxxSTM, APxxETMF, PxxSw, PxxSwSDT, PHwETM, PHwEDTM
Subject-Level Trial Experience Variables	EOSSTT, EOSDT, DCSREAS, DCSREASP

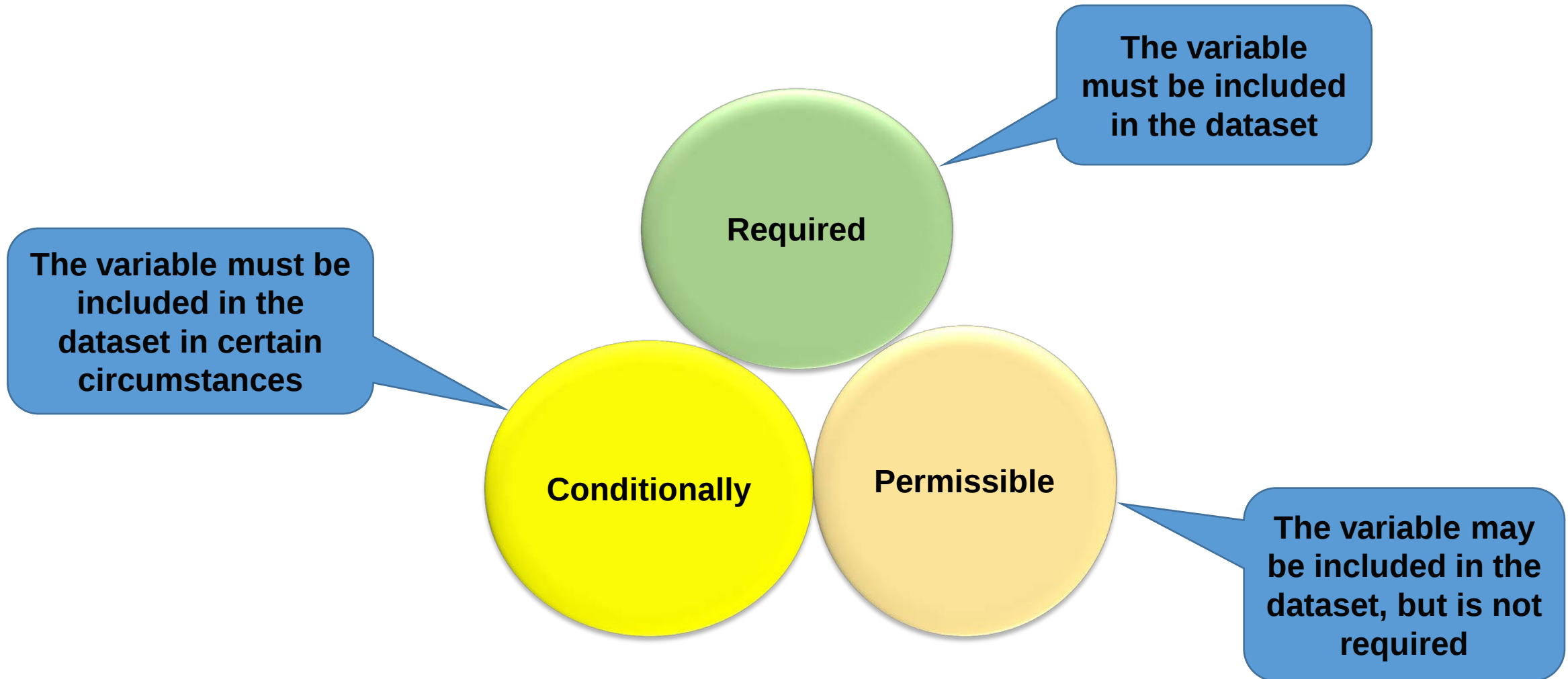
BDS Variables

Variables	Examples
Identifier Variables	STUDYID, USUBJID, SUBJID, ASEQ
Record-Level Treatment and Dose Variables	TRTA , TRTAN, TRTPGy, DOSEP, DOSCUMP, DOSEA, DOSCUMA
Timing Variables	ADT, ATM, ADTM, ADY, AVISIT, APERIODC, APERSDT
Analysis Parameter Variables	PARAM, PARAMCD, PARAMTYP, AVALC, BASETYPE, SHIFTy, CRITy
Analysis Descriptor Variables	DTYPE, AWRANGE, AWTARGET, AWTDIFF, AWLO
Time-to-Event Variables	STARTDT, STARTDTM, CNSR, EVNTDESC, CNSDTDSC
Toxicity and Range Variables	ATOXGR, BTOXGR, ANRIND
Indicator Variables	ABLFL, ANLzzFL, ONTRTFL, ITTRFL, SAFRFL
Data point Traceability Variables	SRCDOM, SRCVAR, SRCSEQ

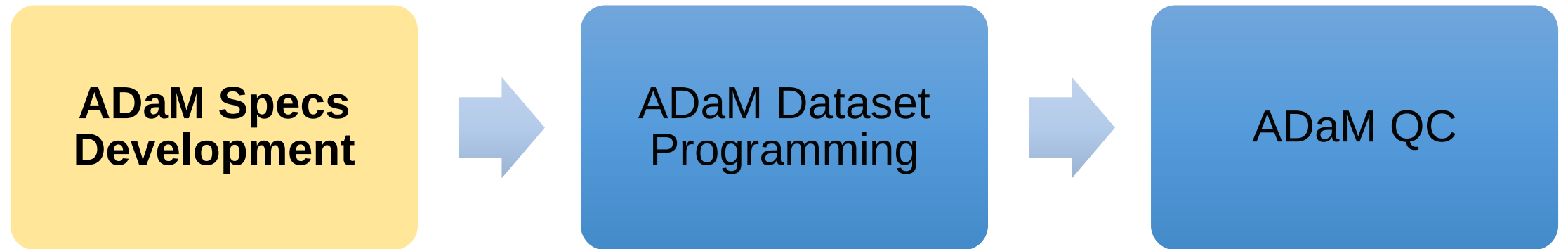
OCCDS Variables

Variables	Examples
Identifier Variables	STUDYID, USUBJID, SUBJID, -SEQ
Dictionary Coding and Categorization Variables	-TERM, -DECOD, -BODSYS, -LLT, -LLTCD, CMCLAS, CMCLASCD, ACATy
Timing Variables	ASTDT, ASTTM, APERIOD, APERIODC, APHASE
Indicator Variables	-OCCUR , --PRESF, ANLzzFL, TRTEMFL, AETRTEM, ONTRTFL, PREFL
Occurrence Flag Variables	AOCCFL, AOCCPFL, AOCCSFL
Treatment/Dose Variables	DOSEON, DOSCUMA
Descriptive Variables	AESER, SEVGRy , AREL, RELGRy , CMINDC

Values of ADaM “Core” Attribute

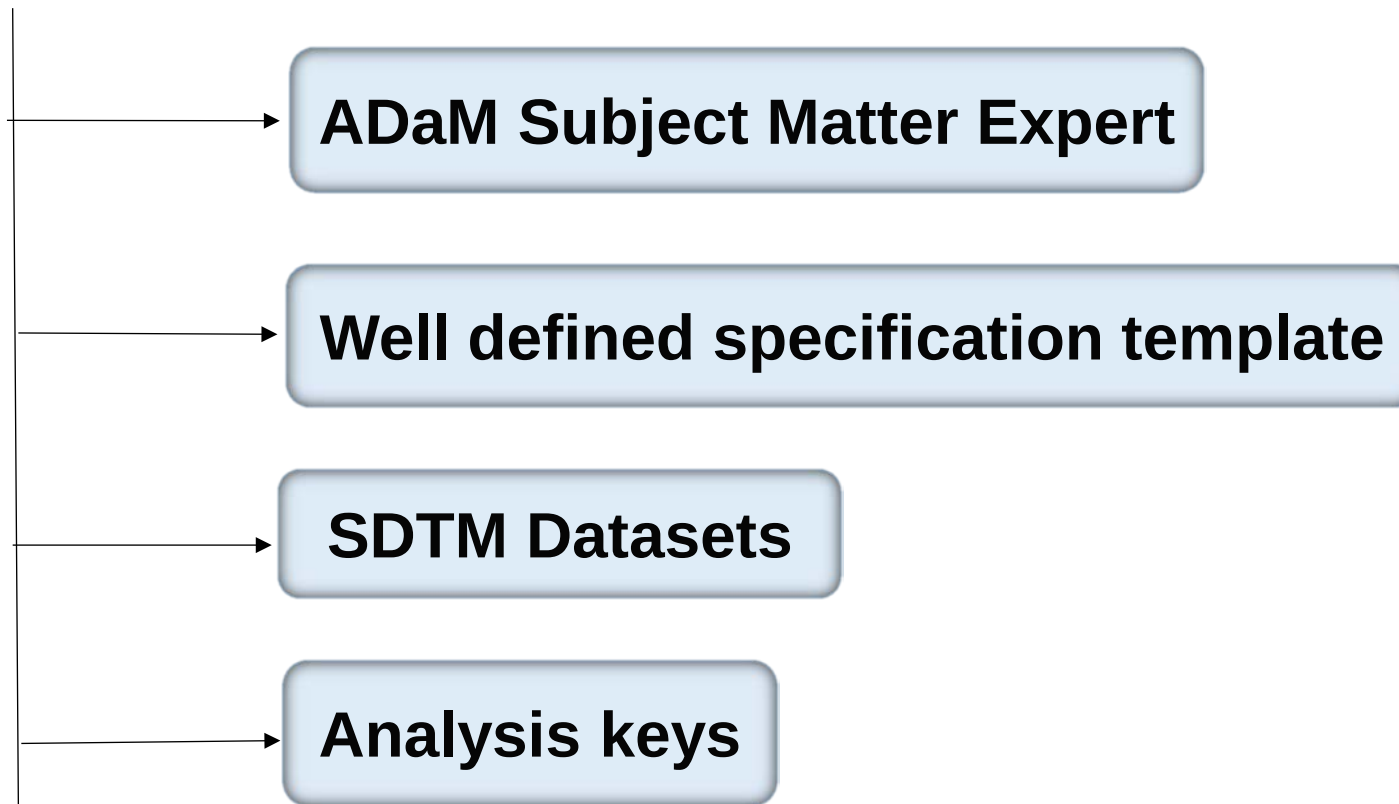


Lets get started...



ADaM Specification Development: Where Do I Start ?

ADaM Specification Development is the starting point for ADaM process after SDTM. Here is the ABC's for specification developments....

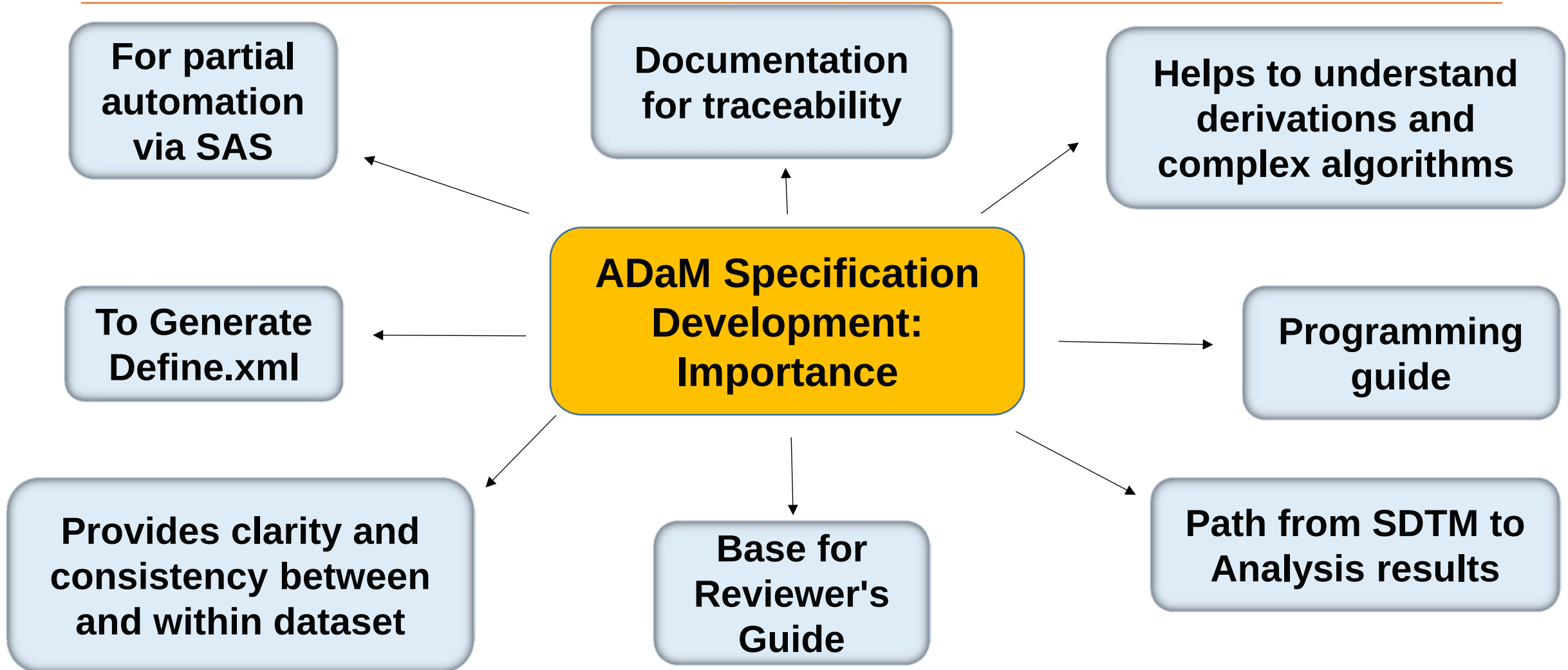


ADaM Specification Development: Considerations

Let's assume we live in an ideal world and creating our specification before starting programming!

- ☐ Identifying requirement of number of datasets based on SAP
- ☐ Readiness of SDTM data
- ☐ Clear specification template to define all required components
- ☐ ADSL as a base for other datasets
- ☐ Understanding of OCCDS and its requirements e.g. MedDRA Coding
- ☐ Must have variables to support Traceability and Analysis Results

ADaM Specification Development: Importance



Can't think about dense datasets like ADaM without specifications!

ADaM Specification Development: Challenges

There's never a road without a turning - we can't expect ADaM to be as simple as SDTM

- ❑ Need to maintain a document which requires frequent updates as ADaM specification is an important source of traceability
- ❑ Compatible for metadata as specifications usually used as the basis for the generation of ADaM Define.xml and Define.pdf
- ❑ Adherence to ADaM IG for datasets structure and variable attributes
- ❑ Defining complex analysis variables well in advance

ADaM Specification Development: Recommendation

- ❑ Identify any discrepancy in SDTM during specification development level
- ❑ Follow the standard process, prepare specifications first and then start programming
- ❑ Adhere ADaM Implementation Guide and write specifications for ADaM datasets instead ADaM like datasets from beginning
- ❑ ADaM is the extended process and changes or updates are expected at any point of time, document and track the updates applicable to specification without delaying

Example

Domain Sheet

Order	Active	Dataset	Description	Structure	Purpose	Keys	Location
1	Y	ADSL	Subject Level Characteristics	One record per subject	Analysis	STUDYID, USUBJID	ADSL.xpt
2	Y	ADAE	Adverse Events	One record per Event per subject	Analysis	STUDYID, USUBJID, AETERM, AEDECOD, AEBODSYS, ASTDT, AENDT	ADAE.xpt

ADSL Sheet

order	Domain Name	Active	Variable Name	Variable Label	Type	Codelist	Origin of Variable (Protocol, Assigned, Derived, eDT, CRF Page no.)	Internal Variable Length	Mapping Rules	Notes provided by CDISC v1.1	Type of CDISC Variable
1	ADSL	Y	STUDYID	Study Identifier	text			12	DM.studyid;	Must be identical to the SDTM variable DM.STUDYID.	Req
2	ADSL	Y	DOMAIN	Domain Abbreviation	text	DOMAIN		4	domain = 'ADSL';		Req

Lets get started...



ADaM Dataset Programming: Where Do I start ?

ADaM is definitely a combined and parallel effort from Specification developer and Programmer. Instead following specifications blindly, inform developer to provide feedback for inconsistency in specifications as and when required.

❑ Look into the source datasets

- *Going through the SDTM datasets provided, having basic understanding about it e.g. total number of subjects in DM, number of subjects which are screen failure, treatment arms and trial design*

❑ ADaM Implementation guide does not define how to write the 'Mapping Rules' column in specification

- *Understand how your specifications written, the simple English or mixed with programming codes*
- *Is there any different columns or indicators for direct SDTM variables or for complex algorithms ?*

ADaM Dataset Programming: Consideration

❑ Basic understanding about ADaMIG or ADaM variables is the key guide along with specification document.

Example 1 - Naming convention for variables

FL	Suffix used in names of character flag variables
DT	Suffix used in names of numeric date variables
TM	Suffix used in names of numeric time variables
DTM	Suffix used in names of numeric datetime variables
DTF	Suffix used in names of date imputation flag variables
TMF	Suffix used in names of time imputation flag variables
DY	Suffix used in names of relative day variables that do not include day 0

Table 1- Naming convention for date and flag variables

Example 2 - Visit windows and Unscheduled visits

- ❑ The record that falls closest to the scheduled visit day is the one that will be analysed and indicated using ANL01FL
- ❑ Used AWTARGET and AWTDIFF to indicate more clearly how the analysed rows were selected from among the candidate rows

ADVS

Row	USUBJID	VISIT	AVISIT	ADY	PARAM	AVAL	DTYPE	ANL01FL	AWTARGET	AWTDIFF
1	1001	Screening	Baseline	-5	SUPINE SYSBP (mm Hg)	144			1	5
2	1001	Baseline	Baseline	1	SUPINE SYSBP (mm Hg)	145		Y	1	0
3	1001	Week 1	Week 1	7	SUPINE SYSBP (mm Hg)	130		Y	7	0
4	1001	Week 2	Week 2	12	SUPINE SYSBP (mm Hg)	133		Y	14	2
5	1001	Week 3	Week 2	17	SUPINE SYSBP (mm Hg)	125			14	3
6	1001	Week 4	Week 4	30	SUPINE SYSBP (mm Hg)	128		Y	28	2

Table 2: Identification of rows used for analysis in ADVS

Example 3 - How to handle missing values

At Week 8, there is a scheduled visit (visit number 6), at that visit, blood pressure should be collected. However, no data on blood pressure were collected. The SAP says that missing post-baseline data should be imputed by two methods: LOCF (last observation carried forward), and WOCF (worst observation carried forward)

ADVS

Row	PARAM	AVISIT	AVISITN	VISITNUM	VSSEQ	ABLFL	AVAL	BASE	CHG	DTYPE	ADY	AWTARGET	AWTDIFF	ANL01FL
1	Systolic BP (mm Hg)	Screening	-4	1	3821		120	114	.		-30	-28	2	Y
2	Systolic BP (mm Hg)	Run-In	-2	2	3822		116	114	.		-16	-14	2	Y
3	Systolic BP (mm Hg)	Week 0	0	3	3823	Y	114	114	0		-2	1	2	Y
4	Systolic BP (mm Hg)	Week 2	2	4	3824		118	114	4		13	14	1	Y
5	Systolic BP (mm Hg)	Week 2	2	4.1	3825		126	114	12		17	14	3	
6	Systolic BP (mm Hg)	Week 4	4	5	3826		122	114	8		23	28	5	Y
7	Systolic BP (mm Hg)	Week 8	8	5	3826		122	114	8	LOCF	23	56	33	Y
8	Systolic BP (mm Hg)	Week 8	8	4.1	3825		126	114	12	WOCF	17	56	39	Y
9	Systolic BP (mm Hg)	Week 12	12	7	3827		134	114	20		83	84	1	Y

Table 3 - Creation of New Rows to Handle Imputation of Missing Values by LOCF & WOCF

Example 4 - All the flag values are not same

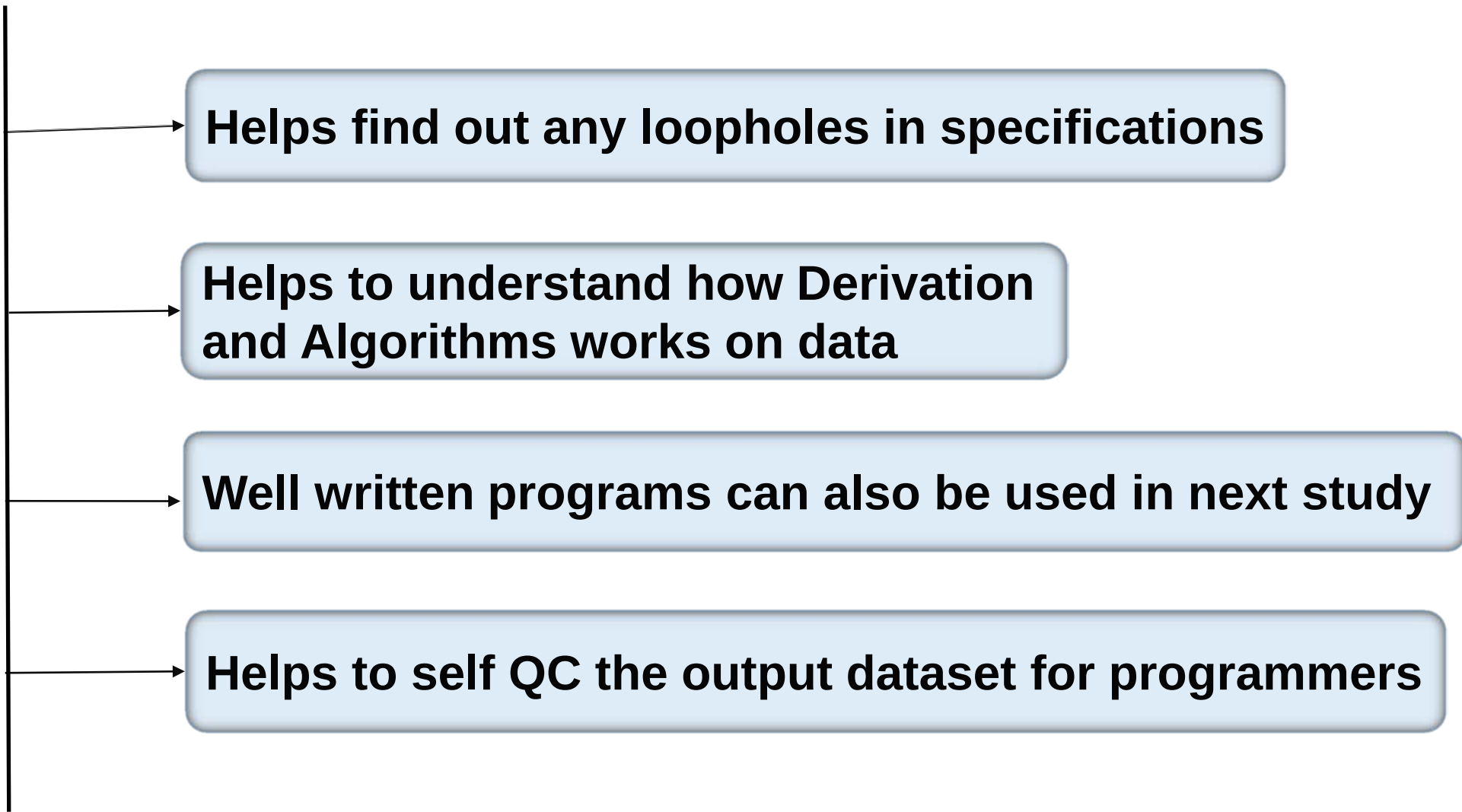
- ❑ Subject-level character population flag values can be 'Y'/'N' or Null while parameter-level and record-level character population flag are 'Y' and Null.
- ❑ How about date imputation flag values ?

BDS

Row	USUBJID	ITTFL	PPROTFL	VISIT	AVISIT	PARAMCD	AVAL	ANL01FL	PPROTRFL
1	1001	Y	Y	Week 0	Week 0	TEST1	500	Y	Y
2	1001	Y	Y	Week 1	Week 1	TEST1	400	Y	Y
3	1001	Y	Y	Week 2	Week 2	TEST1	600	Y	Y
4	1002	Y	N	Week 0	Week 0	TEST1	500	Y	
5	1002	Y	N	Week 2	Week 1	TEST1	48	Y	
6	1002	Y	N	Week 2	Week 2	TEST1	46	Y	
7	1003	Y	Y	Week 0	Week 0	TEST1	999	Y	Y
8	1003	Y	Y	Week 1	Week 1	TEST1	999		Y
9	1003	Y	Y	Retest	Week 1	TEST1	49	Y	
10	1003	Y	Y	Week 2	Week 2	TEST1	499	Y	

Table 4: ADaM Dataset with Subject-Level and Record-Level Indicator Variables in BDS

ADaM Dataset Programming: Importance



Helps find out any loopholes in specifications

Helps to understand how Derivation and Algorithms works on data

Well written programs can also be used in next study

Helps to self QC the output dataset for programmers

ADaM Dataset Programming: Challenges

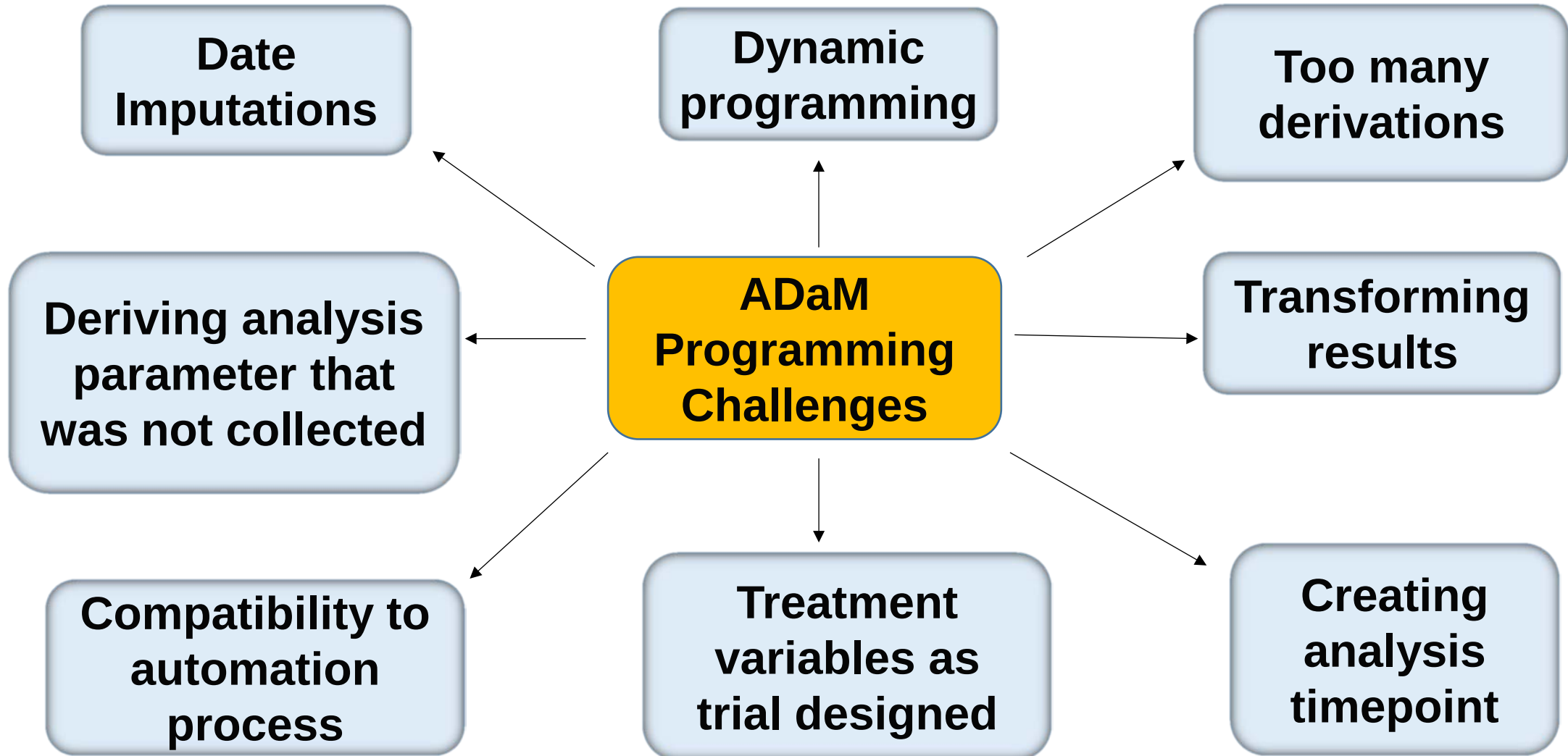
- ❑ FDA CDER draft guidance for Study Data Technical Conformation Guide stats that

“One of the expected benefits of analysis datasets that conform to ADaM is that they simplify the programming steps necessary for performing an analysis”

Does this mean to compensate programming efforts in ADaM ?

May be YES!.....

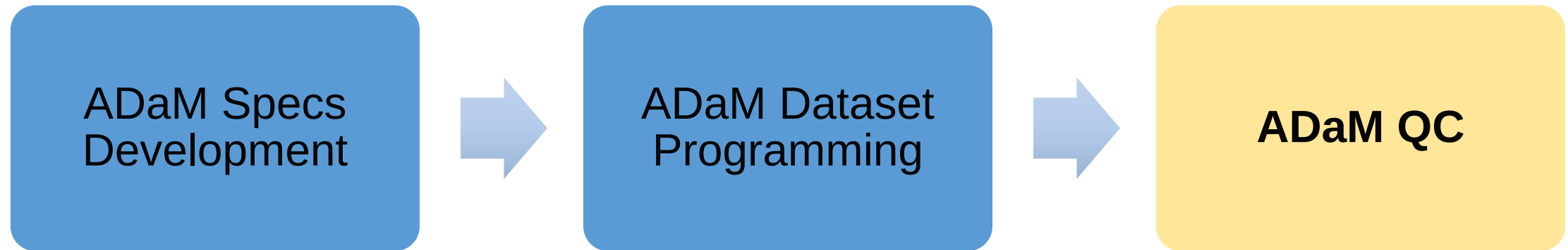
- ❑ Correct interpretation and implementation of specification document
- ❑ Consider all situations in advance without having such data e.g. missing or partial values during early development
- ❑ ADaM requires update when SDTM datasets are updated or TLF shells updated with new requirement from ADaM



ADaM Dataset Programming: Recommendation

- ❑ Basic knowledge about ADaM implementation guide
- ❑ Running Pinnacle21 Validator individually
- ❑ Simple programming which can be edited easily at any point without major changes
- ❑ Develop generic SAS Macro utilities which be utilized across all programming
- ❑ Wherever possible go for automation

Lets get started...



ADaM QC: Where do I Start ?

❑ QC goal should be

- Tracking ADaM data back to SDTM
- Checking compliance with the ADaM Implementation Guide
- Checking compliance of analysis variables and its derivation with required analysis results

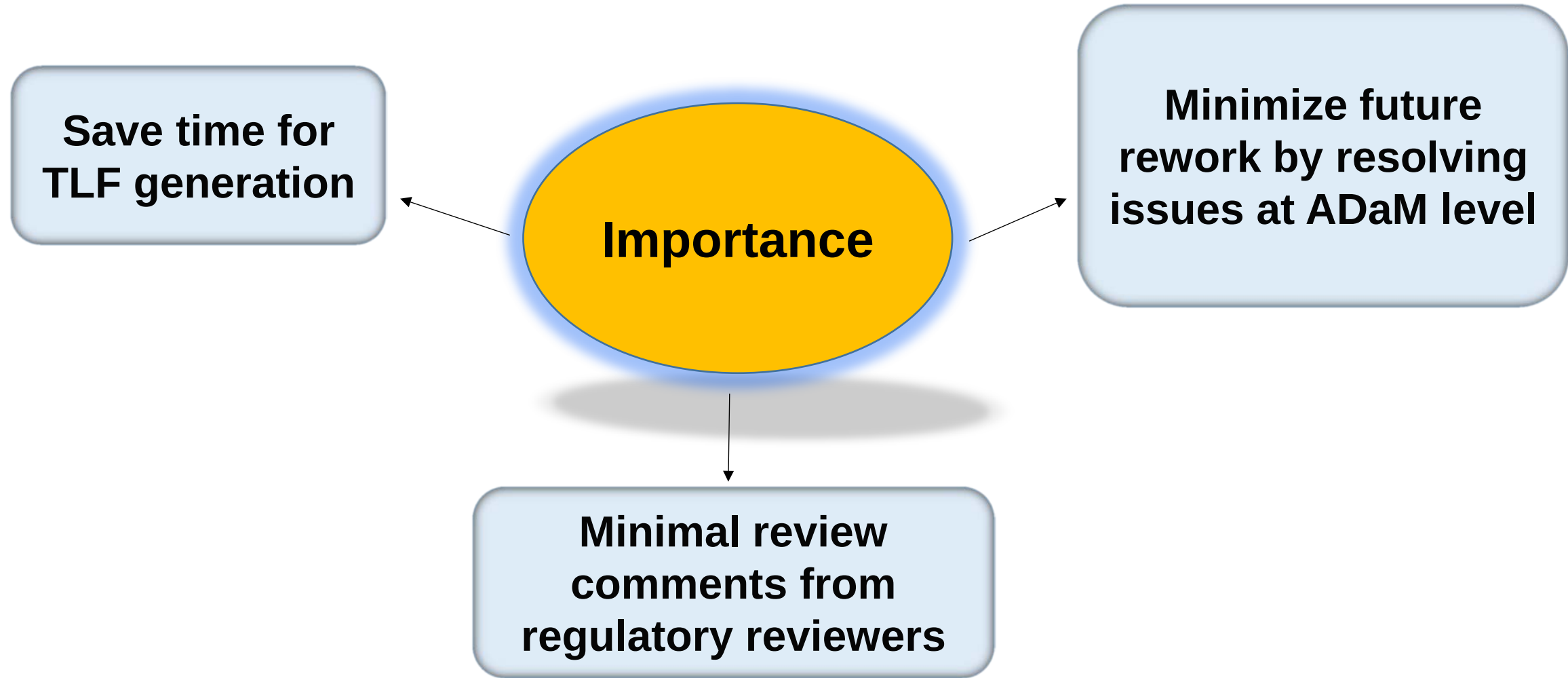
❑ QC process may involve

- Pinnacle21 Validator report
- Using standardized macro which can also generate report
- Identify independent QC resource with ADaM implementation knowledge
- Create a study specific QC checklist which can check basic ADaM principle
- Documentation of QC process and QC findings

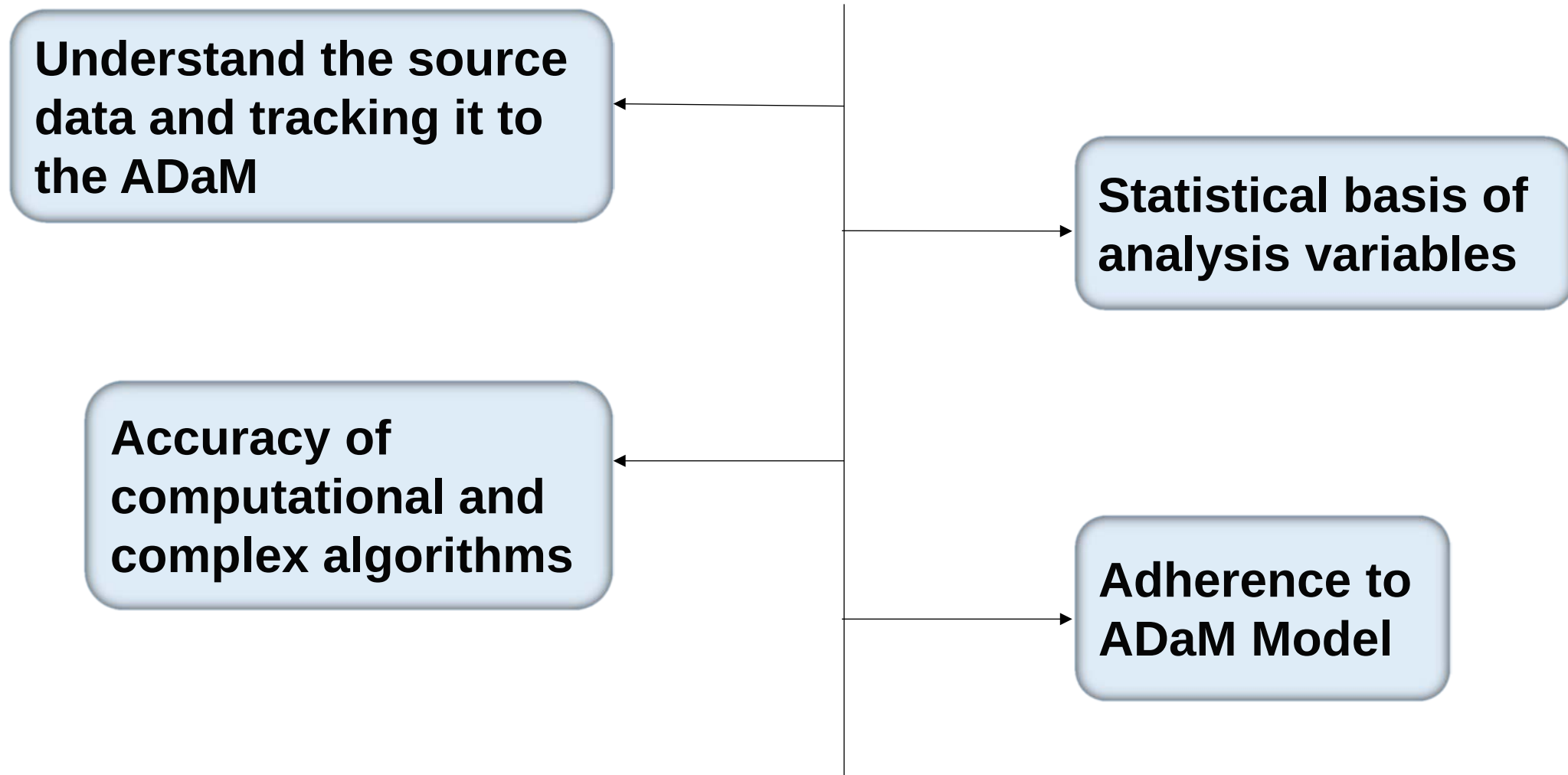
ADaM QC: Consideration

- ❑ **Traceability** is the major concern for FDA and so for QC process in organization as well
- ❑ **QC of ADaM specification against SAP**
 - Here SAP is the Master and specification is support
 - Fitness of computational algorithms with required analysis result
- ❑ **QC of ADaM dataset**
 - Pinnacle21 is just blessings to check CDISC compliance! (for FDA as well)
 - Need an additional check list except standard CDISC checks
- ❑ **QC ADaM Define file**
 - Computational Algorithms, Comments, datasets and controlled terminology definitions
 - Links to the supporting documents and links within the documents
 - Requirement of reviewer's guide

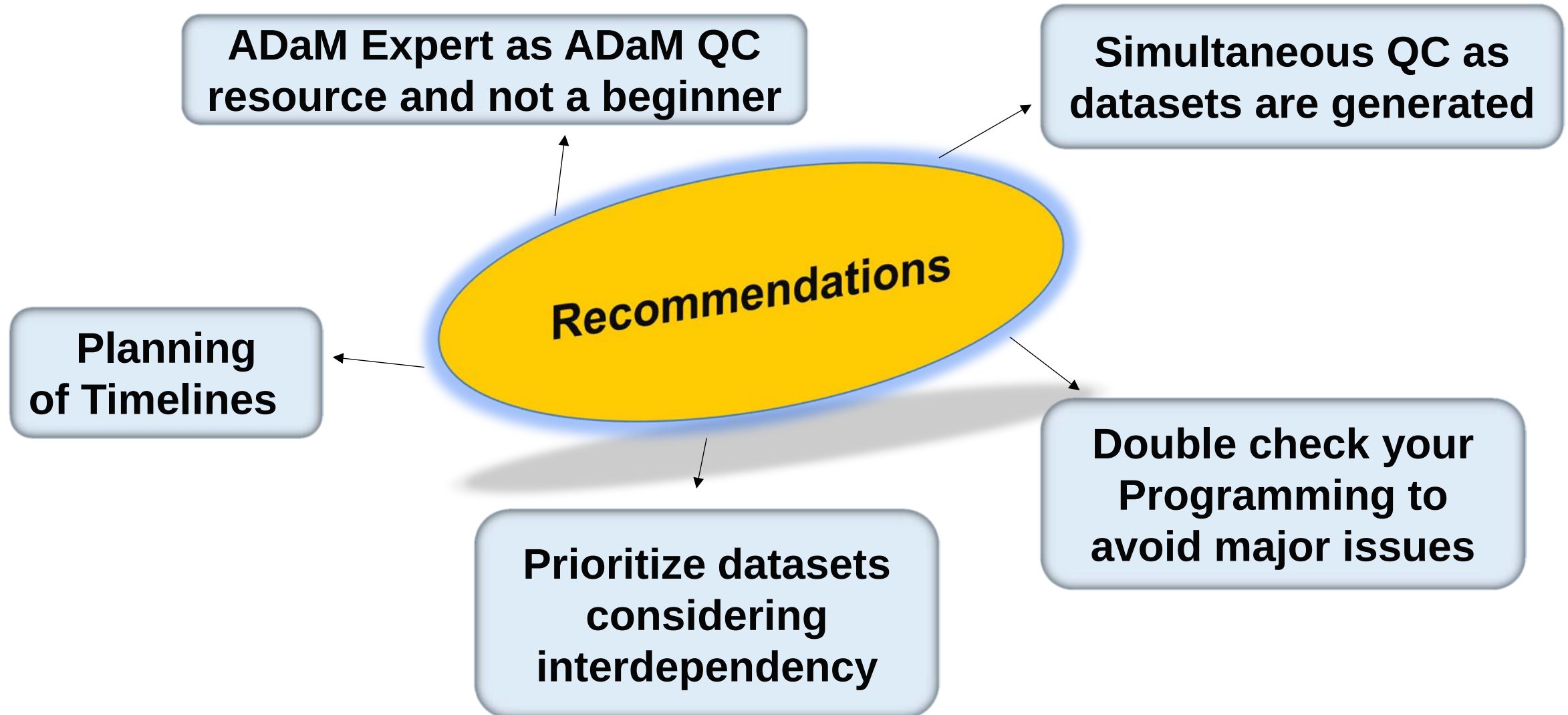
ADaM QC: Importance



ADaM QC: Challenges



ADaM QC: Recommendations



References

- ❑ Analysis Data Model Implementation Guide Version 1.1
- ❑ ADaM Structure for Occurrence Data (OCCDS) Version 1.0
- ❑ CDISC Analysis Data Model Version 2.1
- ❑ How to build ADaM from SDTM: A real case study (JIAN HUA (DANIEL) HUANG, FOREST LABORATORIES, NJ PharmaSUG2010 - Paper CD06)
- ❑ Challenges of Questionnaire Data from Collection to SDTM to ADaM and Solutions using SAS® (Karin LaPann, PRA International, Horsham, PA Terek Peterson, MBA, PRA International, Horsham, PA PharmaSUG 2014 – DS08)
- ❑ An Innovative ADaM Programming Tool for FDA Submission Xiangchen (Bob) Cui, Min Chen Vertex Pharmaceuticals, Cambridge, MA
- ❑ Challenges in Validation of ADaM data, Presented at PhUSE SDE April 18, 2013

Reach out to us @

Krupali Ladani

Senior Clinical SAS Programmer

doLoop Technologies India Pvt. Ltd.

Email: krupali.ladani@dolooptech.com

www.dolooptech.com

Dr.Sangram Parbhane

Associate Clinical SAS Programmer

doLoop Technologies India Pvt. Ltd.

Email: sangram.parbhane@dolooptech.com

www.dolooptech.com



doLoop Technologies

Business Analytics. Delivered.

Thank You !!