The Chicken, the Egg, and the ADaM: Cracking the PRE-ADSL vs ADSL Sequence

Carla Santillan - Dubravka Georgijević

PHUSE Connect EU 2025 - Standards Implementation - Paper SI03

Fortrea

AGENDA

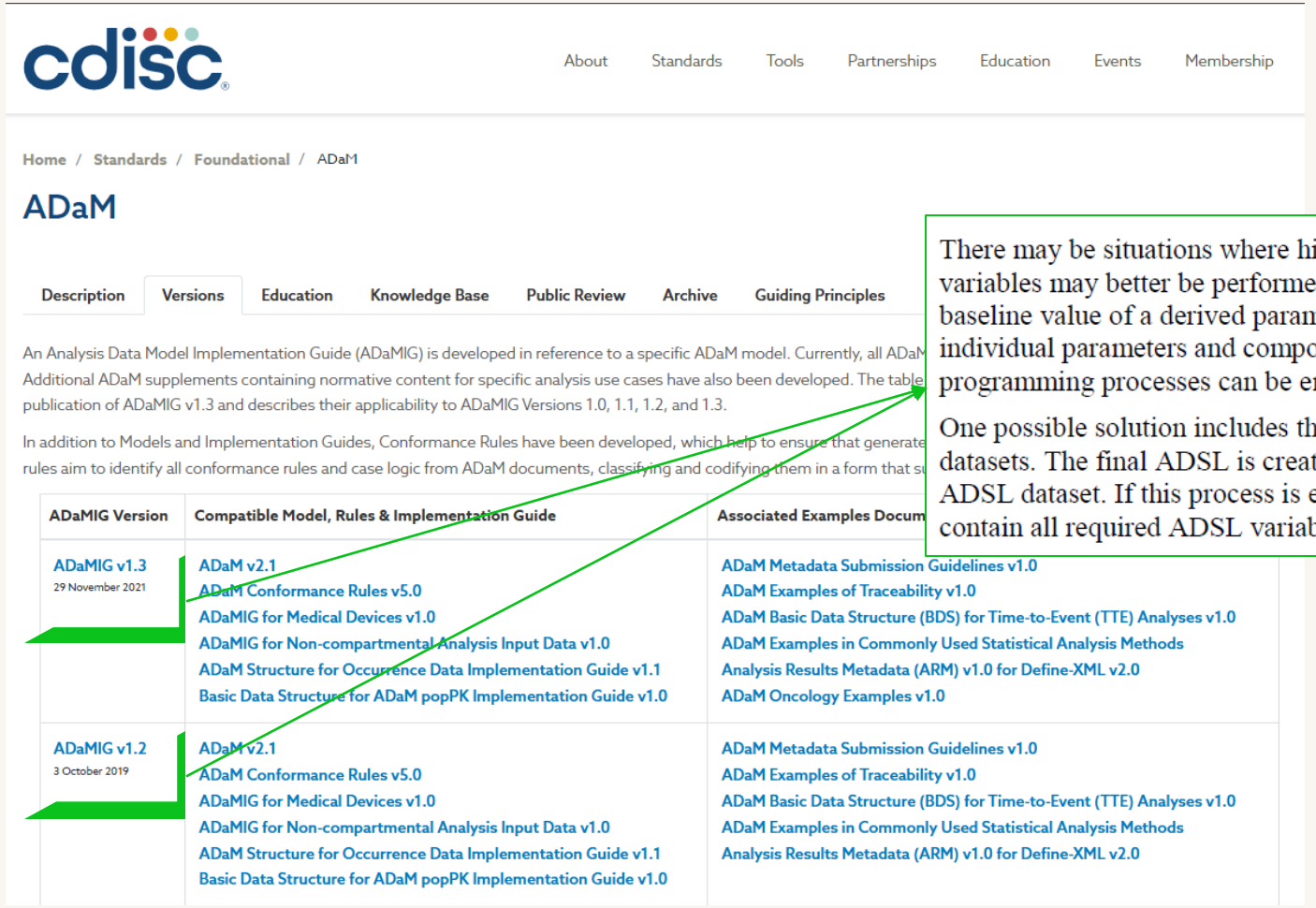
Fortrea

Documentation Available at CDISC

Fortrea

Documentation Available at CDISC

Where Is Pre-ADSL Mentioned?



There may be situations where highly derived variables are to be included in ADSL yet the derivation of these variables may better be performed in another ADaM dataset. For example, consider the analysis need to include the baseline value of a derived parameter that is a composite score based on up to seven other parameters. These individual parameters and composite parameter may best be created in a separate analysis dataset. Different programming processes can be employed to address this analysis need, and the ADaM does not dictate process.

One possible solution includes the creation of a "pre-ADSL" dataset that is used as input into subsequent analysis datasets. The final ADSL is created and collates variables as needed from any analysis dataset including the pre-ADSL dataset. If this process is employed, the pre-ADSL dataset should follow ADaM principles but may not contain all required ADSL variables.

Fortrea

Documentation Available at CDISC

Where Is Pre-ADSL Mentioned?

cdisc

About Standards Tools

Home / Standards / Therapeutic Areas / Rheumatoid Arthritis / Rheumatoid Arthri

Rheumatoid Arthritis Therapeutic Area Us

Release Information | Files & Links

Files

Rheumatoid Arthritis Therapeutic Area User Guide v1.0

rules aim to identify all conformance rules and case logic from ADaM documents, classifying and codifying them in

ADaMIG Version	Compatible Model, Rules & Implementation Guide	Associated Exa
ADaMIG v1.3 29 November 2021	ADaM v2.1 ADaM Conformance Rules v5.0 ADaMIG for Medical Devices v1.0 ADaMIG for Non-compartmental Analysis Input Data v1.0 ADaM Structure for Occurrence Data Implementation Guide v1.1 Basic Data Structure for ADaM popPK Implementation Guide v1.0	ADaM Metadat ADaM Example ADaM Basic Da ADaM Example Analysis Result ADaM Oncolog
ADaMIG v1.2 3 October 2019	ADaM v2.1 ADaM Conformance Rules v5.0 ADaMIG for Medical Devices v1.0 ADaMIG for Non-compartmental Analysis Input Data v1.0 ADaM Structure for Occurrence Data Implementation Guide v1.1 Basic Data Structure for ADaM popPK Implementation Guide v1.0	ADaM Metadat ADaM Example ADaM Basic Da ADaM Example Analysis Result

5.1.1 Additional Highly Derived Baseline Values

The ADaM model describes ADSL as one of the first analysis datasets created, since it serves as the source of important subject-level baseline values and analysis covariates that are included in other analysis datasets. However, it is not an uncommon situation where the desire is to have ADSL contain variables associated with highly derived concepts, some of which would be derived within the primary efficacy analysis datasets. If ADSL is created prior to the programming of the efficacy analysis datasets, then it poses programming challenges to include these derived measures in ADSL. This TAUG proposes a potential solution to this process issue by creating a "pre-ADSL" analysis dataset first and creating the final ADSL at the end. Version 1.2 of the ADaM IG will be adding language to allow this approach.

The following are examples of highly derived baseline variables that may need to be in ADSL. In the Source / Derivation column, the metadata shown utilizes the analysis dataset names and parameters that are described elsewhere in this document. For more details on DAS28, CDAI, and SDAI, see Section 5.4, Analysis Data for Primary Efficacy Endpoints.

CDISC Therapeutic Area Data Standards User Guide for Rheumatoid Arthritis (1.0 Provisional)

ADSL Variable Metadata: Variable Metadata for Other Derived Baseline Values

Variable Name	Variable Label	Type	Source / Derivation
DAS28CBL	Baseline DAS28-CRP	float	ADEFFRA.AVAL where PARAMCD=DAS28CRPand ABLFL=Y.
DAS28EBL	Baseline DAS28-ESR	float	ADEFFRA.AVAL where PARAMCD=DAS28ESR and ABLFL=Y.
FACITBL	Baseline FACIT-Fatigue	float	ADQS.AVAL where PARAMCD=FACIT and ABLFL=Y
HAQnnBL	Baseline HAQ-DI Version nn	float	ADHAQSUM.AVAL where PARAMCD=HAQnnDI and ABLFL=Y, where "nn" refers to the version of the HAQ questionnaire. For example, HAQ02DI indicates version 2 of the HAQ questionnaire.
SF36MCBL	Baseline SF36 Mental Component	float	ADQS.AVAL where PARAMCD=SF36MC and ABLFL=Y
SF36PCBL	Baseline SF36 Physical Component	float	ADQS.AVAL where PARAMCD=SF36PC and ABLFL=Y
SJC28BL	Baseline Swollen Joint Count 28	integer	ADJNTSUM.AVAL where PARAMCD=SJC28 and ABLFL=Y
TJC28BL	Baseline Tender Joint Count 28	integer	ADJNTSUM.AVAL where PARAMCD=TJC28 and ABLFL=Y
SJC66BL	Baseline Swollen Joint Count 66	integer	ADJNTSUM.AVAL where PARAMCD=SJC66 and ABLFL=Y
TJC68BL	Baseline Tender Joint Count 68	integer	ADJNTSUM.AVAL where PARAMCD=TJC68 and ABLFL=Y
CDAIBL	Baseline CDAI	float	ADEFFRA.AVAL where PARAMCD=CDAI and ABLFL=Y.
SDAIBL	Baseline SDAI	float	ADEFFRA.AVAL where PARAMCD=SDAI and ABLFL=Y.

Fortrea

Documentation Available at CDISC

Fundamentals of the ADaM - Standard (ADaM IG V1.3)

- ADaM datasets and associated metadata must clearly and unambiguously communicate the content and source of the datasets supporting the statistical analyses performed in a clinical study.
- ADaM datasets contain both source

Described in Define as Predecessor data

and derived data

Described in Define with algorithm

- It should be clear how the variables and observations were derived
- ADaM datasets and associated metadata must provide traceability to show the source or derivation of a value or a variable (i.e., the data's lineage or relationship between an element and its predecessor(s)). Traceability is built by clearly establishing the path between an element and its immediate predecessor.

Fortrea

Documentation Available at CDISC

ADSL, From ADaM IG V1.3 Section 2.3.1

CDISC Analysis Data Model Implementation Guide (1.3 Final)

although the ADSL contains subject-level variables, there is no requirement that every ADSL variable be present in other analysis datasets

any ADSL variable needed to enable analysis (e.g., statistical model covariates, population flags, subgrouping variables) should appear in the analysis dataset

A variable that is present in both ADSL and any other ADaM dataset must have the same values, type, and label.

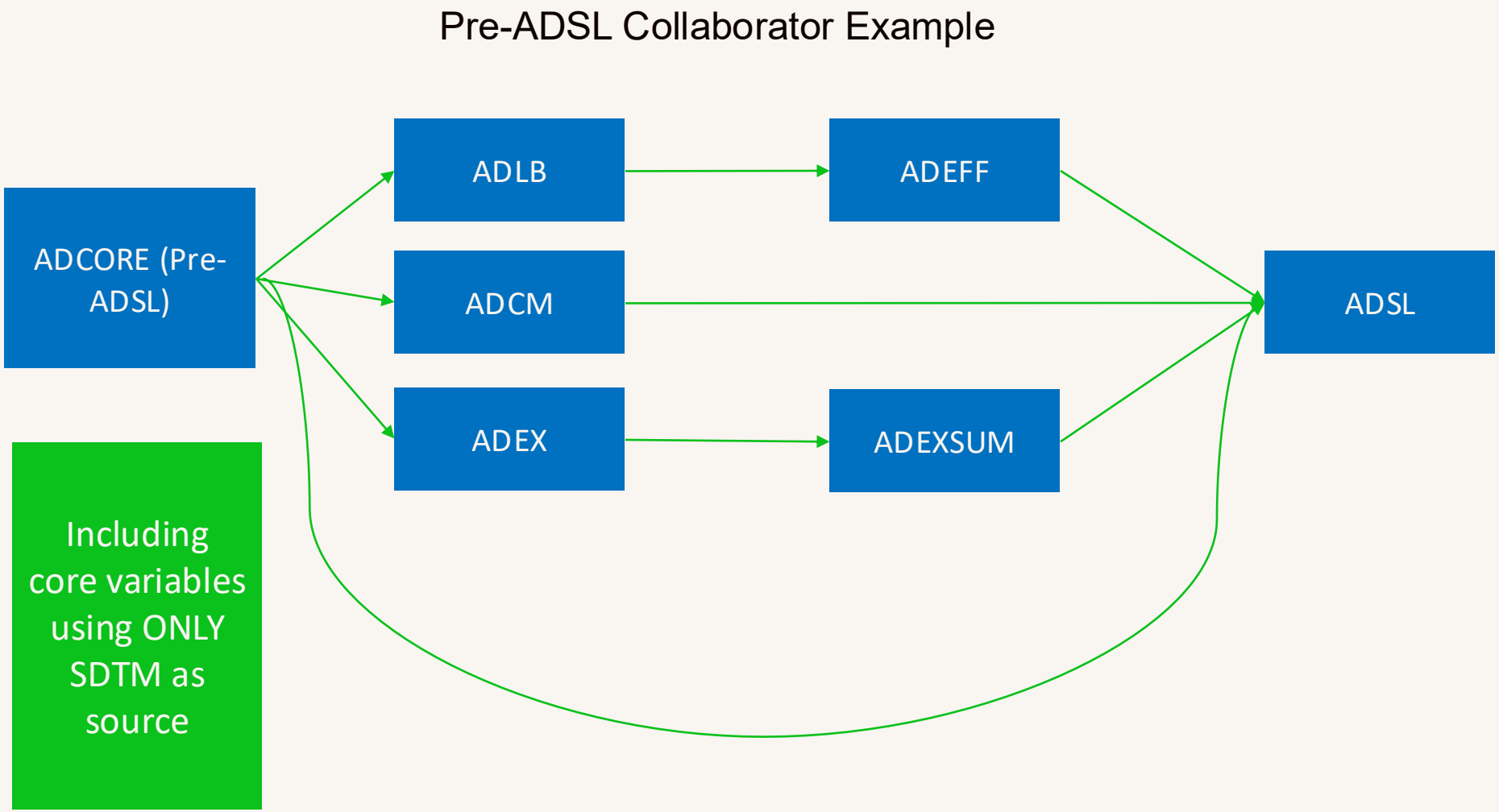
Note that the FDA sdTCG requests that "core" subject-level variables be present in all analysis datasets.

Fortrea

Pre-ADSL Collaborator Implementation

Fortrea

Pre-ADSL Collaborator Implementation



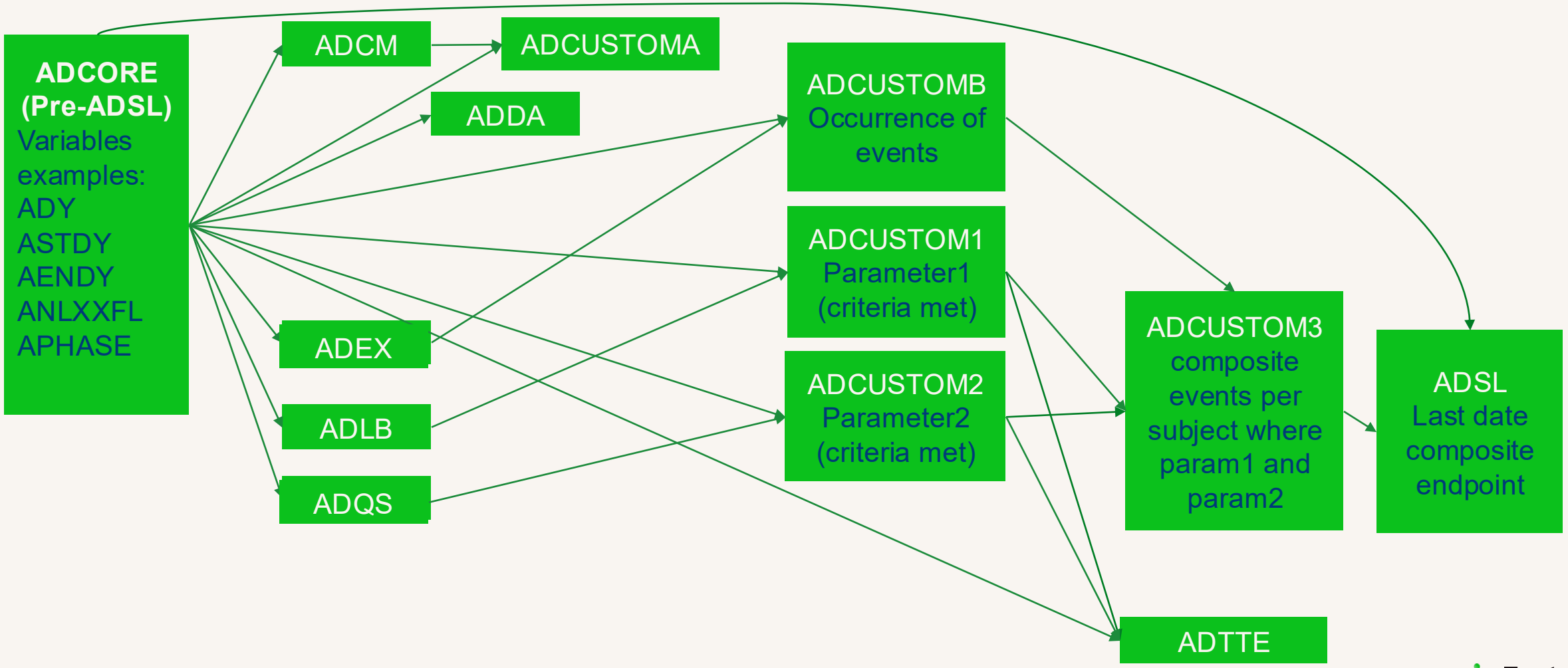
Fortrea

Implementation Example

Fortrea

Implementation Example

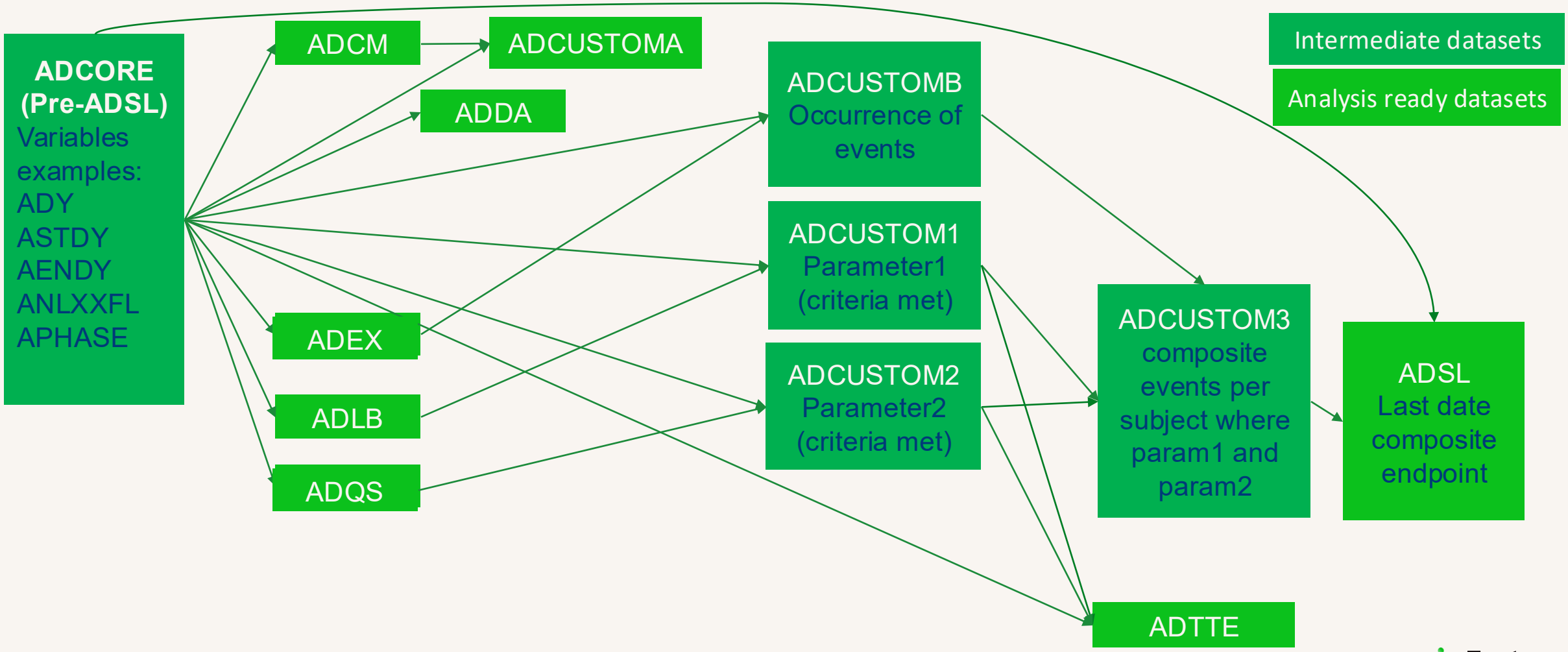
First Example: ***Complex Calculations Composite Endpoint Consisting Of Multiple Event Types***



Fortrea

Implementation Example

First Example: ***Complex Calculations Composite Endpoint Consisting Of Multiple Event Types***



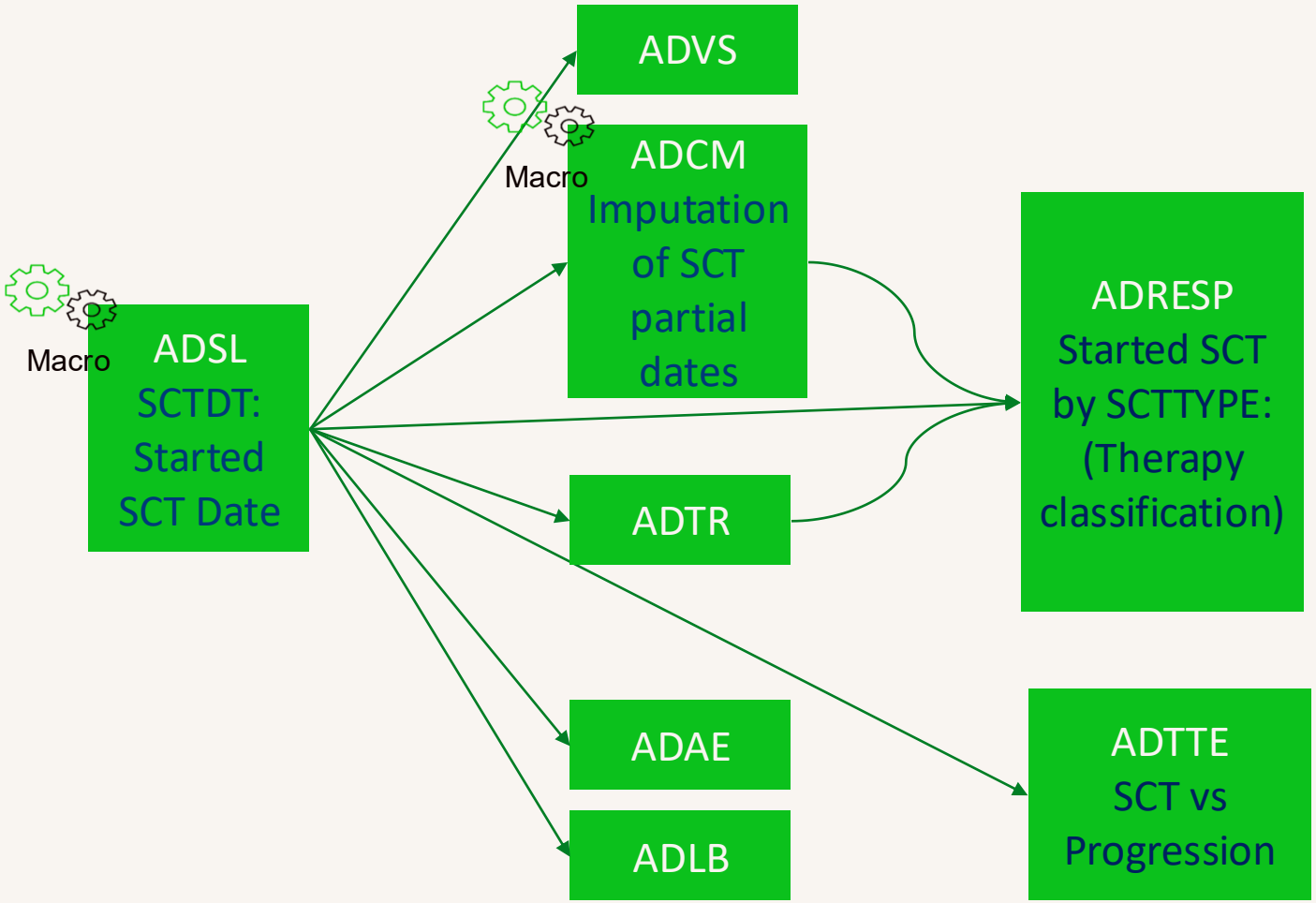
Fortrea

Potential Situations Benefiting from this Approach

Fortrea

Potential Situations Benefiting From This Approach

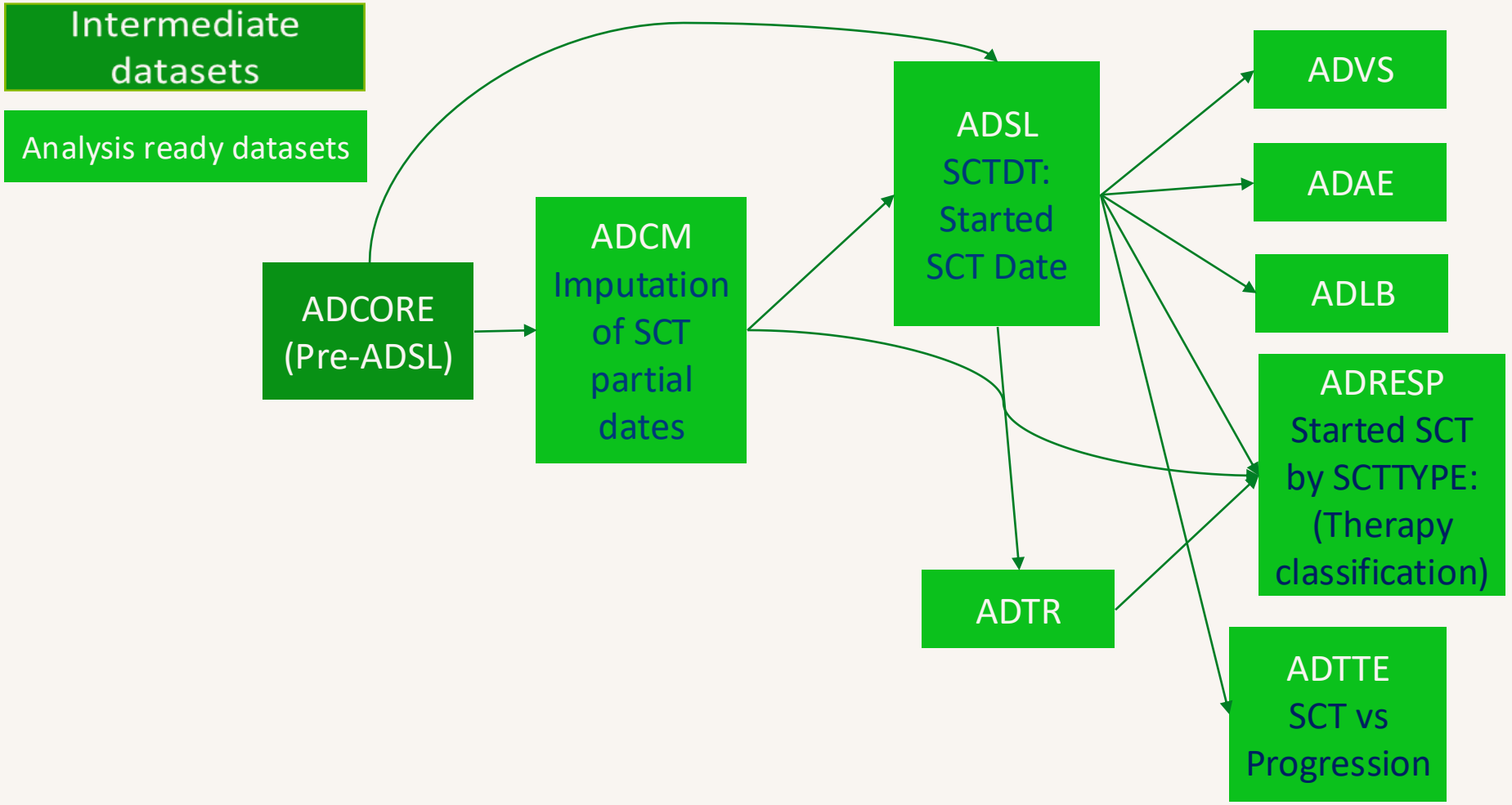
SECOND Example (A): *Handling Subsequent Anticancer Therapy (SCT) Variables*



Fortrea

Potential Situations Benefiting From This Approach

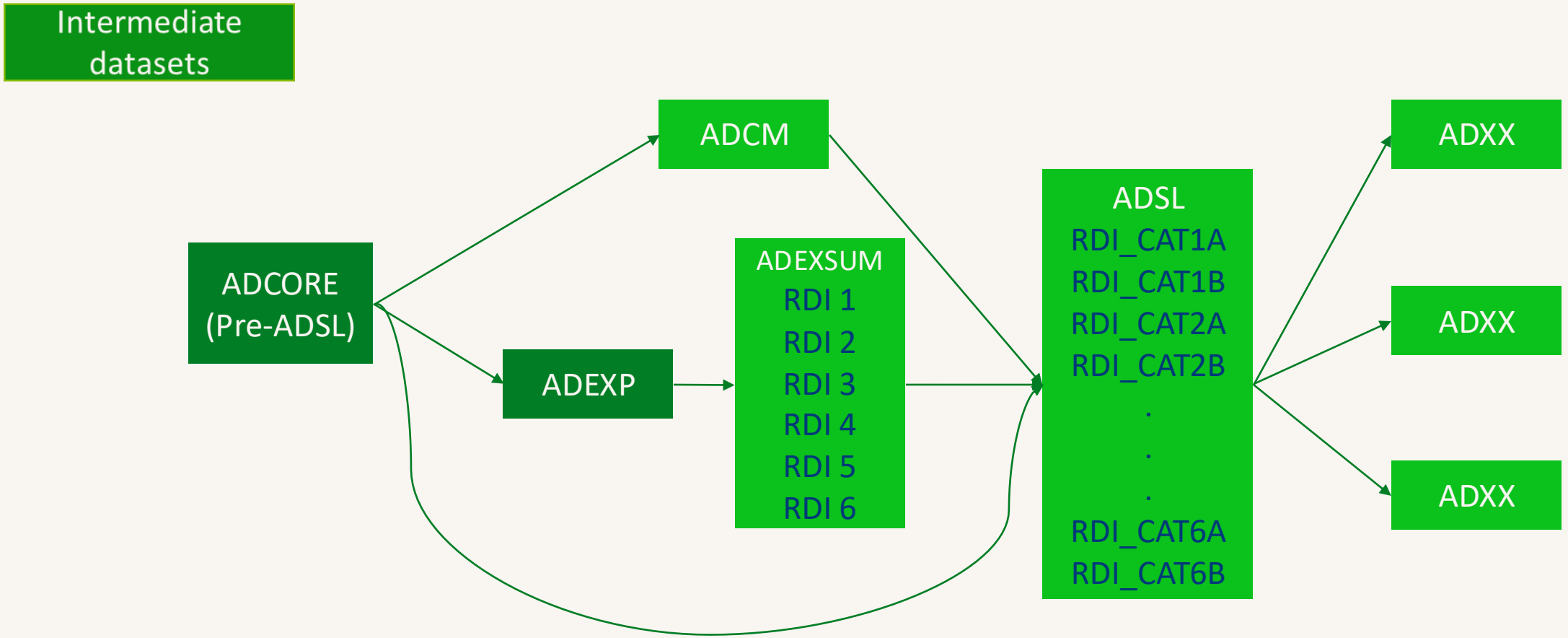
SECOND Example (B): *Handling Subsequent Anticancer Therapy (SCT) Variables*



Fortrea

Potential Situations Benefiting From This Approach

Third Example: *Handling Relative Dose Intensity (RDI) And Related Variables*



Fortrea

Documentation Best Practices

Fortrea

Documentation Best Practices

Define.xml

Derivation/Comment Field

- how the variable was derived
- derivation was performed in a pre-ADSL dataset
- source SDTM domains used (e.g., EX, QS, FA)
- logic applied (e.g., imputation, aggregation)
- a reference to programming specifications or SAP if applicable

Dataset-Level Documentation

at the ADSL dataset description, you may include a note like:

"*Some variables in ADSL were derived using a pre-ADSL staging dataset that followed ADaM principles but did not include all required ADSL variables*"

Documentation Best Practices

ADRG

Section: “Data Derivations and Traceability”

This section should describe how key variables in ADSL were derived.

If a pre-ADSL dataset was used to stage complex derivations (e.g., treatment start dates, baseline scores, population flags), it should be explicitly mentioned here.

Section: “Analysis Dataset Overview”

When describing the ADSL dataset, a note can be included that a pre-ADSL dataset was used in its construction. Mentioning the purpose of the pre-ADSL (e.g., staging derivations, imputations, or complex logic).

Section: “Programming Notes or Assumptions”

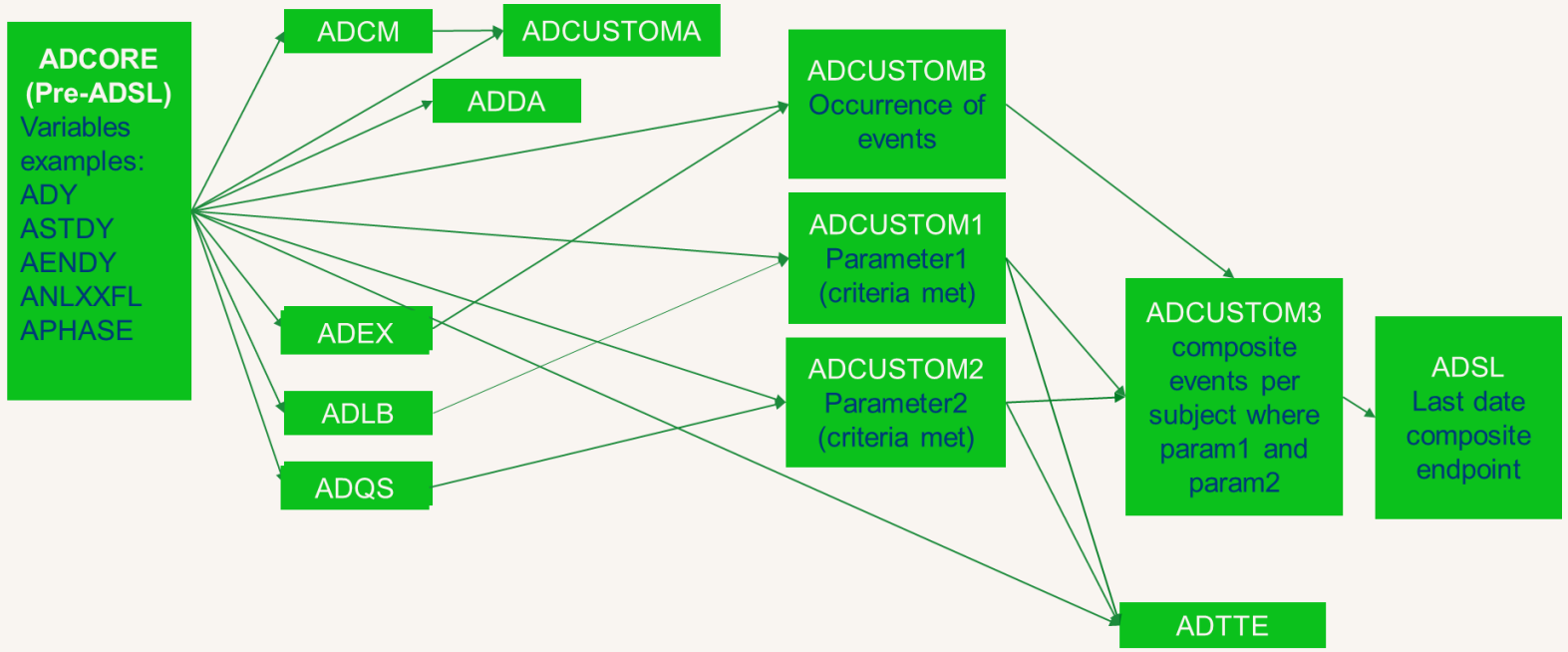
If any assumptions or imputation rules were applied in pre-ADSL (e.g., handling partial dates), that were not documented clearly anywhere else, e.g. SAP then document them here.

Comparison of Pre-ADSL Approaches

Fortrea

Comparison of Pre-ADSL approaches

First Example: *Complex Calculations Composite Endpoint Consisting Of Multiple Event Types*

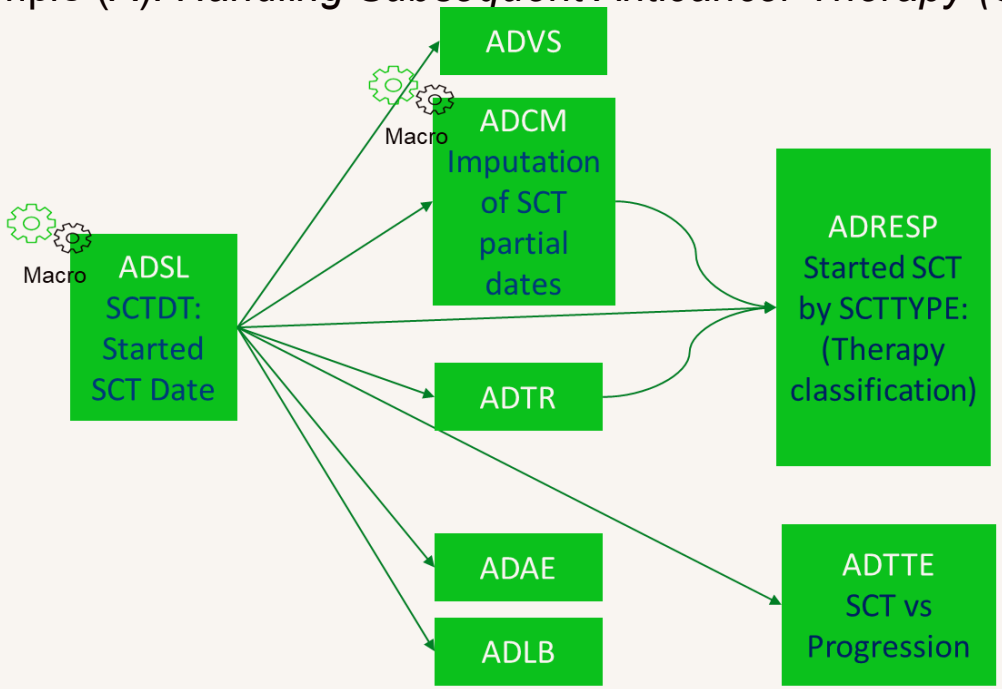


Approach	Pros	Cons
Final derivation in ADSL after other domains	- Improved traceability through modular derivations	- Requires careful management of interdependencies - Changes to final derivation must be propagated backward

Fortrea

Comparison of Pre-ADSL approaches

SECOND Example (A): *Handling Subsequent Anticancer Therapy (SCT) Variables*

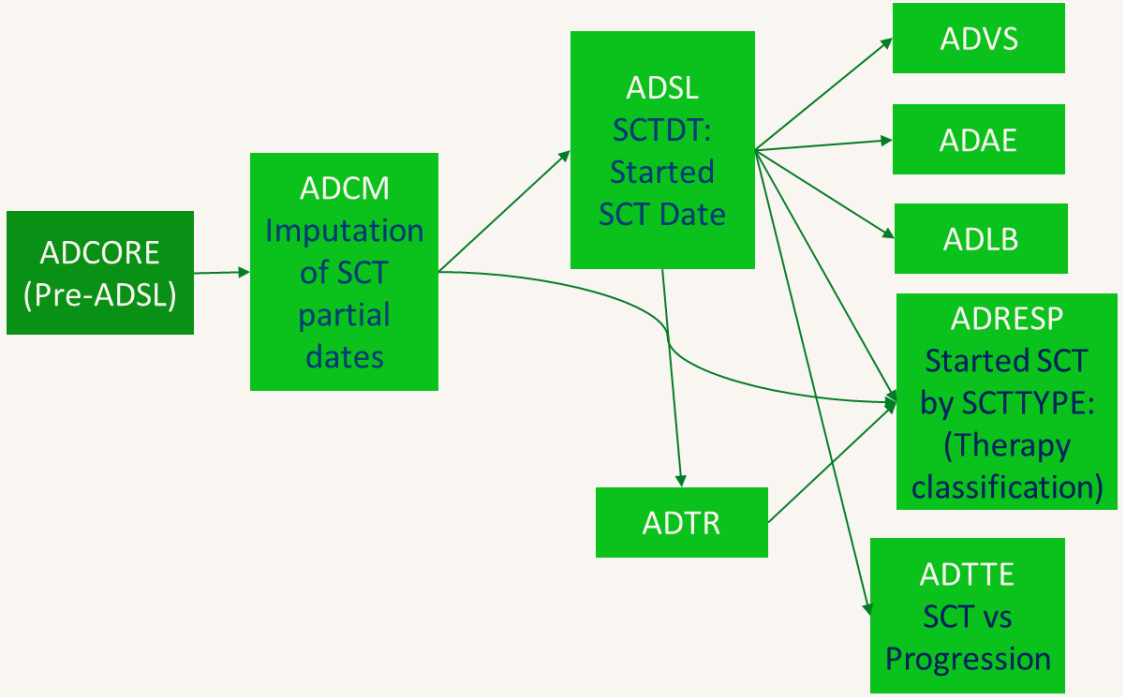


Approach	Pros	Cons
Macro applied in both ADSL and ADCM	- Efficient for simple derivations - Avoids circular logic - Core variables well documented in ADSL - No additional specifications needed	- Requires consistency checks across domains with duplicated logic

Fortrea

Comparison of Pre-ADSL approaches

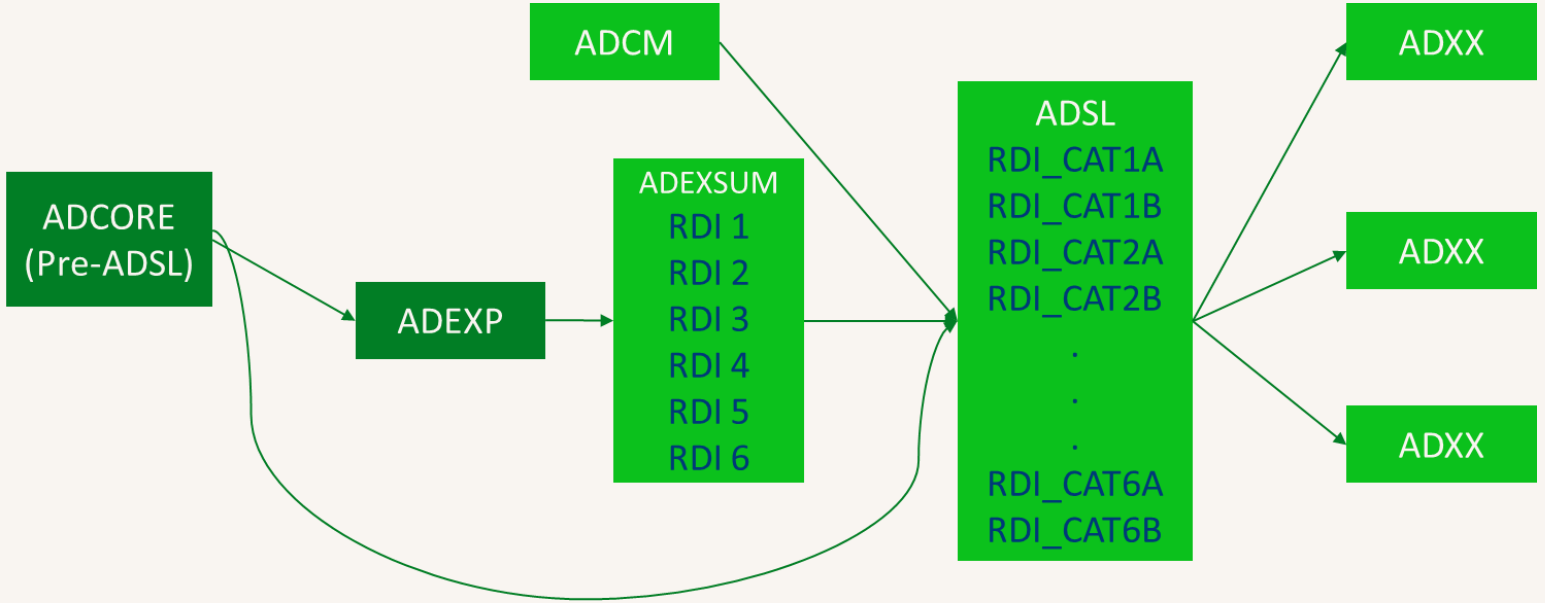
SECOND Example (B): *Handling Subsequent Anticancer Therapy (SCT) Variables*



Approach	Pros	Cons
Pre-ADSL used for staging simple logic	- Core variables documented once in pre-ADSL - Automatic inclusion of core variables downstream domains	- Additional documentation required for pre-ADSL - Limited benefit for simple derivations?

Comparison of Pre-ADSL approaches

Third Example: *Handling Relative Dose Intensity (RDI) And Related Variables*



Approach	Pros	Cons
Pre-ADSL used for staging complex logic, many variables	- Core variables derived once and reused - Simplifies updates when logic changes - Enhances traceability	- Requires careful documentation of pre-ADSL and other intermediate datasets.

Comparison of Pre-ADSL approaches

Balancing Clarity And Effort In Intermediate Dataset Use

- Where should we draw the line?
- Does introducing an intermediate dataset add unnecessary complexity?

The decision depends on the complexity of the derivations and the difficulty of tracing them directly within ADSL. When derivations are intricate or span multiple domains, an intermediate dataset - often referred to as *pre-ADSL* - can be a valuable tool.

Thank You!

Questions?

Carla Santillan - Dubravka Georgijević
Fortrea
PHUSE EU 2025

Fortrea

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the authors at:

Carla Santillan

Fortrea Belgium SRL

Avenue du Bourget 50

Brussels 1130

Belgium

- carla.santillan@fortrea.com

Dubravka Georgijević

Fortrea Netherlands BV

Schipholweg, 111 2316 XC

Leiden, The Netherlands

- dubravka.georgijevic@fortrea.com

Fortrea