

# CLARIFYING THE LEGACYDATA CONVERSION PLAN& SDSP

# INTRODUCTION

Study & Analysis Data Reviewer's Guides and the Study Data Standardization Plan. These documents describe routine situations where FDA endorsed data standards are deployed at the time a study is initiated; additional support is needed when the provenance of the data is not as straightforward.

The FDA's Study Data Technical Conformance Guide calls out for the need to provide a Legacy Data Conversion Plan & Report when legacy data is the source of deliverables based on FDA endorsed data standards, but it is not very clear as to when you must provide one. This presentation will leverage recent PhUSE efforts to develop a template and completion guidelines for this document to clarify when it must be provided and introduce the concept of the Study Data Traceability Guide.

# STUDY DATA STANDARDIZATION PLAN

The Study Data Standardization Plan (SDSP) is described in the FDA binding guidance *Providing Regulatory Submissions In Electronic Format – Standardized Study Data* and the *Study Data Technical Conformance Guide* (Study Data TCG). The SDSP is the beginning of the traceability story for a compound. The document is a high level summary of the exchange and terminology standards for nonclinical and clinical studies submitted as part of the IND or NDA filed under an indication. This is an opportunity for FDA to look at the standards for planned, ongoing, and completed studies early in the development of the compound development cycle.

# LEGACY DATA CONVERSION PLAN AND REPORT

The Legacy Data Conversion Plan & Report (LDCP) is described in section 8.3.2 of the Study Data Technical Conformance Guide which is associated with the FDA binding guidance [Providing Regulatory Submissions In Electronic Format — Standardized Study Data](#). Traceability is important, especially to a regulatory reviewer, when collected study data that is non-standardized data is converted to standardized data. This will aid the reviewer to follow the data from collection to analysis when the analysis was originally done with legacy analysis data. Non-standardized data and legacy data are used interchangeably to refer to data that does not conform to standards currently supported in the Data Standards Catalog. Standardized data, on the other hand, refers to data that conforms to standards currently supported in the Data Standards Catalog. While sponsors work through the period of providing standardized study data based on the rule when the requirement begins, legacy data conversions will occur for studies in which legacy analysis data was used.

The Legacy Data Conversion Plan was a byproduct of an FDA project initiated in the early part of this decade. The FDA engaged with an organization to produce select SDTM domain datasets & ADaM analysis datasets from data that had been provided to the agency in support of regulatory submissions. They focused on a number of studies from a number of vendors within a single therapeutic area/ indication to support a proof of concept that SDTM & ADaM could be successfully used to analyze data across studies and Sponsors. This activity primarily involved creating SDTM & ADaM from legacy tabulation data guided by legacy analysis assets that were present. As the agency sought to confirm results and assess usability they ran into issues where existing tools to record data migration decisions did not adequately explain all they needed to know in order to perform their work. From this a process and template was developed to produce what is now the predecessor to the Legacy Data Conversion Plan to document detailed information about conversion decisions.

The Legacy Data Conversion Plan & Report should be added to the appropriate reviewer's guide for clinical studies (cSDRG, ADRG). The Technical Conformance Guide describes three situations in which the LDCP should be included in the reviewer's guide.

# LEGACY TABULATION DATA CONVERTED TO SDTM

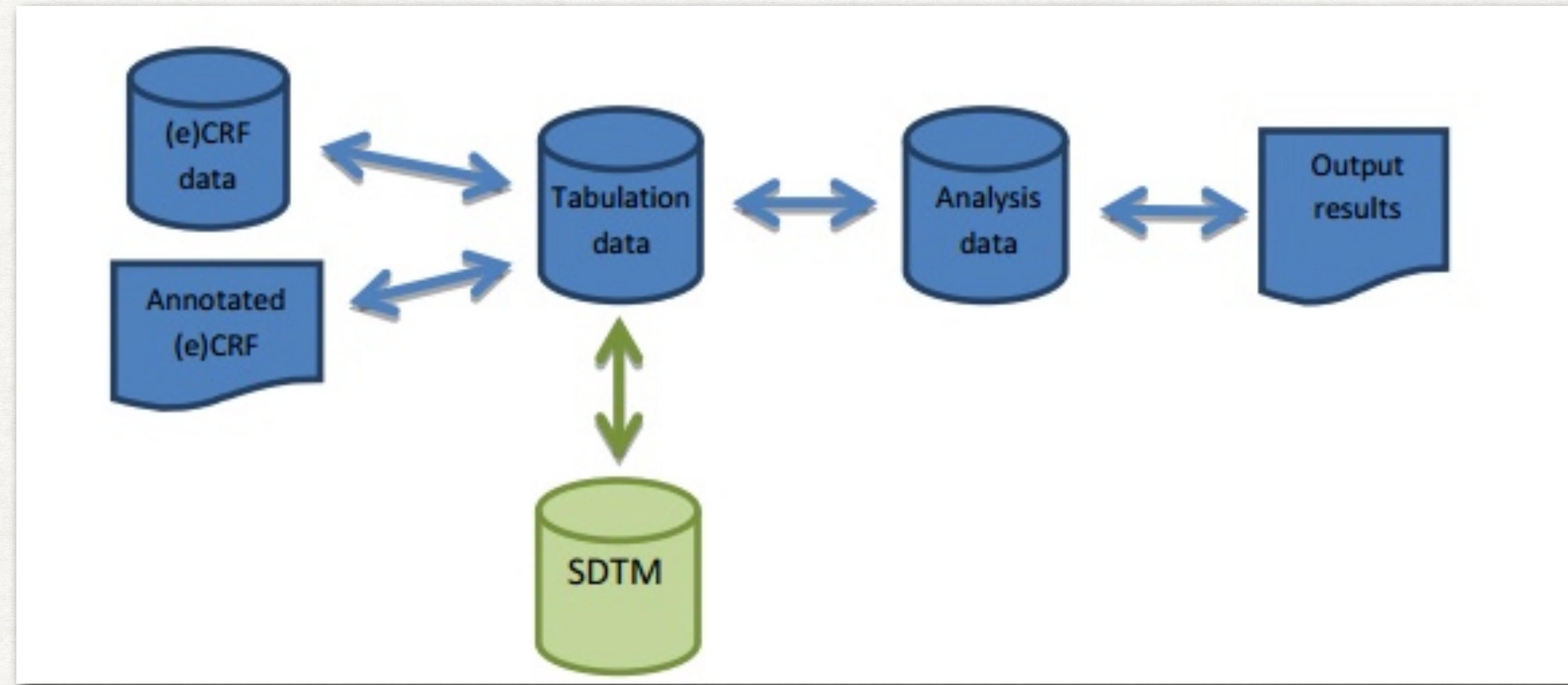


Figure1- LegacyData Conversion to SDTM

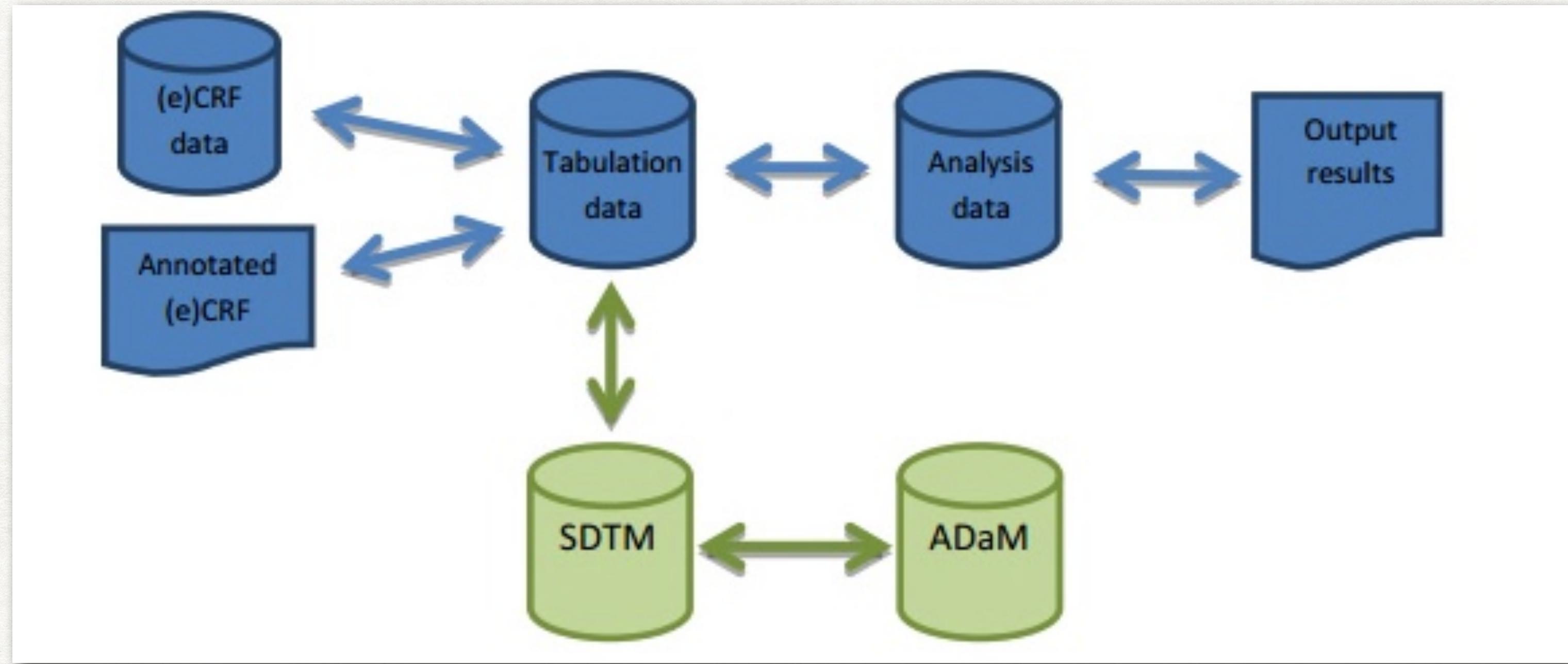
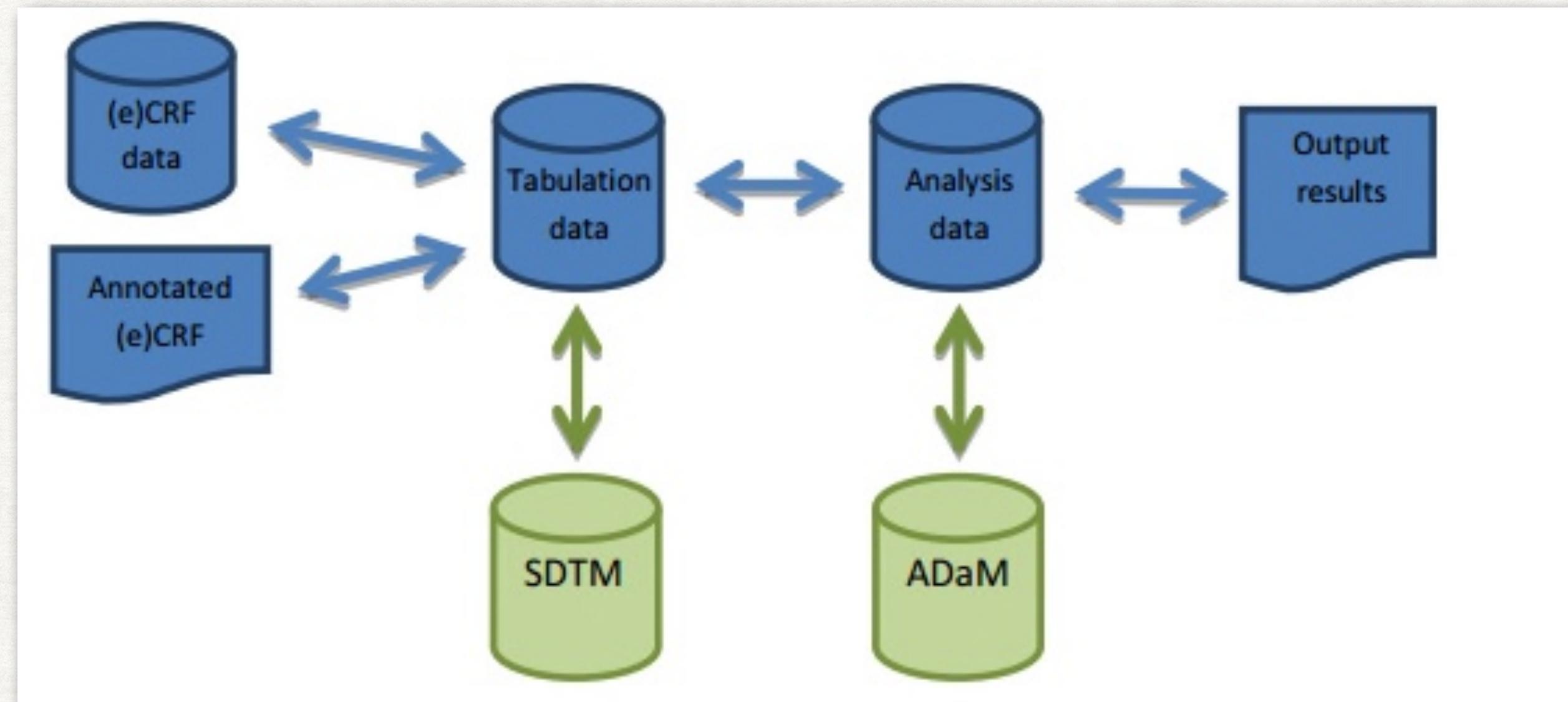


Figure2- Legacy Tabulation Data Converted to SDTM, ADaM Created from SDTM



**Figure3- Legacy Tabulation & Analysis Data Converted to SDTM & ADaM in Parallel**

# WHEN THE LEGACY DATA CONVERSION PLAN SHOULD BE CONSIDERED

When a version of standardized data is no longer supported per the Data Standards Catalog (e.g. SDTM IG 3.1.1) and is up-versioned to a version that is supported per the catalog (e.g. SDTMIG v3.1.3, v3.2).

Every conversion of legacy data to standardized data is different. The decision falls to the sponsor to determine when the LDCP should be included in a reviewer's guide. If the reviewer cannot tell the story of the data via the eCRF, the define document, and the reviewer's guide, then the extra information per the LDCP should be included.

1. The data definition file (define.xml) cannot adequately explain where data present in SDTM came from
2. The aCRF to SDTM is not clear enough on how variables were transformed from how they were collected to how they were represented in SDTM
3. The reviewer's guide cannot adequately document the migration strategy and results without using the Legacy Data Conversion Plan & Report section of the guide to convey this information

## WHAT ELSE SHOULD ACCOMPANY THE LEGACY DATA CONVERSION PLAN

- 1.Two versions of the aCRF: One version that traces the collection field to the legacy variables and one that links the collection field to the correct SDTM variable
2. Legacy tabulation and analysis datasets may need to be provided in addition to the standardized data to support traceability to the standardized data, legacy CSR, and/ or TLFs

# CONCLUSION

There is discussion that should take place within the sponsor organization to determine what trials should be included based on consultation with the compound team and regulatory group. FDA does not want to be surprised as it may be too late in the development of the compound to make needed changes. It is important to be in alignment with FDA and the standards that are supported in the FDA's Data Standards Catalog. As sponsors move into providing standardized data (as opposed to legacy analysis data) and there is a version of SDTM that is no longer supported per the Data Standards Catalog, the information provided in the SDSP is an early look at the standards that are being/will be applied to studies within a compound/indication. The standards provided in the SDSP align with the information that is in the catalog. Information about pooling of data is also included in the SDSP, when it is known. The SDSP is a living document and should be updated appropriately prior to the submission. It can be submitted as a stand-alone document or should be included in the General Investigational Plan when it is part of an IND application.

The SDSP is also useful when an NDA has been submitted and the compound has a new line indication. The SDSP is the beginning of the traceability story for the new indication. Information about the new indication is included in the SDSP with a reference to the original NDA.

“

THANKS

”