

Conversion of Prospective Cohort Study Data to OMOP CDM for Advanced Breast Cancer: Strengths, Limitations, and Future Directions

Authors Jimmy Toulas, Christine Rossin, Alexa Parliyan

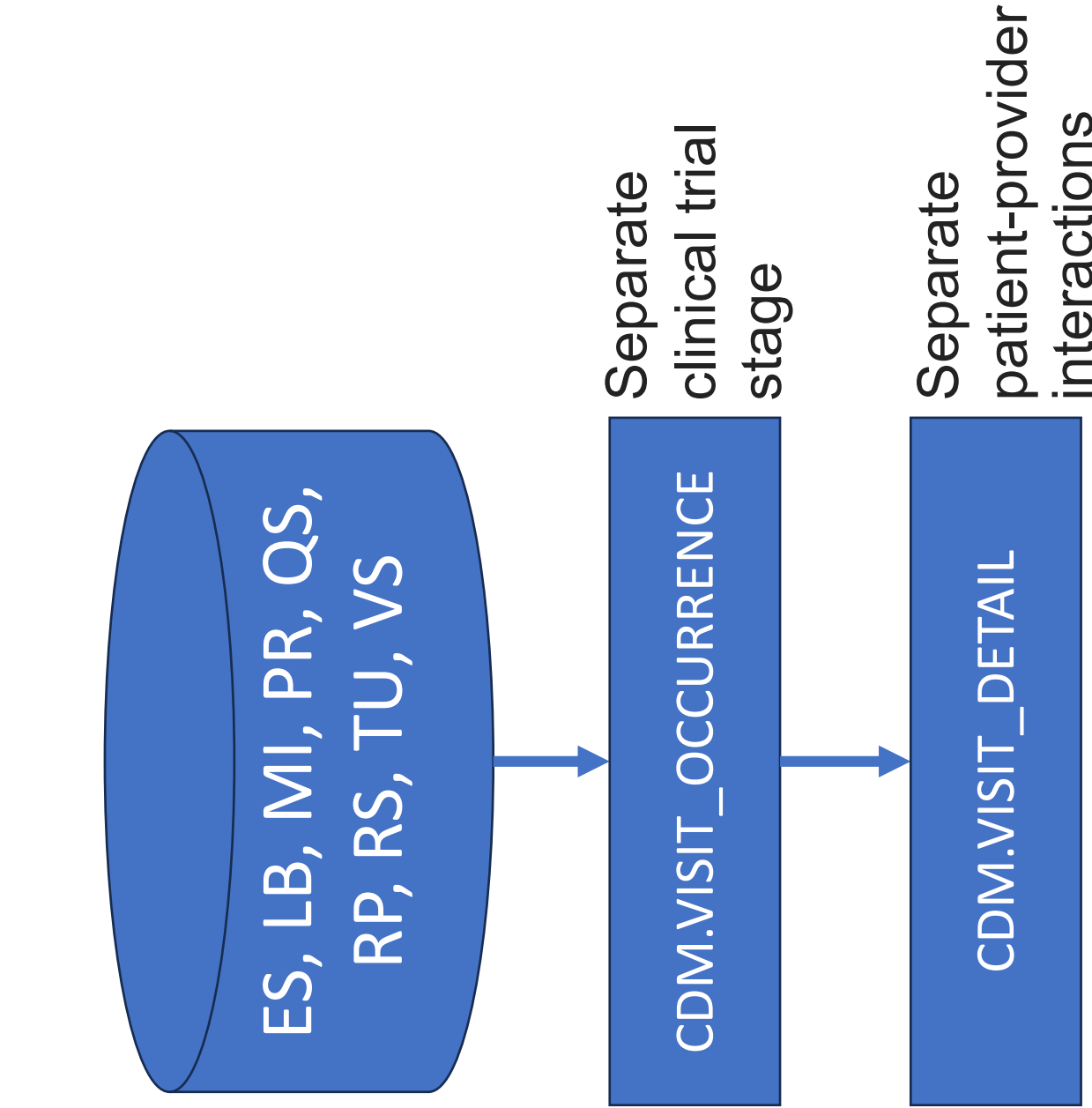
Introduction

This poster focuses on the conversion of data from a prospective cohort study on advanced breast cancer patients, from SDTM to the OMOP Common Data Model (CDM), for integration into OPTIMA, a European federated data network. The diversity of the federated nodes, as well as data privacy restrictions in EU countries raised the need for all data partners to standardize their data. Disparate data sources were made available for standardized analyses, by being converted to the CDM. Typically, the conversion of EHR and Claims data is more common in the industry, however in this case we examine the feasibility and quality of mapping SDTM data to the CDM.

Study Features of Collected Data

- Patient informed consent
- Inclusion/Exclusion criteria to the study
- Patient demographics
- Insurance status Prior antineoplastic therapies
- Comorbidities
- Tumor characteristics, TNM staging, localization of metastases
- Menopausal status (women)
- Vital signs, weight, height (only baseline)
- Clinically relevant prior/concomitant medication
- ECOG or Karnofsky performance status
- Clinically relevant laboratory parameters
- Adverse event reporting
- Treatment and reason for EOT
- Tumor response under and at the end of first/second-line treatment
- Date and methods of assessment of disease progression
- Duration (staff and end date) and of next-line therapies
- Physician's treatment decisions
- Physician's routine assessment of patient's overall health status
- Physician's assessment of patient management
- Handout of patient questionnaires
- Biomarkers/(liquid) (re)biopsies
- Family history of gBRCA1/2-related cancer(s)
- Date and reason for EOS
- Overall survival status

Example: Patient Visits Logic



Acknowledgements

David Eleuteri, Sharon Niedecken, Pfizer OPTIMA team, Odysseus Data Services Inc.

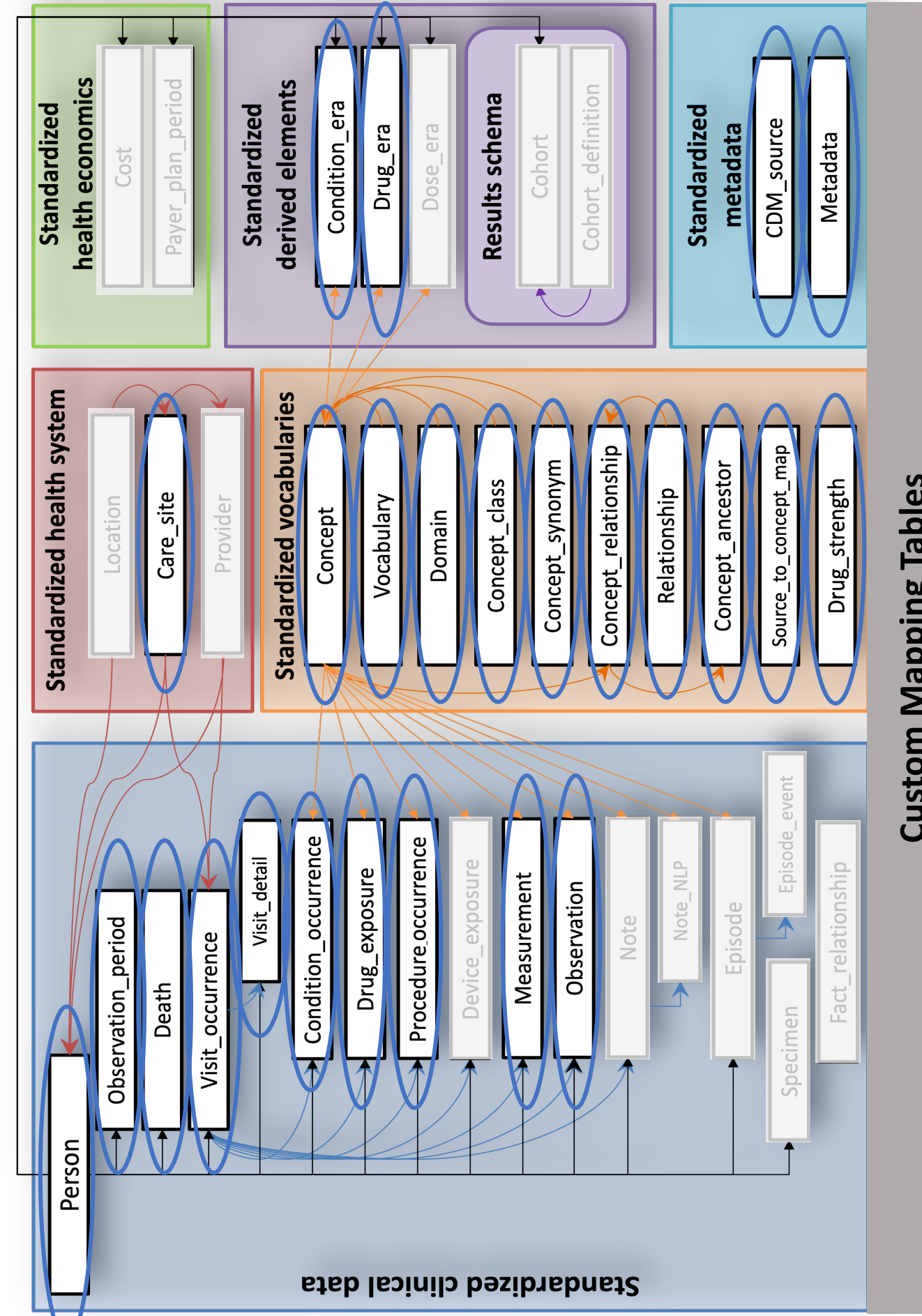
References

The book of OHDSI – Data Standardization, OHDSI, [Data Standardization – OHDSI](#) , Accessed online March 2024

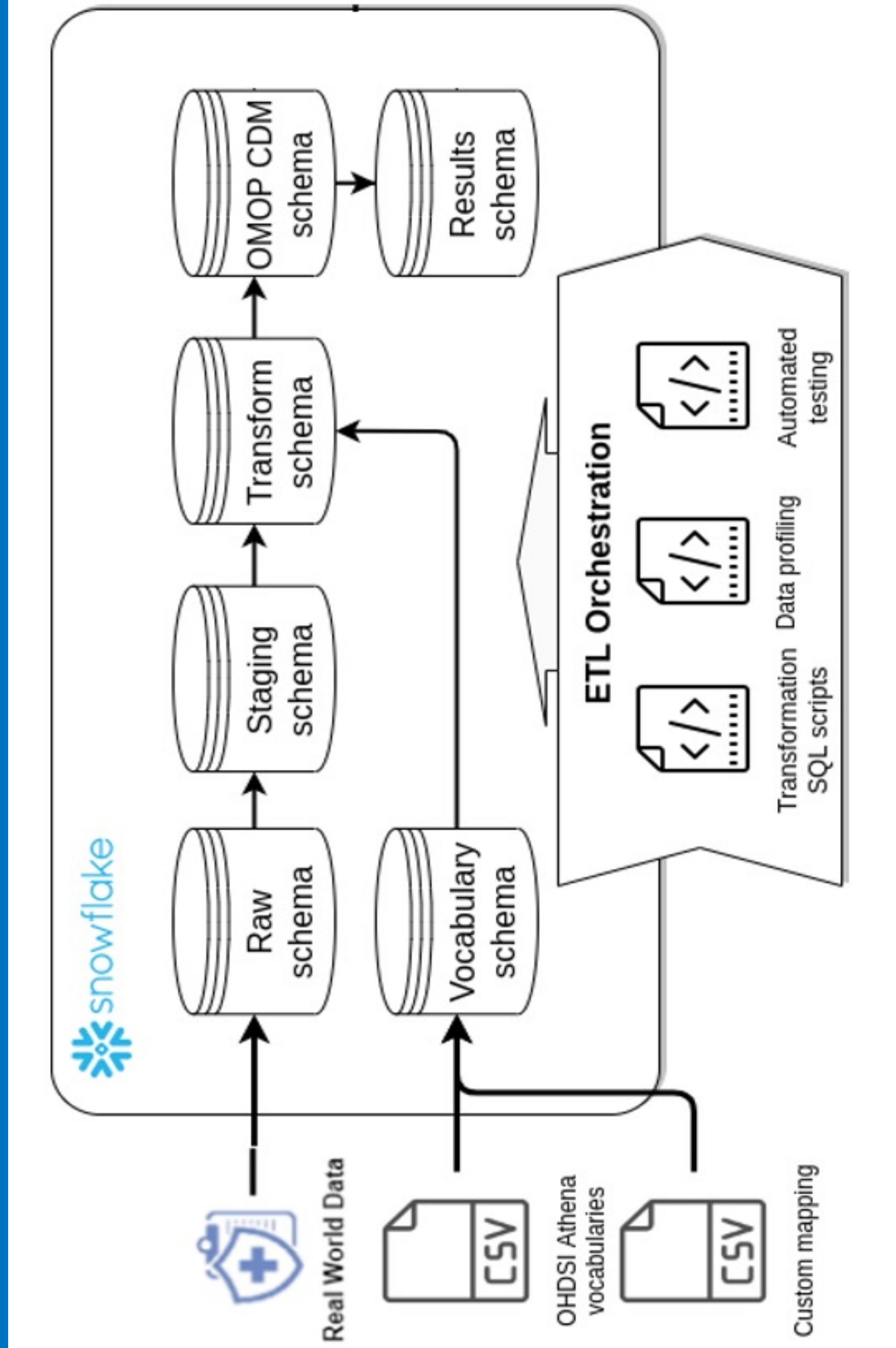
Contact Information

Jimmy Toulas (dimitrios.toulasgourgiotis@pfizer.com), Christine Rossin (christine.rossin@pfizer.com), Alexa Parliyan (alexa.parliyan@pfizer.com)

OMOP CDM Tables Used



Architecture

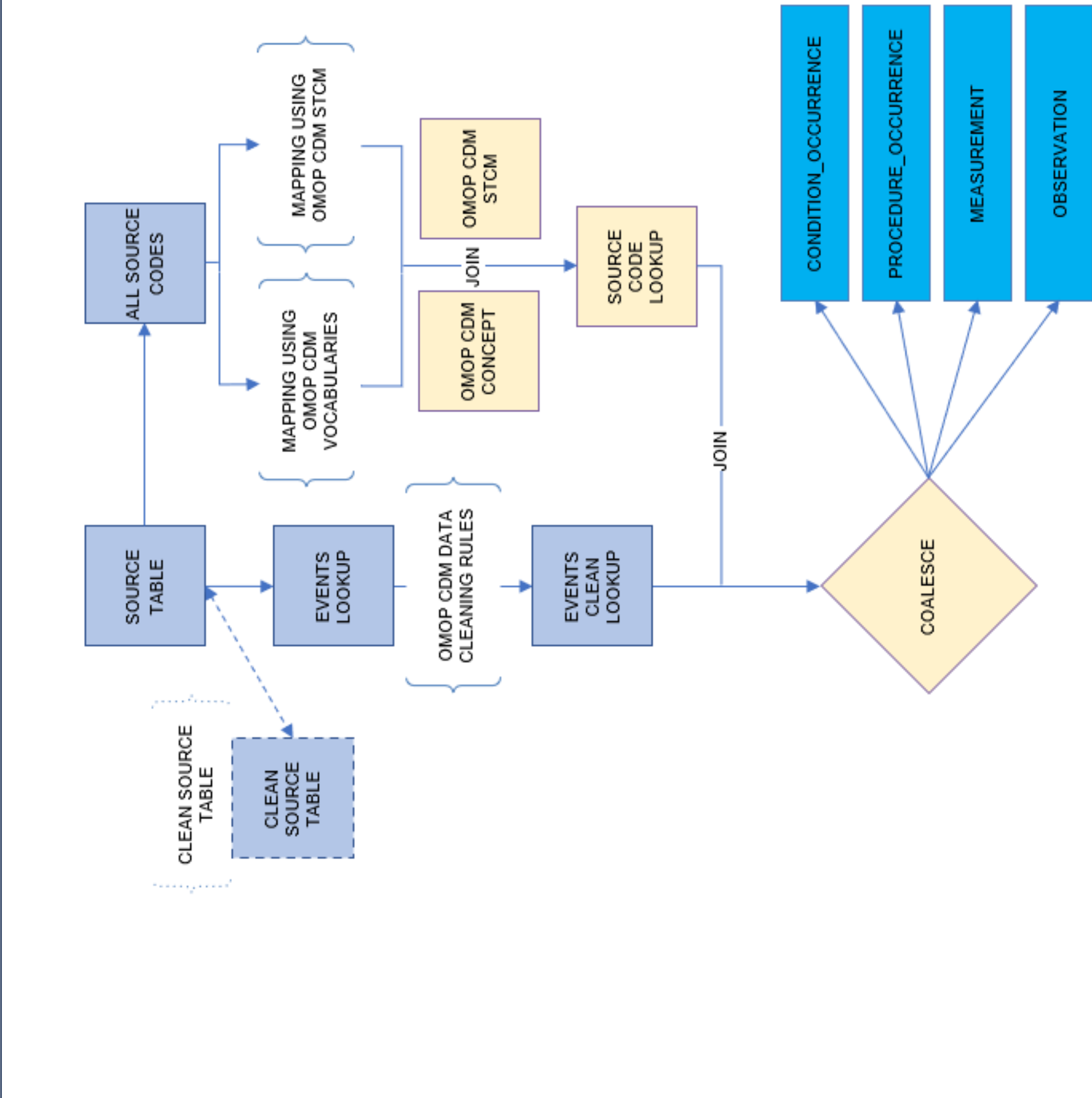


SDTM Data Variables (Raw Schema)				Converted Values (Transform Schema)			
row number	USUBID	VISIT	DOMAIN	--DTC / --ENDTC	visit_order	visit_source_value	visit_detail_source_value
1	001	Baseline	[QS]	4/4/2011	1	001 Baseline	001 Baseline 1
2	001	Baseline	[VS, LB, QS]	10/20/2020	2	001 Baseline	001 Baseline 2
3	001	Baseline	[QS]	11/13/2020	3	001 Baseline	001 Baseline 3

OMOP CDM Schema

CDM.VISIT_OCCURRENCE			
visit_occurrence_id	PERSON_ID	VISIT_SOURCE_VALUE	VISIT_CONCEPT_ID
1	111	001 Baseline	581477 – 'Office visit'
CDM.VISIT_DETAIL			
visit_detail_id	PERSON_ID	VISIT_DETAIL_SOURCE_VALUE	VISIT_DETAIL_CONCEPT_ID
1	111	001 Baseline 1	581477 – 'Office visit'
2	111	001 Baseline 2	581477 – 'Office visit'
3	111	001 Baseline 3	581477 – 'Office visit'

Example: Patient Adverse Event Logic



SDTM Data Variables (Raw Schema)				Converted Values (Transform Schema)	
USUBID	AETERM	DOMAIN	ASTDTC	AELLT	event_start_date
001	Pressure/ache in legs	[AE]	2/19/2021	Leg pain	2/19/2021
001	Loss of Appetite	[AE]	3/19/2021	Appetite lost	3/19/2021
001	Pressure in both legs	[AE]	5/14/2021	Leg discomfort	5/14/2021

OMOP CDM Schema

CDM.CONDITION_OCCURRENCE			
PERSON_ID	CONDITION_CONCEPT_ID	CONDITION_SOURCE_VALUE	CONDITION_STRT_DATE
1	442165	Clinical Indication for CM loss of appetite	3/19/2021
1	4059010	Clinical Indication for CM pressure/ache in both legs	2/19/2021
1	761158	Clinical Indication for CM pressure/ache in her legs	2/19/2021
1	442165	Clinical Indication for CM loss of appetite	3/19/2021
1	4059010	Clinical Indication for CM pressure/ache in her legs	2/19/2021

CDM.OBSERVATION			
PERSON_ID	OBSERVATION_ID	OBS_EVENT_FIELD_CONCEPT_ID	OBSERVATION_SOURCE_VALUE
1	124257	1147127	the adverse event serious? N
1	18460	1147127	the adverse event serious? N
1	85940		the adverse event serious? N

Conversion Data Quality Report

Type	Failed	Passed	Error	N/A
Table is present as expected	0	29	0	0
Completeness	0	12	0	5
Null values	0	80	0	54
Expected data type according to specification	0	91	0	98

Conclusions

The conversion process revealed strengths in the OMOP CDM's ability to standardize and harmonize diverse data sources, not just real world data sources. However, challenges included the complexity of mapping SDTM variables to OMOP due to differences in terminologies, and the loss of certain granular data inherent to the SDTM format. Despite these limitations, the conversion process was successful, enhancing the usability and interoperability of the data. Future efforts should focus on improving methods for data conversion to further optimize the utilization of real-world oncology data in research networks.