

Mapping Real-World Data (RWD) to SDTM in External Control Arm (ECA) Studies and Registry Studies

Yanhui (Kevin) Yang, Joe Xi

Feb 2024

Disclaimer

This presentation shares only the author's experience and opinions, and not necessarily those of Bristol Myers Squibb (BMS). BMS does not guarantee the accuracy or reliability of the information provided in the slides.

Outline

- **Challenges in Mapping RWD to SDTM**
- **AE vs Other Safety Domains**
- **Define Reference/Treatment Starting Dates**
- **EX vs CM (Mapping of Other Therapies)**
- **Use Customized Domain for Baseline Characteristics**
- **Sample of Pn21 Compliance Check**
- **Take-away**

Challenges in Mapping RWD to SDTM

- SDTM is expected for HA submission, but industry still lacks detailed guidance or experience
- RWD is collected by external vendor/institution. The collection process is different from clinical trial, such as:
 - Safety data structure is different, how to parse safety data into SDTM structure to facilitate clinical analysis request
 - Lack of clear definition of reference start and end date
 - Baseline information in RWD are lumped together, how to utilize them in clinical analysis

AE vs Other Safety Domains

- Safety data collected in RWD has different structure
- Dichotomic approach for the parsing safety data into SDTM.AE (general safety summary) + Other domain(e.g. IS for Infection details).

SDTM.AE & IS mapping

STUDYID	USUBJID	DOMAIN	AESEQ	AETERM	AEDECOD	AEBODSYS	AESER	AEACN	AEOUT	AETOXGR	AESTDTC	AEENDTC	AESTDY	AEENDY
Study-1	Study-1-001	AE	1	CLINICAL SIGNIFICANT INFECTION	CLINICAL SIGNIFICANT INFECTION	INFECTION	Y	XX	XX	XX	XXXX-XX-XX	XXXX-XX-XX	XX	XX

STUDYID	USUBJID	DOMAIN	ISSEQ	ISGRPID	ISTESTCD	ISTEST	ISORRES	ISSTRESC	ISRESCAT	ISLOC	ISDTC	EPOCH	ISDY
Study-1	Study-1-001	IS	1	REPORTED ORGANISM 01	VIRIDENT	Viral Organism Identification	Norovirus	Norovirus	INFECTING	GI_LOWER	XXXX-XX-XX	Post Infusion	XX
Study-1	Study-1-001	IS	2	REPORTED ORGANISM 02	VIRIDENT	Viral Organism Identification	Respiratory syncytial virus (RSV)	Respiratory syncytial virus (RSV)	INFECTING	SINUS_RESP	XXXX-XX-XX	Post Infusion	XX
Study-1	Study-1-001	IS	3	REPORTED ORGANISM 03	BCTIDENT	Bacterial Organism Identification	Clostridium difficile	Clostridium difficile	INFECTING	GI_LOWER	XXXX-XX-XX	Post Infusion	XX

Define Reference/Treatment Starting Dates

- RWD lacks clear definition of reference start date
- For registry study, designated therapy start date used
- For ECA and other types of study, clinical adjudication/AI/data-visualization used to identify index therapy. The start date of index therapy used as reference start date

Line of Therapy	Treatment					
	Index	Line	Comment	Treatment	Start Date	End Date
1		1		rituximab	2016-11-08	2016-11-18
		2		bendamustine,rituximab	2016-11-29	2017-01-17
	1	3		cisplatin,cytarabine,etoposide,methylprednisolone,rituximab	2017-02-10	2017-02-10
		1		rituximab	2017-02-13	2017-06-20
		4		doxorubicin,prednisone,rituximab	2017-06-13	2017-07-14
		5		cyclophosphamide,doxorubicin,prednisone,rituximab,vincristine	2017-07-28	2017-07-28

EX vs CM (Mapping of Other Therapies)

- Treatment of index line of therapies --> SDTM.EX
- Other line of therapies -->SDTM.CM

STUDYID	USUBJID	DOMAIN	EXSEQ	EXTRT	EXCAT	EXSTDTC	EXENDTC	EXSTDY	EXENDY
Study-1	Study-1-001	EX	1	cisplatin	SYSTEMIC THERAPY	2017-02-10	2017-02-10	1	1
Study-1	Study-1-001	EX	2	cytarabine	SYSTEMIC THERAPY	2017-02-10	2017-02-10	1	1
Study-1	Study-1-001	EX	3	etoposide	SYSTEMIC THERAPY	2017-02-10	2017-02-10	1	1
Study-1	Study-1-001	EX	4	methylprednisolone	SYSTEMIC THERAPY	2017-02-10	2017-02-10	1	1
Study-1	Study-1-001	EX	5	rituximab	SYSTEMIC THERAPY	2017-02-10	2017-02-10	1	1

Use Customized Domain for Baseline Characteristics

Why use customized BL domain?

- In RWD collected from registry studies, baseline for clinical characteristics was captured and provided by medical sites. No dates were collected for some of the clinical characteristics, which makes it difficult to apply the baseline derivation logic.
- All the baseline data is collected in one dataset, which is different from clinical trial. It is very tedious to categorize and separate them into each individual SDTM domain.

As a result, we decide to map all pre-treatment clinical characteristics into one customized domain - baseline (BL). Below advantages were achieved with this approach:

- No need to categorize each clinical characteristics assessment
- Simplify SDTM mapping process
- Beneficial to subsequent analysis

Sample of Pn21 Compliance Check

Key issues found in Pn21 report:

AE

STUDYID	DOMAIN	USUBJID	DDSEQ	DDTESTCD	DDTEST	DDORRES	DDSTRESC	DDUTC	DDDY
Study01	DD	Study01-001-00000001	1	COCPTH	Contributing Cause of Death	Cardiac failure	Cardiac failure	2022-12-23	10
Study01	DD	Study01-001-00000001	2	COCPTH	Contributing Cause of Death	Pulmonary failure	Pulmonary failure	2022-12-23	10
Study01	DD	Study01-001-00000001	3	COCPTH	Contributing Cause of Death	Cytokine release syndrome (CRS)	Cytokine release syndrome (CRS)	2022-12-23	10
Study01	DD	Study01-001-00000001	4	PRCPTH	Primary Cause of Death	Recurrence/persistence/progression of	Recurrence/persistence/progression of primary disease	2022-12-23	10

Sample of Pn21 Compliance Check

Key issues found in Pn21 report (*continued*):

LB

Take-away

In this presentation, we shared our experience from RWD mapping on:

- How to organize safety data into proper domains
- How to define the reference date in an ECA study
- How to handle complicated baseline information
- A quick review of Pn21 compliance report

This experience is preliminary and not complete, but we hope it can help shaping industry standards.

Reference

- CDISC 2023-06 ***Considerations and Examples of Using SDTM for Observational Studies***
- FDA 2021-10 ***Data Standards for Drug and Biological Product Submissions Containing Real-World Data Guidance for Industry***

Acknowledgement

Great thanks to Lorraine Fang, Joy Wang and Samar Noor

Thank you