PharmaSUG 2021 - Paper SS-072

Challenges and Solutions for an ISM Development

Siru Tang, Yanhong Li, MSD, Beijing, China

ABSTRACT

The integrated summary of microbiology (ISM) is the special integrated summary of efficacy (ISE) in antibacterial drug submission for regulatory approval. Designing integrated datasets and presenting microbiology results are very challenging for programmers due to the complexity of data structures and logics.

This paper discusses the challenges we met recently when working on an ISM. It elaborates how we transferred external excel data to internal SDTM format data, and how we adopted the data standards and pooled multiple studies' data sets together. It provides the method about how to merge treatment groups for a better display format in the reports. It also shares some tips about how we designed ADMB and ADMS datasets to include microbiological responses, pathogen names, and treatment groups so that we could generate them all in one table. It then introduces how we composed the submission document-ADRG for this ISM.

INTRODUCTION

Clinical microbiology is a science of interpretive judgment that is becoming more complex in developing antibacterial drugs. The ISM will be a fundamental instruction for both regulatory and industry in the development of next generation of antibacterial drugs.

Here is the detailed section introduction of this paper:

•Firstly, we discussed the strategy of how to build up the analysis data sets to qualify CDISC standard compliance. There were two obstacles to resolve: the consistency of data format of each study and the unavailability of the study data. This paper will elaborate how we transfer external excel data to internal SDTM format, how we adopt the data standards, and how we pool data sets together.

•Secondly, we had a challenge to merge treatment groups for better presentation in the reports, which is caused by various dosages and treatment groups in different studies in this ISM. Do we want to keep the unique visit window from each study, or can we create a new window to unify all of them? This section clarifies the method we introduced and created to utilize the best analysis visit window for a better sense to interpret the results.

•Thirdly, we proactively tested some special cases that we might have and well prepared some solutions and tips, such as how to design our ADMB and ADMS datasets so that we can easily generate such a report to include microbiological/clinical responses, pathogen names and treatment groups information all in one table. Some useful tips might help you to work on similar task from the design stage of your study.

•And finally, we will introduce how we composed the submission document-ADRG for this ISM.

1. BUILD UP ANALYSIS DATASETS

1.1. POOL DATA STRATEGY

Before we started the ISM analysis datasets design, we should have the statistical analysis plan (SAP) and all tables, listings and graphs shell in hand. The purpose of the SAP is to provide a statistical framework for the analyses of the pooled microbiology data. Therefore, how to construct data sets and define analysis population will be clarified in SAP clearly. After that, statistician should also provide the mockups of all tables, listings and graphs, which is the main document for programmers to design their analysis datasets. Since the ISM analysis is the summary of studies which have been done before, the original efficacy logic will be investigated and utilized. Extra analysis of integrated data will be considered as well.

The analysis datasets of ISM can be built in two ways:

Combine different SDTM datasets then build analysis datasets: at least DM, DS, MB, MS, FA domains should be included to develop the analysis datasets of ADSL/ADMS/ADMB;

Use different analysis datasets directly: this approach can be applied only if all studies in the ISM have similar design for analysis datasets and the data sets are easy to combine.

In this ISM, we applied both ways to develop our analysis datasets. We used the first way to combine extra MS data with original MS data to get new SDTM MS data, then develop ADMS for integrated analysis. For other datasets, which can be combined without adding new data, we used the analysis datasets directly. To help you better understanding our solution, please refer to Figure 1 as follow:
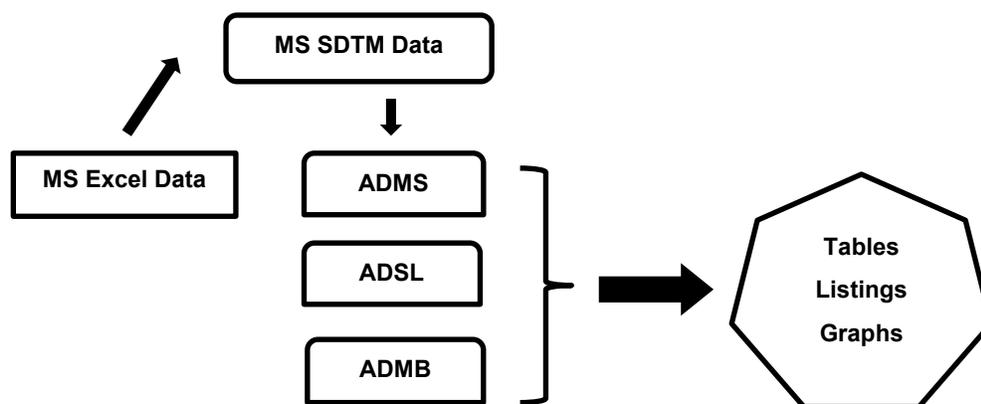

Figure 1 Pooling strategy

1.2.   TRANSFER MS DATA

One obstacle which we need to deal with was that some MS data were not available at the database lock stage for each study. We later received the external MS data in Excel format as follow:

LNKID	ORGANISM	IMIREDIS	IMIREDIU	MEROMIC	MEROMICU	REFERECE
1234567	Escherichia coli	30	mm			D454588
1345678	Pseudomonas aeruginosa	25	mm	15	µg/mL	
1456789	Staphylococcus epidermidis	10	mm			D038828

Table 1 External MS data

However, MS data should be submitted as follow (only some columns presented):

USUBJID	MSREFID	MSGRPID	MSTEST	MSCAT	MSSCAT	MSSTRESC	MSSTRESU	MSLNKID
1234-003_0001	C500314	UBCS	meropenem	Urine and Blood Culture Susceptibility	Escherichia coli	7	mcg/mL	
1234-003_0001	EF33188	EURO	Imipenem	Eurofins	Escherichia coli	0.24	microgm/mL	

Table 2 SDTM MS data

In order to keep our data clean and compliant, we used the following code to convert Table 1 to Table 2:

```
data sdtm_ms;
  set extra_ms;
  *To create similar structure of MS;
  DOMAIN='MS';
```

```
      MSGRPID='GEFCF';
    *If the record is for minimum inhibitory concentration result, then
create a record in MS as follow:
        MSTESTCD='MICRESLT';
        MSTEST="MINIMUM INHIBITORY CONCENTRATION RESULT";
        MSCAT="SUSCEPTIBILITY";
        MSSCAT="MEROPENEM";
        A=ROUND(MEROMIC,0.001);
  ...
        IF MEROMIC<A THEN MSORRES='<='||STRIP(PUT(A,BEST.));
        ELSE IF MEROMIC>A THEN MSORRES='>='||STRIP(PUT(A,BEST.));
        ELSE MSORRES=STRIP(PUT(A,BEST.));
        MSLNKID=STRIP(PUT(LNKID,BEST.));
        output;
    * end;
    * If the record is for zone diameter, then create a record in MS as
follow:
        MSTESTCD='ZONEDIAM';
        MSTEST="ZONE DIAMETER";
        MSCAT="DISK DIFFUSION";
        MSSCAT="IMIPENEM/RELEBACTAM";
        MSORRES=STRIP(PUT(IMIREDIS,BEST.));
        ...
        MSLNKID=STRIP(PUT(LNKID,BEST.));
        MSREFID=REFERENCE;
    * end;
  RUN;
```

Then, we used MSLNKID or MSREFID to combine the USUBJID for each record. And the Table 1 turned into:

USUB JID	MSRE FID	MSGR PID	MSTE ST	MSCAT	MSSCAT	MSSTR ESC	MSSTR ESU	MSLN KID
1234-003_0 001	D4545 88	GEFCF	Zone Diame ter	DISK DIFFUSI ON	IMIPENEM/RELEB ACTAM	29	mm	12345 67

Table 3 SDTM MS data From External MS

2. ANALYSIS DETAILS COMBINATION

2.1. TREATMENT GROUPS COMBINATION

In FDA Microbiology Analysis Guidance for Industry APPENDIX C, it requests at least the following 12 reports for every ISM submission:

- **Four summary tables** are requested for each proposed indication：
    - MIC by microbiological response (MR) for each baseline bacterium by indication
    - MIC by clinical response (CR) for each baseline bacterium by indication
    - Zone diameter by microbiological response for each baseline bacterium by indication
    - Zone diameter by clinical response for each baseline bacterium by indication
- **Two listing reports** for isolates within specified range of AST criteria (S/I/R)
- **Two listing reports** for isolates associated with microbiological/clinical failure
- **Two summary reports** of isolates at MIC associated with each micro/clinical response, grouped by bacterial type
- **Two histograms** show the # of isolates at each MIC from clinical studies overlaying isolates from

nonclinical studies associated with each micro/clinical response

In our ISM, we had three studies to present, two Phase II studies and one Phase III study. To better display in the output, we combined treatment groups with dosages and included all placebo or controlled arms into one group as 'All Comparators'. We created three treatment groups in our ADSL and retained protocol-level of treatment groups in a new treatment group variable to keep data transparency. Regarding the purpose of producing four summary tables by indication, we decided to use the following table header:

cUTI			cIAI			HABP/VABP		All Indications		
Dose 1 n (%)	Dose2 n (%)	Comparator n (%)	Dose1 n (%)	Dose2 n (%)	Comparator n (%)	Dose1 n (%)	Comparator n (%)	Dose1 n (%)	Dose2 n (%)	Comparator n (%)

Table 4 Table Design

2.2. ANALYSIS VISIT WINDOW COMBINATION

In the integrated summary analysis, we need to combine all studies' different analysis visit window together. Before starting program development for efficacy data, the statistician and clinician decided the reasonable combination of analysis visits. In our study, we have following protocol-level analysis visits:

Analysis Visit	Study 1	Study 2	Study 3
Screening	Screening	Screening	Screening
EOT	DCIV		EOT
EFU	EFU	EFU	EFU
Day 28		GFU	Day 28

Table 5 Protocol-level Analysis Visits

Because we had different indication types in each study, to make analysis visit window clear, we used Phase III study (study 3 in Table 5) analysis visit to keep consistency. In some studies, we may encounter different defined analysis visits in each study, if the analysis day has overlaps, we suggest programmers use analysis day to re-define new analysis visits instead of using the above re-naming strategy.

3. SPECIAL CASES

3.1. ADAMS DESIGN

3.1.1. RESPONSE VARIABLES

In each study we included, we had three analysis datasets for efficacy, ADMB, ADMS, and ADEFF. ADMB included all microbiological specimen information which showed in the output as baseline pathogens. ADMS contained all drug susceptibility information which showed in the output as MIC results. ADEFF consisted of all responses (clinical response/ microbiological response/ overall response) information which showed in the output as clinical response or microbiological response. According to our output reports' design, we decided to combine ADMB and ADMS, and include response variables inside ADMS. In practice, we suggest programmers create four variables: per subject microbiological response, per pathogen microbiological response, per subject clinical response, per pathogen clinical response. Microbiological response and clinical response are required responses in your output, but if you have more, please add more variables based on study design. To explain the purpose of combing ADMS and ADEFF, we put one of the tables as below:

	Drug 1				Drug 2			
	MIC (µg/mL)	Dose1 n/N(%)	Dose2 n/N(%)	Comparator n/N(%)	MIC (µg/mL)	Dose1 n/N(%)	Dose2 n/N(%)	Comparator n/N(%)
Subjects in population, N Gram-negative aerobic bacilli		75	73	78		75	73	78
Acinetobacter baumanii complex		0	1/2 (50.0)	2/2 (100.0)		0	1/2 (50.0)	2/2 (100.0)
	16	0	0	1/1 (100.0)	16	0	0	1/1 (100.0)
	32	0	0	1/1 (100.0)	32	0	0	1/1 (100.0)
	≥ 32	0	1/2 (50.0)	0	≥ 32	0	1/2 (50.0)	0

Table 6 Favorable Microbiological Response n/N (%) at the EOT Visit for Key Pathogens from Infection Site or Blood Culture by Baseline MIC (µg/mL)
Phase 2/3 (PN001 and PN002)

3.1.2. ZONE DIAMETER DATA

Since we already transferred these data as stated in Section 1.2 into MS SDTM data, now the data structure of zone diameter for study drug regarding each pathogen in MS was as same as drug susceptibility. Therefore, we designed zone diameter as one of the parameters in ADMS dataset:

	Study Drug			
	Zone Diameter (mm)	Dose 1 n/N (%)	Dose 2 n/N (%)	Comparator n/N (%)
Subjects in population Enterobacteriaceae		20	25	35
	10	0	0	0/1 (0.0)
	18	0	0	1/1 (100.0)
	24	0/1 (0.0)	0	

Table 7 Favorable Microbiological Response n/N (%) at the EFU Visit for Key Baseline Enterobacteriaceae from Infection Site or Blood Culture
by Baseline Zone Diameter (mm) From Pooled Subjects
Phase 2/3 cUTI (PN001 and PN002) and Phase 2 cIAI (PN003)

3.2. GROUP VARIABLES

In ISM submission, we always have several group variables based on phenotype and genotype. In our study, we had five group variables which needed to be put into several tables。 Some tables just needed ADMB and the others only needed ADMS, so we incorporated five group variables into both ADMB and ADMS to satisfy our output need. For tables which were generated from ADMB, please see the follow:

	cUTI		cIAI		HABP/VABP		All Indications	
	Dose1 n (%)	Comparator n (%)	Dose1 n (%)	Comparator n (%)	Dose1 n (%)	Comparator n (%)	Dose1 n (%)	Comparator n (%)
ESBL-producer								
Subjects in population, N	8	4	0	0	0	0	8	4
Gram-negative aerobic bacilli	8 (100.0)	4 (100.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	8 (100.0)	4 (100.0)
Enterobacter cloacae	2 (20.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	2 (20.0)	0 (0.0)
Klebsiella pneumoniae	2 (20.0)	1 (25.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	2 (20.0)	1 (25.0)
KPC-producer								
Subjects in population, N	4	1	1	0	0	0	6	1
Gram-negative aerobic bacilli	4 (100.0)	1 (100.0)	1 (100.0)	0 (0.0)	0 (0.0)	0 (0.0)	6 (100.0)	1 (100.0)
Citrobacter freundii	0 (0.0)	0 (0.0)	1 (100.0)	0 (0.0)	0 (0.0)	0 (0.0)	2 (25.0)	0 (0.0)
Enterobacter cloacae	1 (33.3)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	2 (25.0)	0 (0.0)

Table 8 Baseline Pathogens Identified from Infection Site Culture by Genus and Species by Beta-Lactamase Status
PN003

For table created from ADMS, it is similar to Table 7 as follow:

	Zone Diameter (mm)	Study Drug n (%)
ESBL-producer		
Subjects in population		6
Gram-negative aerobic bacilli		
Enterobacter cloacae, N		1
	30	1 (100.0)
Klebsiella pneumoniae, N		3
	23	1 (33.3)
	24	1 (33.3)
	25	1 (33.3)
KPC-producer		
Subjects in population		6
Gram-negative aerobic bacilli		
Citrobacter freundii, N		1

Table 9 Baseline Zone Diameter (mm) of Imipenem/REL of Key Pathogens from Infection Site by Genus and Species by Beta-Lactamase Status
PN003

3.3. SPECIAL REQUIRED FIGURE

After submission, we were required by agency to provide a group of figures about Minimum Inhibitory Concentration (MIC) distributions of study drug for key Enterobacter ales from clinical (Phase II Trials and Phase III Trials) and 2016-2018 global surveillance isolates comparation. The surveillance isolates data can be downloaded from related clinical website and reviewed by clinicians. The output is as following figure:
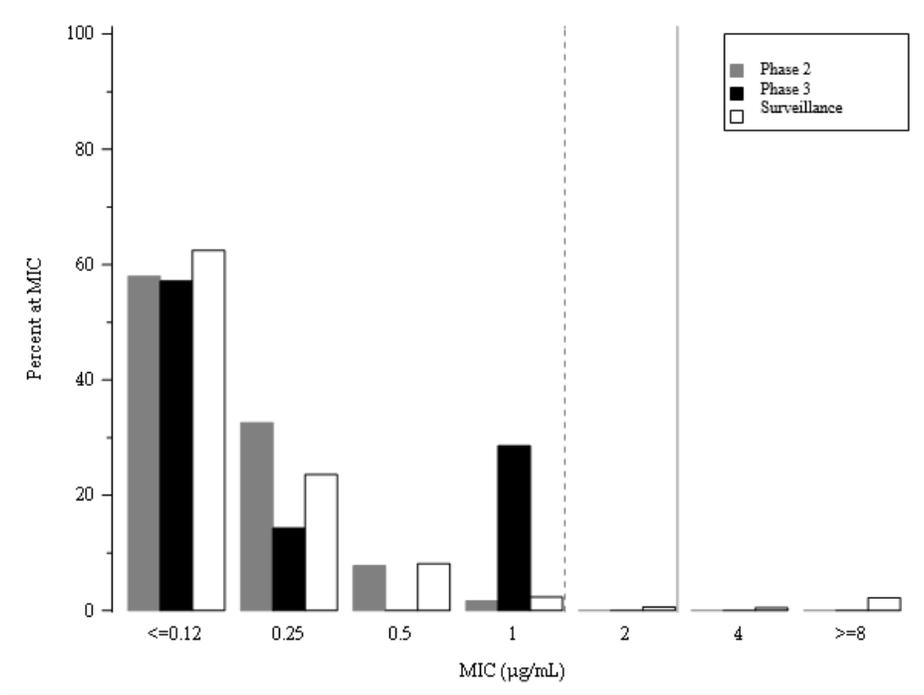

Figure 2 MIC Distribution Figure

The dotted line and solid line are required for: The proposed CLSI broth microdilution susceptible breakpoint for Enterobacterales is ≤1 μg/mL (dotted line). The provisional EUCAST susceptible breakpoint for Enterobacterales is ≤2 mg/L (solid line).

4. ADRG PREPERATION

For ISM submission, after compliance check in Pinnacle 21, stats and programmers should use Analysis Data Reviewer's Guide (ADRG) template to prepare ISM ADRG, which follows the same process as each protocol submission. To be specific, the ISM ADRG has the following four items need both stats and programmers to pay more attention to:

Make sure each individual study's ADaM datasets and SDTM datasets version, Medical Events Dictionary, and Define documents version are all clear and correct. This will be filled in Section 1.3 in ADRG;

If any excel data used in the study design, we should put all of them in the Appendix in ADRG. If the data set is too large, we should submit the excel or SAS format data to agency instead of putting into ADRG;

Besides the final ADaM datasets, the intermediate data used to develop ADaM datasets should also be added into ADRG and explained in Section 4 (Data Dependencies) and Section 5 (Analysis Dataset Descriptions);

If you put protocol-level treatment group into new self-defined TRTxxA/TRTxxP variables, this will be identified as an error in Pinnacle 21. We need to add explanation in Section 6.3 Issue Summary. Please see the following explanation as a reference:

Rule ID	Dataset(s)	Diagnostic Message	Severity	Explanation
AD0062 AD0073	ADSL	Illegal variable name: xx is not in [01-99] for TRTPLA, TRTPLP, TRTPLPN and TRTPLAN	Error	These treatment group variables are customized variables for keeping the original values in the protocol-level analysis dataset. which would be used to display original treatment group information in the listings.

Table 10 Data issue example

CONCLUSION

The ISM is an important part of efficacy summary in antibacterial drug submission, especially when the sample size of individual Phase III study is small. Through ISM analysis, we can find more valuable information by combining multiple studies together. Overall, our paper shared the practices on how to deal with specific situations during preparing ISM analysis datasets, outputs and submission along with other considerations that should be taken while working on ISM.

ABBREVIATION

Abbreviation	Description
ISM	Integrated Summary of Microbiology
ISE	Integrated Summary of Efficacy
ADMB	Analysis Dataset of Microbiology
ADMS	Analysis Dataset of Microbiology Susceptibility
MS	Microbiology Susceptibility
MB	Microbiology
ADSL	Subject-level Analysis Dataset
DM	Demographics
DS	Disposition
ADRG	Analysis Data Reviewer's Guide
CLSI	Clinical and Laboratory Standards Institute
EUCAST	European Committee on Antimicrobial Susceptibility Testing
MIC	Minimum Inhibitory Concentration

REFERENCES

Bharath Donthi, Lingjiao Qi, 'Best Practices for ISS/ISE Dataset Development'
PharmaSUG 2019

CONTACT INFORMATION

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

Author Name: Siru Tang
Company: MSD

City / Postcode: Beijing, China/100012
Email: si.ru.tang@merck.com

Author Name: Yanhong Li
Company: MSD
City / Postcode: Beijing, China/100012
Email: yanhong.li1@merck.com

Any brand and product names are trademarks of their respective companies.