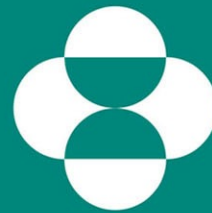


Paper TA05

Programming Approaches for Integrated Summary of Efficacy (ISE) Development in Oncology

Narendra Babu Kolisetty
Saigovind Chenna



MERCK

INVENTING FOR LIFE

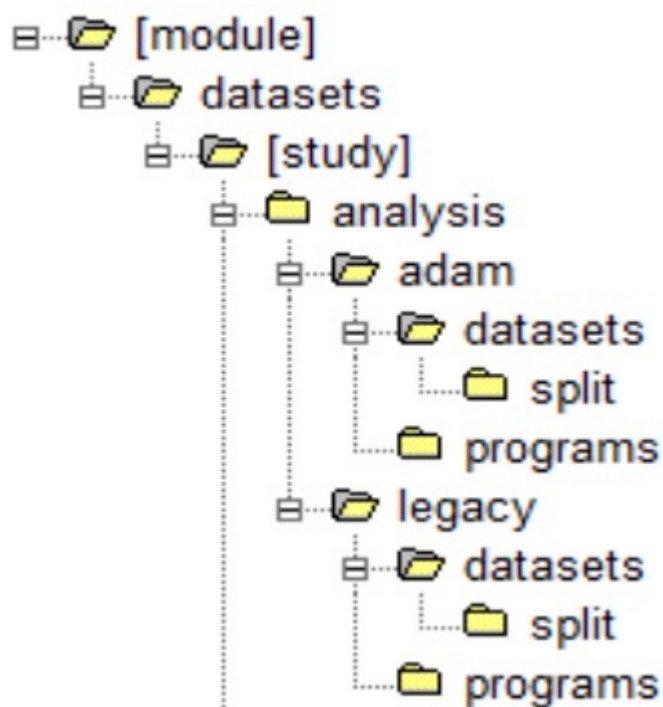
Agenda

- Integrated Summary of Efficacy (ISE)
- Integration strategies
- ISE Programming Approaches
- Conclusion
- Acknowledgements & References
- Q&A

Integrated Summary of Efficacy (ISE)

- ISE provides a comprehensive evaluation of a study drug's benefit-risk profile, extending beyond the findings of individual clinical trials.
- Consolidates the drug's efficacy across various population subgroups, demographics, dosing regimens, and dose-response relationships.
- Enhances the accuracy of results due to larger sample size, and the application of robust statistical methods for efficacy data analysis.
- The eCTD's Module 5, section 5.3.5.3, titled "Reports of analyses from multiple studies," should include integrated data from multiple studies.

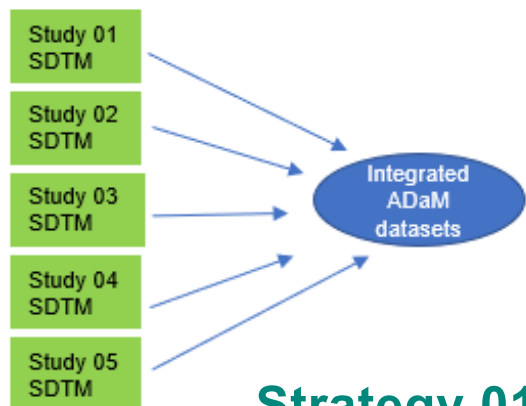
ISE eCTD Folder Structure



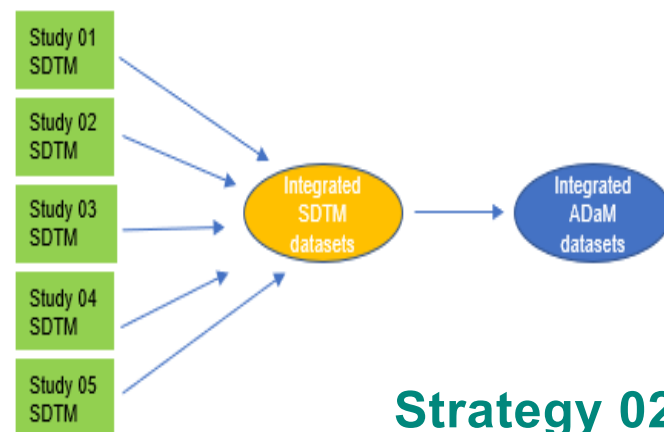
Folder Name	Folder Level	Description/Contents
 [module]	1	Refers to the eCTD module in which study data are being submitted. Name this folder m4 for nonclinical data and m5 for clinical data. Do not place files at this level.
 datasets	2	Resides within the module folder as the top-level folder for study data (nonclinical or clinical) being submitted for the specified module (m4 or m5). Do not place files at this level.
 [study]	3	Name this folder with the study identifier or analysis type performed (e.g., study123, iss, ise). Do not place files at this level.
 analysis	4	Contains folders for analysis datasets and software programs; arrange in designated level 6 subfolders. Do not place files at this level.
 adam	5	Contains subfolders for ADaM datasets and corresponding software programs. Do not place files at this level.
 datasets	6	Place ADaM datasets in this subfolder.
 split	7	Place any split ADaM datasets in this subfolder.
 programs	6	Place software programs for ADaM datasets, tables and figures in this subfolder.
 legacy	5	Contains legacy formatted analysis datasets and corresponding software programs. Do not place files at this level.
 datasets	6	Place legacy analysis datasets in this subfolder. In m4 place tumor.xpt and its associated define.pdf in this folder.
 split	7	Place split legacy analysis datasets in this subfolder.
 programs	6	Place software programs for legacy analysis datasets, tables and figures in this subfolder.

eCTD file folder structure and their associated descriptions

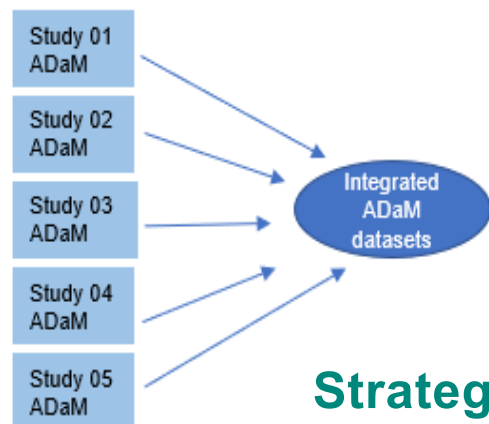
ISE Integration Strategies



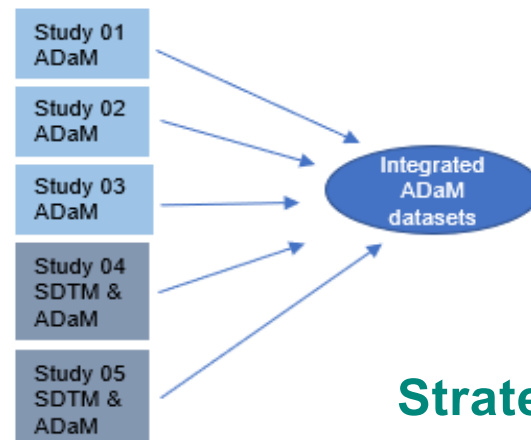
Strategy 01



Strategy 02



Strategy 03



Strategy 04

Tumor Lesion Data

STUDYID	PARAM	PARAMCD	AVAL	BASE	CHG	PCHG
Study 01	Minimum of Sum of Target/Index Lesions in Longest Diameter	MSTLDIAM	30	55	-25	-45.4545
Study 02	Minimum of Sum of Target Lesions in Longest Diameter	MSTLDIAM	100	105	-5	-4.7619
Study 03						
Study 04	Minimum of Sum of Target Lesions in Longest Diameter (Radiology)	RADMSMTL	55	125	-75	-60
Study 05						
Study 06	Minimum of Sum of Target Lesions in Diameter Based on Central Radiology Assessment (m...	MSTLDIRC	43	30	13	43.3334
Study 07	Minimum of Sum of Target Lesions in Longest Diameter (Radiology)	RADMSMTL	96.5	67	29.5	44.0299
Study 08	Minimum of Sum of Target Lesions in Longest Diameter Per RAD	MRADTLM	67	89	-22	24.7191
Study 09	Minimum of Sum of Target Lesions in Longest Diameter per BICR	IRCMSTL	30	34	-4	11.7647
Study 10	Minimum of Sum of Target Lesions (BICR per RECIST 1.1)	MSTLMIRC	50	57	-7	12.2807

- Target or non-target or new tumor lesion data related parameter/variable can be derived during integration by utilizing SDTM data from the input studies.
- Any identified data-related Issues that cannot be fixed and there is no impact to analysis must be clearly documented in the Analysis Data Reviewer's Guide (ADRG).

Tumor Response Data

STUDYID	PARAM	PARAMCD	PARAMN	AVALC	AVAL
Study 01	Best Overall Response per RECIST 1.1 (IRC Radiology Only)	BORRT	111	PR	.
Study 02	Best overall response with confirmation (IRC)	BORCFIRC	22	PR	3
Study 03	Best Overall Response (IRC) with Confirmation	BORCFIRC	22	PR	3
Study 04	Best Overall Response per RECIST1.1 (Radiology)	RADBOR	22	PR	2
Study 05	Best overall response with confirmation (Central Radiology per RECIST 1.1)	BORCFIRC	22	PR	2
Study 06	Best Overall Response with Confirmation Based on RECIST 1.1 per Central Radiology Assessment	BORCIRC	202	PR	3
Study 07	Best Overall Response per RECIST1.1 (Radiology)	RADBOR	22	PR	2
Study 08	Best overall response confirmed - RECIST 1.1, Imaging Vendor	BORRAD	21	PR	.
Study 09	Best overall response with confirmation (BICR per RECIST 1.1)	BORCFIRC	22	PR	3

- Tumor response data is typically present in one of the efficacy analysis datasets like ADEFF, ADRS, ADORR, and ADIRC
- Re-derive PARAM, PARAMCD, PARAMN, and AVAL variables

Legacy Data

- Use of an external data beyond SDTM or ADaM datasets
- Examples: biomarker, tumor mapping, or genotype/phenotype data
- ISE datasets development and analysis include deriving new variables like new population flags, grouping various tumor types, classifying biomarker data, generating efficacy TLFs based on tumor type, or biomarker status etc.

Variable Name	Variable Label	Type	Length	Origin	Define Derivation
BLBMRKR	Baseline Biomarker	integer	8	Assigned	Merge the external data file "biomarker.xlsx" to the ADSL dataset by subject. Assign the values from result column by subject.
BLBMRKS	Baseline Biomarker Status	Char	40	Derived	BLBMRKS="Present", if the result result value is >10. BLBMRKS="Not-present", if the result value is not missing and <=10 BLBMRKS value is null if no result value
BLBMRKSN	Baseline Biomarker Status (N)	integer	8	Assigned	Mapped per codelist values
TMRGR	Tumor Type Grouping	Char	100	Derived	Merge the external data file "tumortype.dat" to the ADSL dataset by tumor type. Count participants by each tumor type. TMRGR="Other", if the tumor type count is <=20 Otherwise, TMRGR=TUMOR from tumortype.dat file.
TMRGRN	Tumor Type Grouping (N)	integer	8	Assigned	Sort unique tumor type values alphabetically first, and then assign numeric values from 1 to N, by interval 1. TMRGRN=999, if TMRGR= "Other".

Participants selection criteria

If ISE analysis include only a subset of the total participants from the input studies

- Participants selection criteria for ISE must be clearly documented in the ISE ADRG. For example: Integration is performed only on the data from participants that were received the treatment as a first-line therapy.
- Create a new population flag variable during the development of ISE datasets to identify list of participants, and to analyze the efficacy data.
- sub-group efficacy data analysis might be conducted to compare drug's efficacy results. Ex: by dosing regimens, by tumor type.

SDTM and ADaM IG Differences

- All required modifications, including variable name, label, formats, control terminology, and derivations, should be implemented in accordance with the ADaM IG designated for the ISE.

Harmonization of RACE variable control terminology in ISE ADSL dataset.

▲ STUDYID	▲ RACE	▲ RACE_NEW
Study 01	MULTIRACIAL	MULTIPLE
Study 02	MULTIPLE	MULTIPLE
Study 03	Multi-racial	MULTIPLE
Study 04	Multi-Racial	MULTIPLE

SDTM and ADaM IG Differences

- SRCDOM variable label and derivation as per ADaM IG 1.0.

ADaM IG	Variable Name	Variable Label	Define Derivation
1.0	SRCDOM	Source Domain	2-letter SDTM dataset name that relates to the analysis value

- Changes to SRCDOM variable label & derivation as per ADaM IG 1.1.

ADaM IG	Variable Name	Variable Label	Define Derivation
1.1	SRCDOM	Source Data	ADaM or SDTM dataset name that relates to the analysis value

- Derivation of an actual arm variable in the ISE ADSL dataset due to difference in SDTM IG among input studies.

Variable Name	Variable Label	Type	Length	Origin	Define Derivation
ACTARM	Description of Actual Arm	Char	40	Derived	For Study 01, 03, 04: ADSL.ACTARM. For Study 02, 05, 06: ADSL.ACTLARM

Null Values

▲ STUDYID	▲ VAR1	⊕ VAR1N	▲ VAR1_NEW	⊕ VAR1N_NEW
Study 01	null	999	Null	999
Study 02	not reported	999	Null	999
Study 03	MISSING	9999	Null	999
Study 04		.	Null	999
Study 05	NULL	99	Null	999
Study 06	NOT REPORTED	99999	Null	999

- Null values can be present in both character and numeric variables
- Null values might have reported in various ways like "Missing", "Null", "Not Reported", or even as a blank value in character variables, and as 99, 999, 9999, and 99999 in numeric variables
- Consolidate these differences into one by consistently assigning a unique character and corresponding numeric value across all ISE datasets

Traceability Variables

- Metadata traceability variables SRCDOM and SRCVAR, and data point traceability variable SRCSEQ offer a data point traceability from SDTM to ADaM datasets.
- However, these traceability variables may not be present if an individual study analysis was not adhered to ADaM standards.
- If the value of the AVAL/AVALC/ADT variable is directly carried forward from either an SDTM or another ADaM dataset, it is advisable to include these variables for traceability.

Value Truncation (Length Statement)

STUDYID	CNTRYNAM
Study01	Brazil
Study02	Canada
Study03	Chile
Study04	Czech
Study05	Italy
Study06	Japan
Study07	Korea,

```
data adsl(keep=studyid cntrynam);
  set adsl1 adsl2 adsl3 adsl4
      adsl5 adsl6 adsl7;
run;
```

STUDYID	CNTRYNAM
Study 01	Brazil
Study 02	Canada
Study 03	Chile
Study 04	Czech Republic
Study 05	Italy
Study 06	Japan
Study 07	Korea, Republic of

```
data adsl(keep=studyid cntrynam);
  length cntrynam $20.;
  set adsl1 adsl2 adsl3 adsl4
      adsl5 adsl6 adsl7;
run;
```

Value Truncation (formats and informats)

▲ STUDYID	▲ USUBJID
Study 01	1111-001-100001
Study 02	1111-001-100002
Study 03	1111-001-100003
Study 04	1111-001-100004
Study 05	1111-001-100005
Study 06	1111-001-100006
Study 07	1111-001-100007

```
data adsl;
  length studyid $10 usubjid $40;
  set adsl01 adsl02 adsl03
      adsl04 adsl05 adsl06 adsl07;
run;
```

▲ STUDYID	▲ USUBJID
Study 01	1111-001-1000011
Study 02	1111-001-100002
Study 03	1111-001-100003
Study 04	1111-001-1000042
Study 05	1111-001-100005
Study 06	1111-001-100006
Study 07	1111-001-100007

```
data adsl;
  length studyid $10 usubjid $40 ;
  set adsl01 adsl02 adsl03
      adsl04 adsl05 adsl06 adsl07;
  format _all_;
  informat _all_;
run;
```

eSUB Checks

- The ISE ADaM eSUB package: define.xml, SAS datasets in .xpt format, ADRG, and program files (.txt).
- ADaM compliance checks on ISE datasets & define using Pinnacle21.
- There should not be any technical rejection criteria.
- Check ISE define file derivation column, variable origin, consistency between define and datasets.
- Any identified data-related Issues that cannot be fixed and have no impact to the analysis must be clearly documented in the ADRG.

Conclusion

- Integrating efficacy data from multiple late-phase clinical trials into a high-quality ISE data package is always challenging for programmers and statisticians.
- These challenges can be mitigated by following an effective programming strategy.
- This presentation summarized some efficient programming approaches for integrating efficacy data at ADaM level.

Acknowledgments & References

Acknowledgments

- The authors would like to thank Anand Jayaraman, Himanshu Patel, Madhusudhan Reddy Papasani, Mary Varughese, and Amy Gillespie for reviewing the presentation and providing great suggestions.

References

- FDA Study Data Technical Conformance Guide <https://www.fda.gov/media/153632/download>
- FDA Study Data Standards Catalog <https://www.fda.gov/industry/fda-data-standards-advisory-board/study-data-standards-resources>

Contact Information

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

Author Name: Narendra Babu Kolisetty

Company: Merck & Co., Inc.

Address: 351 N Sumneytown Pike, North Wales, PA 19454

Work Phone: +1 (267) 305 0403

Email: narendra.kolisetty@merck.com

Author Name: Saigovind Chenna

Company: Merck & Co., Inc.

Address: 351 N Sumneytown Pike, North Wales, PA 19454

Work Phone: +1 (267) 305 2162

Email: saigovind.chenna@merck.com

THANK YOU



MERCK

