PHUSE US Connect 2026 – Paper SA14

Data Harmonization for ISS and ISE: a Practical Checklist for Integration Success

Lily Cheng, Qiong Wei, Ji Qi, Lilian Li, Lixin Gao
BioPier LLC, a Veramed Company

ABSTRACT

Integrated Summary of Safety (ISS) and Integrated Summary of Efficacy (ISE) are pivotal components of regulatory submissions, requiring harmonized, CDISC-compliant data across multiple clinical studies. Having a practical checklist helps guide programmers through the integration process, addressing common challenges such as alignment of dataset implementation guides (IGs), harmonization of variable attributes, formats and casing, reconciliation of measurements (e.g. --PARAM/--PARAMCD) and units, and standardization of coding dictionaries (e.g. MedDRA, WHODRUG, LOINC) and controlled terminology, alignment of visit mapping and derivation logic of key analysis variables (e.g., ABLFL, ANLxxFL, AVALCATx), and coordination of subject identifiers for subjects participated in multiple studies. By applying this checklist, teams can improve integration accuracy and ultimately strengthen the overall integrity of ISS and ISE submissions.

INTRODUCTION

The preparation of Integrated Summary of Safety (ISS) and Integrated Summary of Efficacy (ISE) represents a critical milestone in regulatory submissions, particularly for New Drug Applications (NDAs) and Biologics License Applications (BLAs). These integrated analyses combine data from multiple clinical studies to provide a comprehensive evaluation of a drug's safety and efficacy profiles across diverse patient populations and treatment settings. An ISS would typically include patients that were exposed to the investigational drug, while ISE would include efficacy data and efficacy endpoints.

The complexity of integrating data from multiple clinical studies may seem like a daunting task when starting out for programming teams. Each individual study may have been conducted over different time periods by different vendors with varying protocols, SAP, implementation guides, and programming conventions. Starting with the Statistical Analysis Plan (SAP) and Study Data Standardization Plan (SDSP), if available, and working with a study statistician can help streamline the statistical analyses needed to be performed.

This paper presents a practical checklist in preparing ISS and ISE submissions. The items listed to check can provide a structured approach to identifying and resolving data harmonization issues, ensuring that integrated datasets meet CDISC standards and regulatory expectations while maintaining data integrity and traceability.

DATASET IMPLEMENTATION GUIDE ALIGNMENT

First, consider which studies to include and which datasets will be needed for the integrated summary, and which data standard to serve as the basis. The selection of a pivotal study as the reference point can help establish baseline standards for integration. Integrated summaries can include a mixture of legacy, SDTM, and ADaM datasets needed. Work with your sponsor to decide the method of integration needed – to create an integrated ADaM or the need for an integrated SDTM and ADaM.

ASSESSMENT OF STUDY-LEVEL IMPLEMENTATION GUIDES

Begin by collecting all implementation guides from studies included in the integration. Note the domains included in integrated analyses, such as (but not limited to) ADSL (subject-level analysis dataset), ADAE (adverse events), ADCM (concomitant medications), ADLB (laboratory data), ADVS (vital signs), and efficacy-specific domains.

Example: Consider an integration project combining three Phase 2/3 studies:

- Study A: SDTM IG v3.1.2, ADaM IG v1.0
- Study B: SDTM IG v3.2, ADaM IG v1.1
- Study C (pivotal study): SDTM IG v3.4, ADaM IG v1.2

In this scenario, study C is chosen as the pivotal study, requiring studies A and B to be harmonized to match Study C's implementation guide.

RECONCILIATION OF VARIABLE SPECIFICATIONS AND CONTROLLED TERMINOLOGY

After reviewing the implementation guides, evaluate the variable specifications across all studies. This involves checking variable names, types, lengths, labels, and controlled terminology. Use the integrated SAP to determine the

required populations, analysis groups, and endpoints. Identify which studies apply similar derivations based on these specifications and consult the sponsor if any derivation is unclear. If no comparable derivations exist, new variables or parameters will need to be created in the integrated dataset. All decisions and derivations should be documented in the specification.

Example: Differences in treatment dose level across studies:

- Study A has dose levels of 100 mg, 200 mg, 300 mg, 400 mg, 500 mg, 600 mg, 700 mg, and 800 mg.
- Study B has dose levels of 300 mg, 600 mg, and 900 mg.
- Study C has a dose level of 600 mg.

In accordance with the integrated SAP requirements, an integrated variable will be created to categorize doses into the following groups: < 300 mg, 300 mg - <= 600 mg, and > 600 mg.

Example: Differences in Causality terminology across studies:

- Study A uses causality terms of 'Unlikely' and 'Definite'.
- Study B uses causality terms of 'Not related', 'Unlikely', 'Possibly', and 'Related'.
- Study C uses causality terms of 'Not Suspected' and 'Suspected'.

Per CDISC analysis conventions, analysis causality must be standardized into two categories: 'Related' and 'Not Related'. The terms 'Unlikely', 'Not related', and 'Not Suspected' will be mapped to analysis causality value of 'Not Related'. The terms 'Definite', 'Possibly', 'Related', and 'Suspected' will be mapped to analysis causality value of 'Related'.

CODING DICTIONARIES

Different studies may use different versions of coding dictionaries. Note the individual versions used across studies and follow either the latest version or pivotal study version, as determined in consultation with the sponsor.

DICTIONARY VERSIONING

The MedDRA (Medical Dictionary for Regulatory Activities) is used to code adverse events and WHODrug Global to code medications. All adverse events and medications should be coded to a single version of the respective dictionaries in ISS. All decisions and derivations can be documented in the specification and Analysis Data Reviewer's Guide.

Example: MedDRA version harmonization scenario:

- Study A used MedDRA version 20.1
- Study B used MedDRA version 22.0
- Study C (pivotal study) used MedDRA version 26.1
- ISS submission requests MedDRA version 26.1

In consultation with the sponsor, determine which version will be used for the ISS submission; any affected studies will need to be up versioned to the selected version (for example, Version 26.1).

VARIABLE ATTRIBUTES AND METADATA HARMONIZATION

After aligning dataset structures and coding dictionaries, the next consideration is variable attributes and metadata to ensure consistency in how data are represented and stored across integrated datasets. This harmonization extends to all aspects of variable definition including labels, lengths, formats, types, and controlled terminology.

VARIABLE LABELS AND LENGTHS

Variable labels must be standardized across all studies for each variable and must comply with CDISC specifications. Inconsistent labels may create confusion during analysis and may trigger validation warnings. Variable names must not exceed 8 characters, and variable label lengths must be no more than 40 characters. The implementation guide should be used to ensure all required variables and information from the integrated SAP are included. For character variables, the integrated dataset must accommodate the longest value observed, up to a maximum of 200 characters.

Example: Variable label harmonization:

- Study A: TRT01P label = 'Planned Treatment'
- Study B: TRT01P label = 'Planned Treatment for Period 1'
- Study C: TRT01P label = 'Protocol Planned Treatment'

Standardized label for integrated dataset: 'Planned Treatment for Period 1' (consistent with CDISC ADaM standard variable label).

FORMAT STANDARDIZATION

Variable formats and display formats must be standardized to ensure consistent data representation across studies. This includes harmonizing numeric formats for continuous variables, date and datetime formats, and character formats for categorical variables. Date and datetime variables should adhere to ISO 8601 standards. Although date

imputation is typically performed at the individual study level, if the integrated SAP specifies otherwise, those instructions should be followed, and a unified imputation rule should be applied across all integrated studies.

Example: Date format harmonization:

- Study A stores dates as numeric format DATE9. (e.g. 25JAN2025).
- Study B stores dates as character in display as YYYY-MM-DD (e.g. 2024-12-25).
- Study C stores dates as numeric format YYMMDD10. (e.g. 2025-02-15).

A standardization decision must be made for the integrated dataset regarding variable representation. The team may elect to retain numeric values alone or to include both numeric and character formats.

CASE STANDARDIZATION

Case standardization is necessary to ensure consistency across studies for character variables. When a variable represents the same semantic meaning but appears in different casing formats (e.g., “M” vs. “MALE”), a decision must be made regarding the appropriate standardized format for the integrated dataset, based on CDISC conventions and sponsor preference.

Example: Treatment value case standardization:

- Study A: TRTP = ‘DRUG A 100 MG’, ‘PLACEBO’.
- Study B: TRTP = ‘Drug A 100 mg’, ‘Placebo’.
- Study C: TRTP = ‘Drug A 100 mg’, ‘PLB’.

For integrated dataset, standardize TRTP to ‘Drug A 100 mg’ and ‘Placebo’ and TRTPN (numeric treatment code) as: 1 = Drug A 100 mg, 2 = Placebo.

In addition, if the same variable name is represented as a character variable in one study and a numeric variable in another, it will need to be determined which format will be used in the integrated dataset. In some situations, both formats (numeric and character versions) may be retained to support traceability; however, the variable naming must remain consistent across all studies. Documentation of the chosen standardization approach shall be included in the specifications.

Example: Character vs numeric case standardization:

- Study A: Variable PNDTOT (Planned Total Dose) is stored as a character value – ‘3012’.
- Study B: Variable PNDTOT (Planned Total Dose) is stored as a numeric value - 3012.
- Study C: Variable PNDTOT (Planned Total Dose) is stored as a numeric value - 3012.

To ensure consistency across the integrated dataset, determine the appropriate standard format based on analytical needs and CDISC conventions. In most cases, numeric format is preferred for planned dose variables. Therefore, character values should be converted to numeric and appropriately documented in integrated data specifications.

VISIT MAPPING AND TIMING VARIABLES

Standardizing visit mapping and timing variables is useful for any time-based analyses in both ISS and ISE. Because different studies may use different visit schedules, naming conventions, and visit numbering schemes, these elements must be harmonized to support meaningful comparisons and pooled analyses. This process may involve aligning visit windows, standardizing visit names and visit numbers, and reconciling variations in scheduled, unscheduled, and early/late visits. Refer to the SAP for the integrated visit framework and collaborate with the sponsor to determine the appropriate visit windows and mapping approach that is consistent and meaningful across all studies.

VISIT WINDOW DEFINITION AND ALIGNMENT

Where applicable, define consistent visit windows to align equivalent time points across studies. When studies have different visit schedules, an integrated visit structure should be created that accommodates all study designs while maintaining analytical comparability. Follow the visit-mapping guidance outlined in the SAP and collaborate with the sponsor to confirm the final standardized visit window definitions and mapping rules.

Example: Visit window alignment across two studies:

AVISIT (Analysis Visit)	Study A Visit	Study B Visit
Baseline	Baseline	
Week 8	Week 8	
Week 16	Week 16	
Week 24	Week 24	
Week 32	Week 32	
Week 40	Week 40	
Week 48		Week 8
Week 46		Week 16

Week 64		Week 24
Week 72		Week 32
Week 80		Week 40
Continue…		*Continue…*

Some participants who roll over from Study A into Study B will continue with the integrated visit numbering instead of restarting at Baseline. For these subjects, the baseline will remain defined based on the first enrolled study, unless the SAP specifies a different approach for supportive analyses. This method ensures a consistent and accurate longitudinal representation of each participant's timeline across the integrated dataset.

MULTI-STUDY TIMING HANDLING

Harmonizing the calculation of study day and analysis day across all studies is needed for consistent timing-based analyses. Study day is typically derived as the number of days relative to the first dose date, with Day 1 corresponding to the first dose and Day –1 representing the day prior to first dose; per CDISC conventions, there is no Day 0. Analysis day (ADY) may differ from study day if a different reference date is defined in the SAP for analysis purposes.

Depending on the SAP requirements, the integrated first dose date typically refers to the subject's earliest exposure date across all contributing studies. This may differ from the first dose date recorded within a continuation of study alone. As a result, study day calculations must be adjusted at the integrated level to ensure chronological accuracy and prevent misalignment of exposure-based timelines across studies.

For example: Integrated first dose date:

- Subject 001 participated in Study A (first dose: 12JAN2009) and Study B (first dose: 14FEB2012).
- Integrated TRTSDT (Date of First Exposure to Treatment) = 12JAN2009 (at earliest exposure)

STUDYID variable can be used to track which studies the patient was enrolled in, and variables STUDID01, STUDID02, etc. can be used to track which study each record originated from for proper traceability and integrated analyses use.

ANALYSIS FLAG DERIVATION LOGIC

Analysis flags are variables that help identify records to be included in specific analyses. These flags must be consistently derived across all studies in the integrated dataset to ensure a valid pooled analysis. Commonly used analysis flags include baseline flags (ABLFL), analysis flags (ANL01FL, ANL02FL, etc.), and analysis value category variables (AVALCAT1, AVALCAT2, etc.).

Because individual studies may apply different definitions and derivation rules, there is no guarantee that variables with the same name across studies contain the same meaning. This is especially true for CDISC ADaM variables, where standard variable names such as ANLxxFL or AVALCATy may be defined differently from one study to another. Therefore, these flags require careful review and potential rederivation to ensure consistent definitions in the integrated dataset.

BASELINE FLAG LOGIC

The baseline flag (ABLFL) identifies the baseline assessment for each parameter. Baseline definitions can vary across studies—some may define baseline as the last non-missing value prior to the first dose, others may use a predefined visit (e.g., Day –1), or some may use an average of multiple values at the same visit. For the integrated dataset, a single standardized baseline definition must be applied according to the integrated SAP.

For subjects participating in multiple studies (e.g., rollover or continuation study participants), baseline must be recalculated using the integrated first dose date, which represents the earliest exposure date across all studies in which the subject participated. If the baseline definition is based on "last non-missing value before first dose," this must be rederived at the integrated level to align all subjects under a consistent baseline timepoint.

ANALYSIS FLAGS AND CATEGORIES

Analysis flags (ANL01FL, ANL02FL, etc.) identify specific subsets of records for targeted analyses. These may include flags for on-treatment periods, unique or priority records, specific analysis time windows, evaluable populations, or records that meet predefined clinical criteria. Analysis value categories (AVALCAT1, AVALCAT2, etc.) are used to group continuous values into clinically meaningful ranges to support categorical analyses.

Because the derivation rules for these flags and categories may differ across individual study data specifications, a detailed review is required to determine which analysis flags and value categories are appropriate for integration. In many cases, the integrated dataset will need its own standardized set of derived analysis flags and categories to

maintain across all studies. The final set of flags and categories should align with the requirements of the TFLs and follow the definitions outlined in the integrated SAP. All decisions and derivations shall be documented in the ADaM specification and Analysis Data Reviewer's Guide.

UNIQUE SUBJECT IDENTIFIER COORDINATION

Maintaining a consistent USUBJID (Unique Subject Identifier) across all studies is essential to prevent subject-level conflicts while preserving traceability back to each source study. This can be challenging when subjects enroll in multiple studies that use different identification structures.

Although the CDISC-recommended structure is typically STUDYID–SITEID–SUBJID, variations may occur depending on study design and potential identifier conflicts. When inconsistencies in USUBJID exist across studies, reference the pivotal study as the primary source, consult with the sponsor, and update the integrated analysis datasets to use the standardized USUBJID. The original study-specific subject identifier should be retained as a separate variable to maintain full traceability.

Example: USUBJID construction across studies:

- Study A: USUBJID = 'VWFABC123-11-345678'
- Study B: USUBJID = 'ABC123-11-345678'

For integrated datasets, USUBJID values will use Study B structure, and keep original USUBJID as OUSUBJID for Study A to maintain uniqueness and traceability.

PARAMETER AND UNIT RECONCILIATION

Laboratory data, vital signs, and other measurement domains require careful reconciliation of parameter definitions and units across studies. Different studies may measure the same clinical parameter using varying methods, report results in different units, or apply distinct naming conventions for the same parameter.

Variability often arises because different studies may measure the same clinical parameter using different methods, assign different parameter codes for the same test, or report results in non-standard or inconsistent units. To ensure harmonization in the integrated dataset, discrepancies must be identified, standardized, and documented.

PARAMETER CODE MAPPING

Ensure that PARAMCD (Parameter Code) and PARAM (Parameter) are harmonized across all studies, maintaining a strict 1:1 relationship between the two. This standardization process should identify parameters that represent the same measurement test but are coded differently across studies. For any discrepancies detected, document the full mapping specification, including all study-specific parameter codes and their standardized PARAM/PARAMCD assignments for the integrated dataset. In addition, consult with the sponsor to confirm all discrepancies and ensure alignment on the final standardized mappings.

Example: Laboratory parameter code harmonization:

- Study A uses PARAMCD = 'ALB' for Albumin.
- Study B uses PARAMCD = 'ALBC' for Albumin.

Both map to PARAM = 'Albumin' but require standardization to a single PARAMCD.

Example: Vital signs parameter harmonization:

- Study A uses PARAM = 'Systolic Blood Pressure (mmHg)' for PARAMCD = 'SYSBP'.
- Study B uses PARAM = 'Systolic Blood Pressure' for PARAMCD = 'SYSBP'.

Standardize the approach for including units in parameter labels for laboratory and vital signs data in alignment with the sponsor.

Additionally, be aware that PARCAT1 (Parameter Category 1) may also be inconsistently assigned across studies. For example, the HCG laboratory test might be categorized under 'Pregnancy' in one study but under 'Urinalysis' in another. Such differences should be documented in the specification and all discrepancies should be reviewed with the sponsor, and final alignment jointly confirmed to ensure accuracy and consistency.

UNIT STANDARDIZATION AND CONVERSION

All measurements for the same parameter must be converted to a standardized unit to ensure consistency across the integrated dataset. This is especially important for laboratory values, vital signs, and efficacy assessments, where different studies may have collected or reported results using different units. Unit conversions must be performed using scientifically validated conversion factors and fully documented in specifications. To maintain traceability, both the original values and the converted standardized values should be retained in the integrated dataset. Follow consistent variable naming conventions such as AVAL (analysis value), AVALC (analysis value, character), and keep

SDTM-based variables such as –ORRES (original result), –ORRESU (original units), –STRESN (standardized numeric result), –STRESC (standardized character result), and –STRESU (standard units).

Example: Laboratory unit conversion:

- Study A: Creatinine reported as mg/dL.
- Study B: Creatinine reported as mcmol/L.
- Study C: Creatinine reported as mg/dL.
- Conversion: To convert mcmol/L to mg/dL, divide by 88.42.

Since the standard measurement for Creatinine is mg/dL, all measurements reported in different units must be converted accordingly across all studies when preparing the integrated analysis dataset.

Example: Weight unit conversion:

- Study A: Weight in kilograms (kg).
- Study B: Weight in pounds (lbs)
- Study C: Weight in kilograms (kg).
- Conversion: To convert pounds to kilograms, multiply with 0.453592.

Standardize all weights to unit kg, when preparing the integrated analysis dataset.

PROGRAMMING SPECIFICATION DOCUMENTATION

When documenting decisions and mapping rules for ISS & ISE in the specification sheet, certain elements are required to be included in dataset specifications. These components ensure full transparency, traceability, and regulatory compliance within the integrated SDTM and/or ADaM specifications.

Each specification document should include the following:

- Dataset names and descriptions
- All variable names, labels, types, lengths
- Controlled terminology lists for certain variables
- Derivations and computation rules, written in a clear, understandable manner
- Variable origin, such as Derived, Assigned, or mapped from which individual study and variable

Value-level metadata, as required for SDTM findings or ADaM BDS structures

SUBMISSION REQUIREMENT CHECKLIST

Similar to the submission of individual study datasets, regulatory agencies such as the FDA and PMDA require that ISS/ISE data be submitted in accordance with the most current electronic Common Technical Document (eCTD) requirements. The following components are essential for a complete submission package:

INTEGRATED ADAM DATASETS

- ISS/ISE Datasets: These datasets must be prepared to conform to the Study Data Technical Conformance Guide. Adherence to this guide ensures the datasets meet regulatory expectations and maintain technical consistency across submissions.
- Validation Using Pinnacle 21 (P21): It is crucial to validate the integrated datasets using Pinnacle 21 to ensure compliance with CDISC standards. Each issue identified in the P21 validation report should be carefully reviewed and addressed as needed. Any errors or warnings that cannot be resolved must be documented in the Integrated Analysis Data Reviewer's Guide (iADRG), accompanied by clear and justified explanations.

DEFINE-XML DOCUMENTATION

- Metadata Documentation: A define-xml file must be included to document the metadata for each dataset and variable. This file should follow the appropriate version as specified in the latest data standards catalog and comply with the Define-xml specification.
- Define-xml Validation: Validation of the define-xml file using Pinnacle 21 is a mandatory step. The define-xml should typically be free from compliance issues; however, occasional codelist-related issues may persist from the study design or original datasets. Any such unresolved issues must be documented in the iADRG, with appropriate explanations provided.

INTEGRATED ANALYSIS DATA REVIEWER'S GUIDE (IADRG)

The iADRG should comprehensively document significant details and decisions associated with data harmonization for ISS/ISE analysis. PHUSE has introduced a specialized iADRG template for integrated data, which includes thorough completion guidelines and illustrative examples.

While the use of the PHUSE iADRG template is not required by regulatory authorities such as the FDA or PMDA, it has gained broad acceptance within the industry. This template offers regulatory reviewers a standardized and

organized approach specifically designed to facilitate the evaluation of submitted integrated data. Additionally, its guidance provides valuable considerations when planning data integration activities.

CONCLUSION

The successful preparation of Integrated Summary of Safety (ISS) and Integrated Summary of Efficacy (ISE) datasets requires careful planning, systematic harmonization, and close collaboration among programmers, statisticians, and data managers. Working closely with statisticians and thoroughly reviewing both the Statistical Analysis Plan (SAP) for the integrated summary and the SAPs for individual studies as well as the study-specific annotated CRFs can be fundamental to understanding the analytical requirements and ensuring that integrated datasets support all planned analyses.

Often the most time-consuming component of the ISS/ISE process is the creation of the integrated datasets themselves. This work becomes particularly challenging if having to integrate legacy datasets, reconciling different SDTM/ADaM versions, addressing variations in controlled terminology, or managing subjects who participated in multiple studies. Without a structured harmonization strategy with a practical checklist, these inconsistencies can lead to incorrect pooled analyses, data quality issues, or delays in regulatory submissions.

SUMMARY OF PRACTICAL CHECKLIST APPLICATION

The checklist presented in this paper can be applied following this systematic approach:

1. **Planning and Assessment Phase**:
   Identify studies to be included, collect all implementation guides, evaluate data availability and quality, determine whether integration will occur at the SDTM-level or ADaM-level, and establish timeline and resource requirements.
2. **Dataset and Variable Level Harmonization**:
   Align dataset structures across studies; standardize variable names, types, lengths, and labels; harmonize formats and casing rules; and ensure consistent controlled terminology.
3. **Content Harmonization**:
   Up-version coding dictionaries (MedDRA, WHODrug, CTCAE (Common Terminology Criteria for Adverse Events)), reconcile parameters and units; align visit mapping and timing variables, if applicable; standardize analysis flag derivation logic; and ensure consistency in subject identifiers values across studies.
4. **Implementation and Quality Control**:
   Develop and execute programming specifications, perform data integration, conduct thorough quality control checks, verify data integrity by reconciling counts against individual studies, and ensure traceability is maintained throughout the process.
5. **Validation and Documentation**:
   Perform Pinnacle 21 validation and resolve all feasible issues; prepare comprehensive documentation (cSDRG, iADRG, Define.xml); document all harmonization decisions and mapping rules; and collaborate with the sponsor to ensure their review and approval along every step.

Engaging experienced professionals who have prior ISS/ISE submission expertise can provide valuable insight and guidance when navigating complex integration challenges.

By applying this checklist and following best practices for data harmonization, programming teams can develop high-quality integrated datasets that adhere to CDISC standard, satisfy regulatory expectations, and provide a solid foundation for comprehensive safety and efficacy analyses. For convenience, a summary checklist is provided in the Appendix to support consistent implementation. Ultimately, the goal is to deliver robust integrated summaries that can support regulatory decision-making and drug approval decisions and help advance patient care.

REFERENCES

U.S. Department of Health and Human Services Food and Drug Administration, Center for Drug Evaluation and Research (CDER) and Center for Biologics Evaluation and Research (CBER). “Study Data Technical Conformance Guide”. FDA website. March 2022. Available at: https://www.fda.gov/industry/fda-data-standards-advisory-board/study-data-standards-resources

Donthi B, Qi L. Best Practices for ISS/ISE Dataset Development. PharmaSUG 2019. Available at: https://pharmasug.org/proceedings/2019/AD/PharmaSUG-2019-AD-299.pdf

Patel PH. Navigating FDA Requirements: ISS/ISE Build Strategies. PhUSE Connect 2019. Available at: https://www.lexjansen.com/css-us/2019/PP34.pdf

Pothuri K, Vajjala K. Overcoming Challenges of your Integrated Summary ISS/ISE Submissions. PhUSE 2022. Available at: https://www.lexjansen.com/phuse/2022/ds/PAP_DS08.pdf

U.S. Food and Drug Administration. Guidance for Industry: Integrated Summaries of Effectiveness and Safety: Location Within the Common Technical Document. 2009. Available at: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/integrated-summaries-effectiveness-and-safety-location-within-common-technical-document
Tinazzi A., Cedric M. An FDA Submission Experience Using the CDISC Standards. PhUSE 2017. Available at: https://www.lexjansen.com/phuse/2017/rg/RG04.pdf
SDSP (Study Data Standardization Plan) Case Studies and Considerations. PharmaSUG 2019. Available at: SDSP (Study Data Standardization Plan) Case Studies and Considerations
ADaM Data Structures for Integration. Available at: https://www.cdisc.org/standards/foundational/adam
MedDRA Maintenance and Support Services Organization (MSSO). Available at: https://www.meddra.org/
WHO Drug Global. Available at: https://www.who-umc.org/whodrug/
Clinical Data Interchange Standards Consortium (CDISC). CDISC Standards. Available at: https://www.cdisc.org/standards

ACKNOWLEDGMENTS

The author would like to extend appreciation to colleagues, subject-matter experts, and collaborators who generously shared their time, expertise, and perspectives through the writing and revision process. Your valuable feedback and thoughtful reviews are instrumental in strengthening the clarity, accuracy, and completeness of this work. This paper reflects the collective efforts of all who offered support, encouragement, and constructive input.

RECOMMENDED READING

- Integrated Analysis Data Reviewer's Guide Completion Guidelines
- CDISC Analysis Data Model Implementation Guide
- CDISC Study Data Tabulation Model Implementation Guide: Human Clinical Trials

CONTACT INFORMATION

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

Name: Lily Cheng
Company: BioPier LLC, a Veramed Company
Email: lcheng@biopier.com
Location: Burlington, MA

Name: Qiong Wei
Company: BioPier LLC, a Veramed Company
Email: qwei@biopier.com
Location: Burlington, MA

Name: Ji Qi
Company: BioPier LLC, a Veramed Company
Email: jiqi@biopier.com
Location: Burlington, MA

APPENDIX: ISS/ISE PRACTICAL INTEGRATION CHECKLIST

Summary Checklist for Integrated Summary of Safety and Efficacy

1. PLANNING AND ASSESSMENT	
☐	Identify studies to include and select pivotal study as reference
☐	Collect all dataset standards (legacy formats, SDTM IG versions, ADaM IG versions) and review the integrated SAP
☐	Determine the integration approach: (SDTM (s) to iADaM; ADaM(s) to iADaM; SDTM(s) to iSDTM then to iADaM)
☐	Confirm early regulatory expectations (FDA, PMDA) for ISS/ISE submission
☐	Establish project timeline, delivery milestones, and resource requirements
2. DATASET AND VARIABLE ALIGNMENT	
☐	Align dataset structures and variable specifications to the target IG version
☐	Standardize variable attributes (names, types, lengths, labels, casing, formats)
☐	Harmonize variable names (max 8 characters), variable label (max 40 characters), and variable lengths (max 200 characters) to accommodate all values
3. CONTENT HARMONIZATION	
☐	Document MedDRA, WHODrug, CTCAE, and any other dictionary versions per study, and perform up-versioning if necessary
☐	Map integrated treatment groups and standardize terminology (for dose groups, causality categories, demographic information, etc.)
☐	Align visit mapping and timing variables across all contributed studies (analysis visit framework, study day rules, rollover subject adjustments, etc.)
4. SUBJECT IDENTIFIERS	
☐	Standardize USUBJID structure (consult sponsor if conflicts exist) and ensure USUBJID uniqueness across all integrated datasets
☐	Maintain STUDYID and origin-tracking variables (e.g. STUDID01, STUDID02) for traceability to source study
5. DERIVATION LOGIC	
☐	Define integrated baseline (ABLFL) using integrated first-dose date, if applicable (noting rollover subjects)
☐	Standardize ANLxxFL derivation logic across all contributing studies
☐	Derive any analysis value categories (AVALCATx) per SAP requirements
☐	Define and derive population flags (SAFFL, ITTFL, FASFL, PPFL, etc.) per SAP
☐	Reconcile parameters (PARAM/PARAMCD), units, and any SAP defined reference ranges

6. PROGRAMMING AND QC	
☐	Develop complete programming specifications documenting all harmonization decisions and derivation rules, and ensure specifications meet expectations for regulatory agency reviewer reproducibility.
☐	Maintain complete traceability to source data
☐	Perform independent programming validation
7. DOCUMENTATION PREPARATION	
☐	Prepare iADRG, cSDRG (if appliable), and Define.xml
☐	Validate Define.xml and datasets using Pinnacle 21
☐	Address Pinnacle 21 issues when possible; document unresolved compliance issues in the iADRG with clear and justified explanation.
☐	Run P21 validation on all integrated datasets and Define.xml; generate final Pinnacle 21 validation report for submission.
8. SPONSOR REVIEW AND APPROVAL	
☐	Submit integrated datasets and documentation to sponsor for review
☐	Address sponsor feedback and implement requested changes
9. REGULATORY SUBMISSION	
☐	Organize datasets following eCTD Module 5 folder structure
☐	Finalize all documentation (iADRG, cSDRG (if applicable), Define.xml)
☐	Complete regulatory submission checklist and confirm submission meets regulatory guidance (Technical Conformance Guide, Data Standards Catalog, Define-xml specification, etc.)
☐	Archive final deliverables, programming code, and communication trace.

Note: This checklist should be used in conjunction with the Integrated Summary Statistical Analysis Plan (SAP), Study Data Standardization Plan (SDSP), and all applicable CDISC standards (SDTM, ADaM), and regional regulatory guidance (Study Data Technical Conformance Guide, Data Standards Catalog, etc.). All harmonization decisions must be documented in the Integrated Analysis Data Reviewer's Guide (iADRG) and programming specifications.