

Paper SI09

Choosing the right path to follow when integrating ADaM

Laura Phelan, Cytel Inc, Paris, France
Magalie Gallet, Cytel Inc, Geneva, Switzerland

ABSTRACT

Performing an integrated ADaM database (“pooling”) is not an uncommon undertaking in the Biotech/Pharma business, but getting it right remains a challenge when the analysis plan is complex and does not easily fit within the CDISC structure.

While there is a growing reference of presentations, white papers, and the ADaM IGs are relatively stable with regards to the most commonly included domains, rarely can one find examples which exactly match your pooling needs. We would like to share our recent submission experience and describe the many challenges we had to face and the solutions we found, with a special focus on:

- Designing ADaMs where subjects are analyzed in multiple treatment cohorts of varying duration
- The pros and cons of different approaches
- Will the future guidelines have more precise recommendations for this type of analyses?

INTRODUCTION

CDISC provides a comprehensive guidance suited to most of the analyses that can currently be foreseen. However, when it comes to more complex analyses such as a pooling, we sometimes need to get creative in applying the model.

Before describing the complexity of the analysis, the question of whether to pool at the SDTM or ADaM level was discussed. In this particular submission, because the migration to SDTM was done retrospectively, the studies were harmonized at the SDTM level, meaning that ADaMs could be easily pooled from the individual study SDTMs without needing to pool them first.

We faced such a challenge in a recent submission. In our integrated ADaMs, we needed to include several studies where subjects belonged to more than one treatment group and more than one cohort. These subjects came from studies with large databases, and to deal with this, the whole structure of the analysis datasets from ADSL to the BDS, had to be re-thought to conform, as much as possible, to the CDISC standard model and to allow for efficiency and flexibility in case of late changes to the analysis.

In this paper we will present the different approaches we could have applied with regards to structuring the pooled datasets, the solution we chose, and why. We will then discuss the advantages and disadvantages of the chosen approach. As an integrated CDISC guideline is under review at the time of writing this paper, we will also discuss how this guideline could have helped for this specific analysis and how it would have impacted the structure of our ADaMs.

ANALYSIS DESCRIPTION

A leading global science and technology company, further referenced in this document as the ‘sponsor’, was the pharma company responsible for a planned FDA CDER submission in 2018. As a Data Service Provider Cytel Inc. we were contracted to perform the CDISC SDTM conversion and the Integrated Summary of Safety (ISS) analysis. In this document, we will focus on the ISS analysis and specifically on the CDISC ADaMs. We were responsible for the ISS CDISC full package, including the ADaMs, Tables, Listings and Figures (TLFs), define.xml and Analysis Study Data Reviewers Guide (adrg).

The ISS ADaM structure followed the latest CDISC ADaM guidelines available prior to the submission, namely:

- ADaM Implementation Guideline v1.1
- ADaM Structure for Occurrence Data (OCCDS) v1.0
- ADaM Time-to-Event Analyses v1.0
- Sponsor ADaM Guideline

Pinnacle conformance was run using Pinnacle 21 Community version 2.2.0.

The ISS design for this analysis was a pooling including subjects from ten studies with safety data from phases II/III studies and an ongoing observational (registry) study. Almost half of the subjects participated in at least two studies and approximately a third of them participated in three of the ten studies. It included more than 2000 subjects with data spanning up to 10 years.

The analysis was designed as such that there were multiple groupings of studies, referenced in this document as “cohorts”, with subjects contributing sometimes to more than one treatment within a cohort. One subject could be reported in more than one cohort with different treatment groups and different treatment periods between the cohorts. We limit this paper to three of the most distinctly different cohorts: 1, 2 and 3. The approach used for these served as a model for the remaining cohorts.

To face this complex ISS design and comply as much as possible with the CDISC guidelines, we asked ourselves different questions such as:

- How to structure ADSL? What CDISC variables to use to fit the analysis?
- How to incorporate the cohorts in the other ADaM Classes: OCCDS and BDS structured ADaMs.
- How to ensure traceability when pooling multiple studies?
- How to manage pooling studies with different versions of the coding dictionaries?

Subjects could have taken

- Active treatment in both their first and second studies: subject 1 referred to as [s1]
- Placebo in their first study and active treatment drug in their second study: subject 2 referred to as [s2]
- Active treatment or Placebo in their first study and never enter a second study

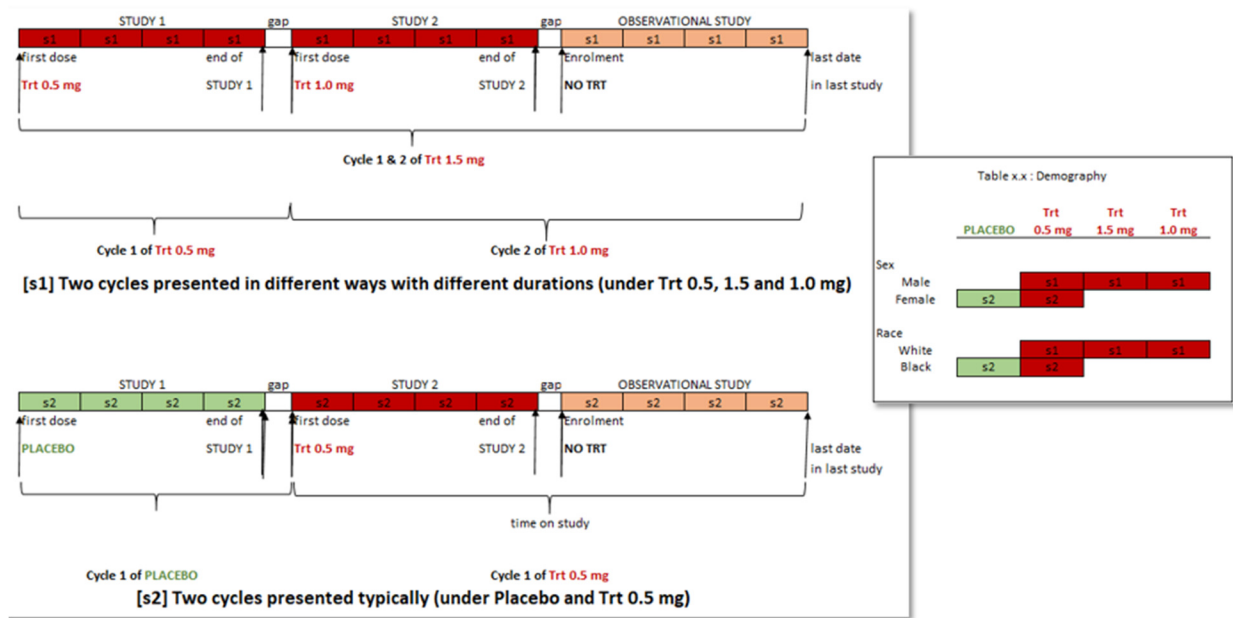
In all three scenarios, subject may or may not have continued to the observational study. This paper will describe the more complex cases where subjects participated in a second study.

COHORT 1

In the Cohort 1 summary tables, subjects who received active treatment in both periods of exposure (referenced in this document as “cycles”) [s1], were summarized under up to three treatment groups, depending on the cycles included. Subjects who received Placebo followed by active treatment [s2] were summarized under their two treatment groups, Placebo and active treatment. Subjects receiving only one cycle, were, naturally, summarized only under that cycle’s treatment.

Figure 1a shows Cohort 1 study design and Figure 1b shows layout example for a summary table.

Figure 1a



s1=Subject 1, s2=Subject 2

Figure 1b

Table x.x: Demography – Cohort 1 – Subject 1 [s1] and Subject 2 [s2]

Parameter	Statistics	Placebo (N=xxx)	Trt 0.5 mg (N=xxx)	Trt 1.5 mg (N=xxx)	Trt 1.0 mg (N=xxx)
Sex, n(%)	N (missing)	xxx (x)	xxx (x)	xxx (x)	xxx (x)
	Male	xxx (xx.x)	xxx (xx.x)	xxx (xx.x)	xxx (xx.x)
	Female	xxx (xx.x)	xxx (xx.x)	xxx (xx.x)	xxx (xx.x)
Race, n(%)	N (missing)	xxx (x)	xxx (x)	xxx (x)	xxx (x)
	White	xxx (xx.x)	xxx (xx.x)	xxx (xx.x)	xxx (xx.x)
	Black or African American	xxx (xx.x)	xxx (xx.x)	xxx (xx.x)	xxx (xx.x)
	Asian	xxx (xx.x)	xxx (xx.x)	xxx (xx.x)	xxx (xx.x)
	Other	xxx (xx.x)	xxx (xx.x)	xxx (xx.x)	xxx (xx.x)

[s2] contributes to 2 treatment groups

[s1] contributes to 3 treatment groups

COHORT 2

In the Cohort 2 summary tables, [s1] was summarized once including all their time in Cycle 1 and 2 – like “Trt 1.5 mg” in Cohort 1, just labelled “Trt”. [s2] was summarized exactly as in Cohort 1 but again, with active treatment labelled differently: “Trt” instead of “Trt 0.5 mg”, as in Figure 2a and 2b.

Figure 2a

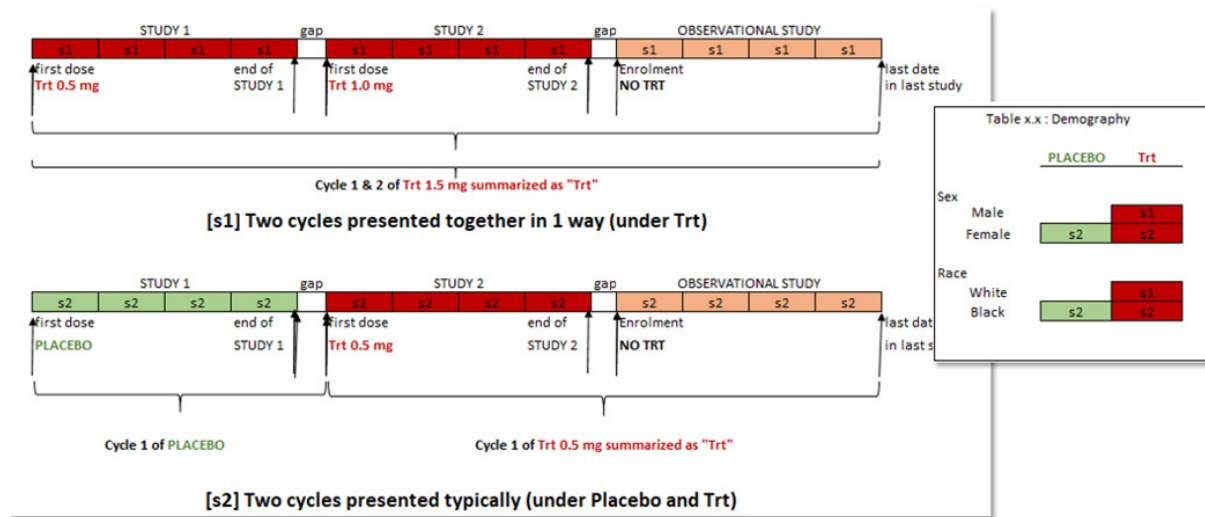


Figure 2b

Table x.x: Demography – Cohort 2 – Subject 1 [s1] and Subject 2 [s2]

Parameter	Statistics	Placebo (N=xxx)	Trt (N=xxx)
Sex, n(%)	N (missing)	xxx (x)	xxx (x)
	Male	xxx (xx.x)	xxx (xx.x)
	Female	xxx (xx.x)	xxx (xx.x)
Race, n(%)	N (missing)	xxx (x)	xxx (x)
	White	xxx (xx.x)	xxx (xx.x)
	Black or African American	xxx (xx.x)	xxx (xx.x)
	Asian	xxx (xx.x)	xxx (xx.x)
	Other	xxx (xx.x)	xxx (xx.x)

[s2] contributes to
2 treatment groups

[s1] contributes to
1 treatment group

COHORT 3

Cohort 3 is the same as Cohort 1 with regards to how treatment groups were labelled, but the duration for [s1] and [s2] is shorter in Cycle 1, as seen in Figure 3a and 3b: the gap from the end of Study 1 to the start of Study 2 was not included.

Figure 3a

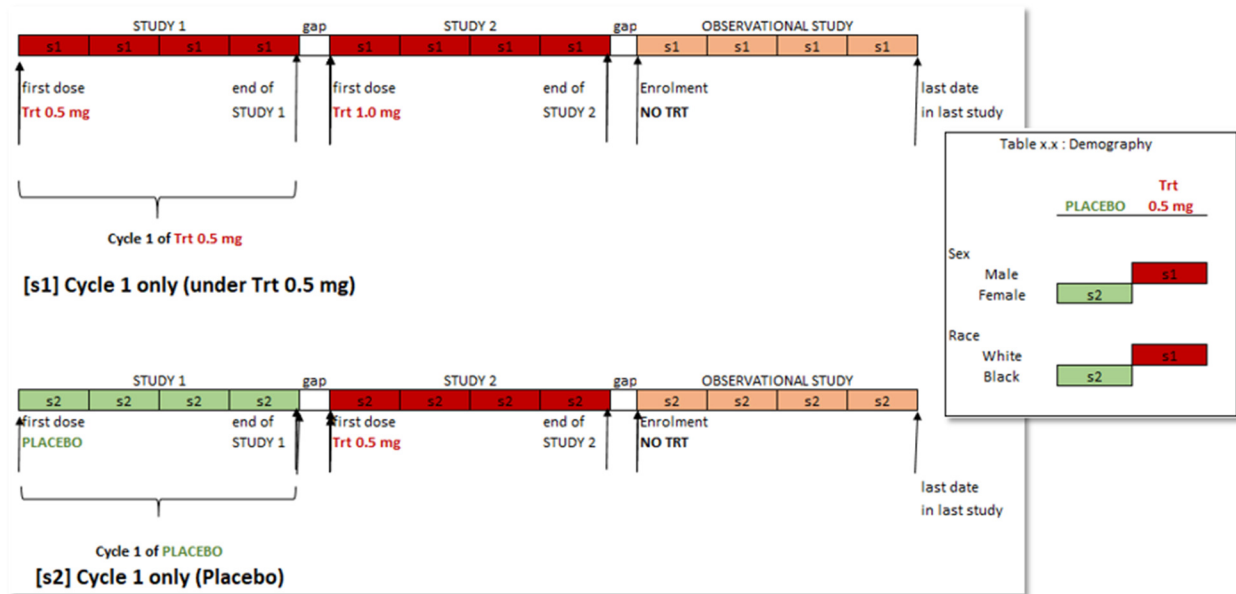


Figure 3b

Table x.x: Demography – Cohort 3 – Subject 1 [s1] and Subject 2 [s2]

Parameter	Statistics	Placebo (N=xxx)	Trt 0.5 mg (N=xxx)
Sex, n(%)	N (missing)	xxx (x)	xxx (x)
	Male	xxx (xx.x)	xxx (xx.x)
	Female	xxx (xx.x)	xxx (xx.x)
Race, n(%)	N (missing)	xxx (x)	xxx (x)
	White	xxx (xx.x)	xxx (xx.x)
	Black or African American	xxx (xx.x)	xxx (xx.x)
	Asian	xxx (xx.x)	xxx (xx.x)
	Other	xxx (xx.x)	xxx (xx.x)

[s2] contributes to
1 treatment group

[s1] contributes to
1 treatment group

Another way of looking at this, to perhaps understand more clearly, is by subject as in Figure 4.

Figure 4

Subject s1

	STUDY 1	STUDY 2	OBSERVATIONAL
Cohort 1	Trt 0.5 mg	Trt 1.0 mg	
	Trt 1.5 mg		
Cohort 2	Trt		
Cohort 3	Trt 0.5 mg (excl. gap)		

Subject s2

	STUDY 1	STUDY 2	OBSERVATIONAL
Cohort 1	Placebo	Trt 0.5 mg	
Cohort 2	Placebo	Trt	
Cohort 3	Placebo (excl. gap)		

HOW TO STRUCTURE ADAMS

To define how to best build ADaMs, we had to consider the setup chosen from A to Z, meaning working backwards from TLF to BDS/OCCDS structure back to ADSL in order to answer the following questions:

1. Would the BDS structure allow easy TLF “one proc away” with so many cohorts?
2. Would all the cohorts fit into one ADaM?
3. Would ADaMs (i.e ADAE or ADLB) work with duplicated records for so many treatment groupings/durations?
4. Would Pinnacle complain about how the BDS/OCCDS ADaMs were “mapped” from ADSL?

ADSL

ADSL is key. It is called in almost all programs. If it is not designed properly it may generate difficulties in building the BDS/OCCDS ADaMs, and later when used in TLF programs. Therefore, as with probably most complex pooled analyses, designing ADSL proved to be very time consuming from the start.

In ADaM IG version 1.1, ADSL is structured with **one record per subject** and the different standard treatment groups variables proposed are:

- ✓ TRxxPGy/AGy: Planned/Actual Pooled Treatment for Period xx
- ✓ TRTxxP/A: Planned/Actual Treatment for Period xx

Therefore we looked at their different definitions to see which one fit the best:

- TRxxPGy IG says “Planned pooled treatment y for period xx. Useful when planned treatments (TRTxxP) in the specified period xx are pooled together for analysis according to pooling algorithm y. Each value of TRTxxP is pooled within at most one value of TRxxPGy.”
- TRTxxP IG says: “Subject-level identifier that represents the planned treatment for period xx.”

At first glance, the existing TRxxPGy/AGy seemed they might have been appropriate for our needs, but when we simulated how it would work in ADSL based on [s1], [s2] and another hypothetical subject [s3], it quickly became obvious that it would not suit (see Figure 5).

Figure 5

ROW	STUDYID***	USUBJID	TRT01A	TRT02A	TR01AG1 (Actual pooled treatment 1 based on period 01 TRT01A)	TR02AG1 (Actual pooled treatment 1 based on period 02 TRT02A)
1	which to use?	s1	Trt 0.5 mg	Trt 1.0 mg	Trt	Trt
2	which to use?	s2	PLACEBO	Trt 0.5 mg	PLACEBO	Trt
3	which to use?	s3	Trt 0.5 mg		Trt	

***see STUDYID/ASTUDYID section further down in this paper for details

TRxxPGy/TRxxAGy are grouping variables over specific periods and not grouping several periods i.e. TRT01A and TRT02A could not be “grouped” under TRxxAGy. These variables are more appropriate to group together treatment groups of a same compound with different dosage than periods. Another approach was necessary.

As the IG 1.1 did not seem to propose other variables for our analysis needs, using other treatment variables not typically intended for pooling was considered: TRTxxP/TRTxxA (see Figure 6).

Figure 6

			Cohort 1			Cohort 2		Cohort 3
ROW	STUDYID	USUBJID	TRT01A	TRT02A	TRT03A	TRT04A	TRT05A	TRT06A
1	which to use?	s1	Trt 0.5 mg	Trt 1.5 mg	Trt 1.0 mg	Trt	null	Trt 0.5 mg
2	which to use?	s2	PLACEBO	Trt 0.5 mg	null	PLACEBO	Trt	PLACEBO

With this approach, all treatment groups/periods per subject were managed and could be easily transposed in a BDS structure, allowing tables to be one proc away. It was consistent with ADaM IG, and Pinnacle did not give errors.

It may not have been the exact intent for these variables. The intent of these variables is to describe the planned/actual treatment of a specific period. With this approach, TRTxxP/xxA were not consecutive, for example. However, it adhered to a standard not designed for such an integrated pooling. The team agreed that the simplicity of this approach could work, and even may have been easier for a reviewer to understand once a clear explanation in the reviewers guide was provided. This was not deviating or going against any ADaM conformance rule, but clearly the model was “forced” (and we were aware of this).

Other approaches such as the use of sub-periods variables (PxxSwSDT/PxxSwEDT) or phase variables (APHASE, PHxxSDT/PHxxEDT) were discussed, but none of them allowed for a different treatment group variable. Due to the different cohorts design and durations, the use of these variables would have made more complex the reading of ADSL and the creation of BDS and TLFs. In big integrated analyses like these ones (i.e. containing different cohorts, treatment groups and durations) it is important to use a method than can be generalized for all analyses that are performed.

BDS/OCCDS DATASETS

As mentioned above, ADSL is not exactly user friendly in identifying easily to which cohort and treatment group a subject belonged. To avoid having to restructure the ADSL in every other ADaM and/or output program and stay one-proc away, we needed another “analysis-ready” ADaM with one row per subject per cohort per treatment group/period.

We also had to consider that several baseline characteristics changed depending on the period, such as Age at baseline. Adding multiple definitions for each of these variables would have led to 100's of variables in ADSL, which is not in anybody's interest. ADBL (AD Baseline) was the BDS used.

This dataset became the key dataset in the analysis. It was the basis of all ADaMs and TLFs, summarizing cohorts, treatment groups and corresponding period durations.

Size became an issue for consideration. The FDA size limit for datasets is 5 gigabytes (GB). Given the multiple cohorts we had, this limitation would have been exceeded notably for the larger datasets like ADLB. Moreover, such large datasets take longer to read and slow the system. In order to speed the analysis run time and facilitate the review process, datasets were split by cohort. ADBL was split per cohort as were all the following ADaMs. This had the advantage that if a definition for any particular cohort changed, we could run the datasets and outputs only for that cohort, thus saving time.

FUTURE/NOT RELEASED GUIDELINE

To meet the needs for this analysis, ideally ADSL could have been structured to allow one row per subject, per cohort and per treatment group/period. The new integrated ADSL (IADSL) would have allowed this (Figure 7), but this wasn't available at the time of pooling, and at the time of writing this paper, it is still in draft status.

IADSL

The planned structure for this dataset is *"one record per subject per pool"*. Or, and as was the case in this pooling, *"...where the integrated statistical analysis plan (SAP) calls for a subject to be counted more than once for a given pool - 1 record per subject per pool per subject identifier for the pool"*.

New variables in this guideline which allow for uniqueness and identification of pooling treatments and periods include:

- ✓ STUDIES IG says: *"...is used if any row in integrated ADSL includes data from more than 1 study", and "If STUDIES is present, do not include STUDYID"*
- ✓ POOL IG says: *"This is a required key variable in the IADSL class to indicate the pool to which a record belongs (e.g., PLACEBO-CONTROLLED, ACTIVE DRUG). The pools are as defined in the integrated SAP. An overall pool record is required when using the IADSL class."*
- ✓ PSUBJID IG says: *"This variable is required if the structure of integrated ADSL is not 1 record per unique subject per pool (i.e., when USUBJID is insufficient to uniquely identify a subject within a pool for analysis reasons)."*
- ✓ PSREAS IG says: *"A brief description of the reason that a unique subject has multiple records within a pool (e.g., "Re-enroller same study different site," "Subject participated in more than 1 pivotal study")."*
- ✓ NUMSTUDY IG says: *"..to identify the number of studies included in each row. The value of NUMSTUDY will match the number of study identifiers listed in STUDIES."*

Figure 7 shows an example on how this IG could have been interpreted for this pooling.

Figure 7

ROW	USUBJID	STUDIES	POOL	PSUBJID	PSREAS	NUMSTUDY	TRT01A
1	s1	STUDY 1	COHORT 1	s11	Cohort 1 Treatment 1 definition	1	Trt 0.5 mg
2	s1	STUDY 1 + STUDY 2 + OBS	COHORT 1	s12	Cohort 1 Treatment 2 definition	3	Trt 1.5 mg
3	s1	STUDY 2 + OBS	COHORT 1	s13	Cohort 1 Treatment 3 definition	2	Trt 1.0 mg
4	s1	STUDY 1 + STUDY 2 + OBS	COHORT 2			3	Trt
5	s1	STUDY 1	COHORT 3			1	Trt 0.5 mg
6	s1	STUDY 1 + STUDY 2 + OBS	OVERALL			3	
7	s2	STUDY 1	COHORT 1	s21	Cohort 1 Treatment 1 definition	1	PLACEBO
8	s2	STUDY 2 + OBS	COHORT 1	s22	Cohort 1 Treatment 2 definition	2	Trt 0.5 mg
9	s2	STUDY 1	COHORT 2	s21	Cohort 2 Treatment 1 definition	1	PLACEBO
10	s2	STUDY 2 + OBS	COHORT 2	s22	Cohort 2 Treatment 2 definition	2	Trt
11	s2	STUDY 1	COHORT 3			1	PLACEBO
12	s2	STUDY 1 + STUDY 2 + OBS	OVERALL			3	

The new guidance also states that “an overall pool record is required for each unique subject. This record describes the entirety of the subject’s treatment experience and allows relevant data to be slotted appropriately”. Given that IADSL was not used in this instance, it is unclear how exactly these “OVERALL” rows should be populated, notably with regards to treatment variables and start-end dates.

IBDS/IOCCDS Structure

The integrated BDS structure (IBDS) is designed to allow one record per pool per subject per parameter. The integrated OCCDS structure (IOCCDS) is designed to allow one record per pool per subject per event.

The IBDS structure is, in fact, almost unchanged compared to the existing BDS except for the definition of some variables. For example, BDS ABLFL is baseline per parameter per baseline type (BASETYPE), and in IBDS it is baseline per parameter per baseline type per pool. Similarly, some IOCCDS variables such as occurrence flags, differ from OCCDS ones in the sense that they are designed to be derived per pool in addition to per subject. Of note, in IOCCDS, only records where the AE.STUDYID is listed in IADSL.STUDIES are supposed to be kept – the IG even provides some SAS code for this selection.

The IADSL variables STUDIES, POOL and PSUBJID are copied to the IBDS/IOCCDS Structures, and according to the draft guideline, only the pools needed in analysis are to be copied, and not POOL=OVERALL.

Having these guidelines in place at the time of this pooling analysis would clearly have saved us time when designing ADSL and BDS datasets. Given that ADSL is usually the first dataset looked at when performing analysis or review, having all the subject treatment and periods summarized there brings added value. Regardless of which guideline was being used, other decisions such as managing traceability had to be taken. This is described in the next section.

TRACEABILITY WHEN POOLING MULTIPLE STUDIES

One of the fundamental principles of ADaMs is traceability, but when it comes to pooling, this traceability principle is not always easy to follow.

STUDYID/ASTUDYID variables

As we have seen in the previous chapter, IADSL allowed for multiple rows per subjects, per pool and per treatment groups within a pool. The variable STUDIES can contain more than one study identifier, whereas in the ADaM Ig 1.1, only one study can be included in STUDYID.

The conformance rules (in Pinnacle21), fire a cross reference ERROR (AD0256) that “USUBJID value does not exist in the ADaM ADSL domain” if your ADSL.USUBJID and STUDYID do not match the USUBJID and STUDYID in another ADaM. This is because USUBJID and STUDYID are key variables from ADSL to the other ADaMs.

However, if this rule is implemented, if for example, the first study entered (STUDY 1) is used to populate STUDYID, then traceability is lost when any parameter or event is derived across more than one study. For example, the BDS parameter (PARAM) “Time on Study “ was derived taking dates from STUDY 1 up to the end of the Observational study. Which STUDYID of the three to use?

To satisfy both constraints:

- STUDYID took the first study the subject entered i.e. STUDY 1
- a new variable ASTUDYID=“Analysis Study Identifier” was added which contained the STUDYID from which records originated (i.e. a lab sample, an adverse event...)

This new variable ensured traceability between SDTM and ADaMs. When the parameter or event came from only one study, that study identifier was used. It was null when they came from more than one study as in the above example for Time on study. It was also null if, for example, records from different studies were averaged and added as a new record.

DATAPOINT TRACEABILITY VARIABLES

ASTUDYID was not sufficient in some ADaMs to ensure full traceability. ASTUDYID should be used, and not STUDYID, to identify the SDTM study package. Some ADaMs were combining multiple sources of SDTMs and/or ADaMs, for example Time-to-event datasets. To ensure the full traceability, as in any conventional study, the variables SRCDOM, SRCVAR and SRCSEQ were also used.

MANAGING DIFFERENT DICTIONARIES

Anyone who has integrated studies is likely to have seen that frequently the AE, MH, CM, etc., SDTM domains are coded with different MedDRA, WHODrug dictionary versions across studies. This, understandably, is because studies were reported at different times, and the latest dictionary from that time is used.

This integration was no exception. The up versioning of dictionaries could have been done at the SDTM level. However, we chose to do it in the ADaMs, given there was no plan to change or integrate any SDTMs and we wanted to maintain traceability back to the original CSR. Also, fortunately for us, the CDISC OCCDS v1.0 had a recommendation that could be followed for this (see Figure 8).

Figure 8

Table 3.2.10.1 Original or Prior MedDRA Coding Variables

Variable Name	Variable Label	Type	Codelist / Controlled Terms	Core	CDISC Notes
DECDORGw	PT in Original Dictionary w	Char	MedDRAw*	Perm	Original preferred term coding of XX --TERM using MedDRA or other dictionary version X.X.
BDSYORGw	SOC in Original Dictionary w	Char	MedDRAw*	Perm	Original body system coding of XX --TERM using MedDRA or other dictionary version X.X.
HLGTORGw	HLGT in Original Dictionary w	Char	MedDRAw*	Perm	Original HLGT coding of XX --TERM using MedDRA or other dictionary version X.X.
HLTORGw	HLT in Original Dictionary w	Char	MedDRAw*	Perm	Original HLT coding of XX --TERM using MedDRA or other dictionary version X.X.
LLTORGw	LLT in Original Dictionary w	Char	MedDRAw*	Perm	Original LLT coding of XX --TERM using MedDRA or other dictionary version X.X.
LLTNORGw	LLT Code in Original Dictionary w	Char	MedDRAw*	Perm	Original LLT code of XX --TERM using MedDRA or other dictionary version X.X.

* For each version of an external dictionary, a different reference name must be used. The individual reference names will point to a dedicated section in the data definition file where all external dictionaries used in the analysis are listed, including dictionary name and version.

Interestingly, the IOCCDS recommends a different approach when up versioning is done in integrated ADaMs. There, the original coding variable should be prefixed with an “O” and not suffixed with “ORGw”.

Another impact of this up versioning is that coding variables like AEDECOD and AEBODSYS could not be used in integrated ADaMs. Remember the ADaM principle of harmonization: “same name, same meaning, same values”.

If SDTM.AEDECOD is not equal ADAE.AEDECOD, traceability is lost. Instead, these were mapped to ADECOD/ABODSYS borrowed from the IG where an A- prefix is used to denote analysis variables. This necessary

deviation from the expected OCCDS structure is detected by Pinnacle - “AD0047: Required variable is not present” – and was explained in the reviewers guide.

Slightly less typical in this pooling, was the situation where Adverse events were coming from different SDTM sources including:

- the usual AE domain
- a ZA domain: as per sponsor decision, after some of the studies had been locked and reported, pharmacovigilance updated some of the adverse event information which was captured in a custom domain following the AE structure but with ZA prefix, e.g. ZATERM for AETERM.
- a CE domain: this was included from the observational study only, where events were mapped as clinical events in SDTM, but were considered as Adverse events in analysis

This meant that not only the standard coding variables, but some descriptive variables, had to be renamed. For example: AESER to ASER, AESEV to ASEV. As per the OCCDS v1.0 for ASEV: “Apply imputation rules for missing severity of adverse events as specified in the SAP or metadata.”, Pinnacle checks that this variable is always filled. Otherwise the error “AD0282A: Analysis severity ASEV value is null” is generated. In this pooling, if the SDTM AESEV was null, so was ASEV because it was not performing an imputation, but being used for traceability. Again, an explanation for this was provided in the reviewer’s guide.

CONCLUSION

Clearly the path followed for this pooling could have been designed in different ways. The existing CDISC guidance allows for a breadth of interpretation. The good news is the draft integration guideline seems to be making progress in facilitating these decisions as more and more pooling is performed. In the end, in fact, only the ADSL/ADBL described in this paper would have been different: replaced by one IADSL, while the other ADaMs would have practically been the same.

Of course, this new guidance may not - yet - provide the one-stop solution to all. Several of the challenges we faced and solved, such as traceability, don’t seem to be fully incorporated yet. But it is a step in the right direction. Especially if the CDISC conformance rules can be published at release for Pinnacle to align the checks.

Other interesting papers that describe alternative approaches to integrate ADaMs for similar designs, i.e. subjects in multiple treatment groups, have been published since the time this analysis was performed. At the risk of repeating ourselves, multiple paths such as those suggested in this and these other papers can be chosen, as long as the importance of a clear and detailed accompanying reviewer’s guide and define.xml remains at the forefront of everyone’s mind to facilitate review.

REFERENCES

ADaM Implementation Guide v1.1

<https://www.cdisc.org/standards/foundational/adam/adam-implementation-guide-v1.1>

ADaM Structure for Occurrence Data (OCCDS) v1.0

<https://www.cdisc.org/standards/foundational/adam/adam-structure-occurrence-data-occds-v1.0>

STUDY DATA TECHNICAL CONFORMANCE GUIDE (March 2019)

<https://www.fda.gov/media/122913/download>

ADAMINT-ADaMDataStructuresforIntegration-Compiled-180919-1017-730 - Version 1.0 Draft (Dated February 2018)
(not online)

ADaM Structures for Integration: A Preview - Wayne Zhong; Kimberly Minkalis; Deborah Bauer – PharmaSUG 2018

<https://www.pharmasug.org/proceedings/2018/DS/PharmaSUG-2018-DS03.pdf>

ADaM Integration for Summary of Clinical Safety: The ‘Unique Patient’ Paradox – Tracy Sherman; Aakar Shah; Brian Fairfield-Carter – PharmaSUG 2018

<https://www.lexjansen.com/pharmasug/2018/BB/PharmaSUG-2018-BB06.pdf>

Best Practices for ISS/ISE Dataset Development – Bharath Donthi; Lingjiao Qi – PharmaSUG 2019

<https://www.pharmasug.org/proceedings/2019/AD/PharmaSUG-2019-AD-299.pdf>

CONTACT INFORMATION

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

Magalie Gallet
Cytel Inc.
ICC, Building H, 2nd floor
Route de Pré-Bois, 20
C.P. 1839
1215 Geneva 15, Switzerland
Email: Magalie.Gallet@cytel.com
Web: www.Cytel.com

Laura Phelan
Cytel Inc.
63, Avenue des Champs Elysée
75008 Paris, France
Email: Laura.Phelan@cytel.com
Web: www.Cytel.com

Brand and product names are trademarks of their respective companies.