PharmaSUG 2018 - Paper BB-06

ADaM Integration for Summary of Clinical Safety: The 'Unique Patient' Paradox

Tracy Sherman, Ephicacy Consulting Group, Inc.;
Aakar Shah, Pfizer Inc.;
Brian Fairfield-Carter, Syneos Health™ Canada

ABSTRACT

There are various ways to structure ADaM data for a Summary of Clinical Safety (SCS) analysis that integrates multiple studies, and the methods chosen depend on a number of factors. One of these is the issue of whether the ADaM data sets can or should be created at the study-level and just stacked together, or if there is a need to create new integration records within each patient. How should you handle patients at the study-level when you also need to integrate patients that were enrolled in multiple studies? That's right: accommodating both types of analysis ('patient experience' and the 'unique patient') is a requirement when structuring analysis data sets to support the TLFs (table, listing and figures) in a SCS.

This paper will propose one method of structuring ADaM safety data sets to incorporate the 'patient experience' in addition to the 'unique patient' analyses. We will outline the paradox of defining a unique patient and include the challenges faced and the decisions that were made to create the integrated analysis data sets and to further support the 'analysis-ready' CDISC principle to streamline TLF programming.

INTRODUCTION

You might well ask, what is a Summary of Clinical Safety (SCS)? A SCS is a global regulatory document required by the FDA as part of a new drug application. In essence, compared to an ISS (Integrated Summary of Safety) it's a smaller, more concise set of safety analyses required for summarizing the safety of a treatment. It usually involves integrating studies that were performed on the same molecule in different patient populations in various study phases.

For our most recent SCS, we were provided with a plan indicating which studies were to be integrated, and a set of TLF mock-ups (note that there was no Statistical Analysis Plan (SAP) provided, which will be discussed later). There were five studies in total: 1 phase III, 2 phase II's, and 2 phase I's. Study level ADaM data would be used for all studies, except for one early phase I study where we would use SDTM data. To keep it interesting, four of the studies were active at the time of programming, one study was blinded and one study was an extension study that allowed patients to roll into after the first study was completed.

The first version of the SCS table mock-ups were structured to provide a 'unique patient' analysis. Unique patient experience is the typical analysis seen where each patient in each study is considered unique and the patients enrolled in the open-label extension study are analyzed in a separate table column. We created ADaM specifications and the data sets based on the unique analysis type, with each patient in each study having a unique USUBJID.

After review of the first dry run of the TLFs, it was decided that the patients in the open-label extension study would be analyzed separately AND would be integrated into the 'All Studies' column to provide the 'patient experience' analysis (Here we are defining patient experience as the safety response during both studies (originating and the open-label extension study)). To do this, we needed to 'combine' the data

from both studies to get the overall picture of how the patient responded to the entire duration of active treatment.

The uniqueness of a patient is a matter of interpretation. Is the patient considered unique because they were enrolled in only one study? Or is a patient unique regardless of how many studies they enrolled in? Or does uniqueness dictate that a given data point for a given patient at a particular point in time should be included in a specific column of a summary table rather than in multiple columns? All patients are unique, but some are more unique than others. As you can imagine interpreting this uniqueness adds another layer of complexity to the common challenges of data integration.

Some of the more common data set integration challenges that we experienced (that are not discussed here) were as follows:

- Different version of MedDRA coding for adverse events between studies
- Baseline, post-baseline and treatment-emergent definitions were not uniform
- Active blinded studies with dummy exposure data
- Data issues (dirty data due to ongoing patients)
- Special interest adverse events were different between studies
- Different study reference dates (randomization versus enrollment date)
- Various study teams and resourcing
- Individual study ADaM were designed to support study CSR (Case Study Report), not SCS
- Implementing the data cutoff date for all studies
- Absence of a SAP

It's also important to mention that ADaM integration, which is essential for submission, is a topic that is not covered in the latest ADaM Implementation Guide V1.1 that was released February 12, 2016.

We will discuss integration challenges as well as how we achieved one of CDISC's ADaM goals; to simplify TLF programming with 'analysis-ready' data sets. Additionally, we will demonstrate one method for structuring specific ADaM safety data sets (ADSL, ADAE, ADLB, ADVS and ADEX) to provide the 'unique patient' and the 'patient experience' analyses.

OVERVIEW OF INTEGRATED STUDIES

Figure 1 summarizes a work flow diagram of the studies that were to be integrated and the source data for each study. We used study level ADaM for all studies, except Study 3, where we used SDTM.

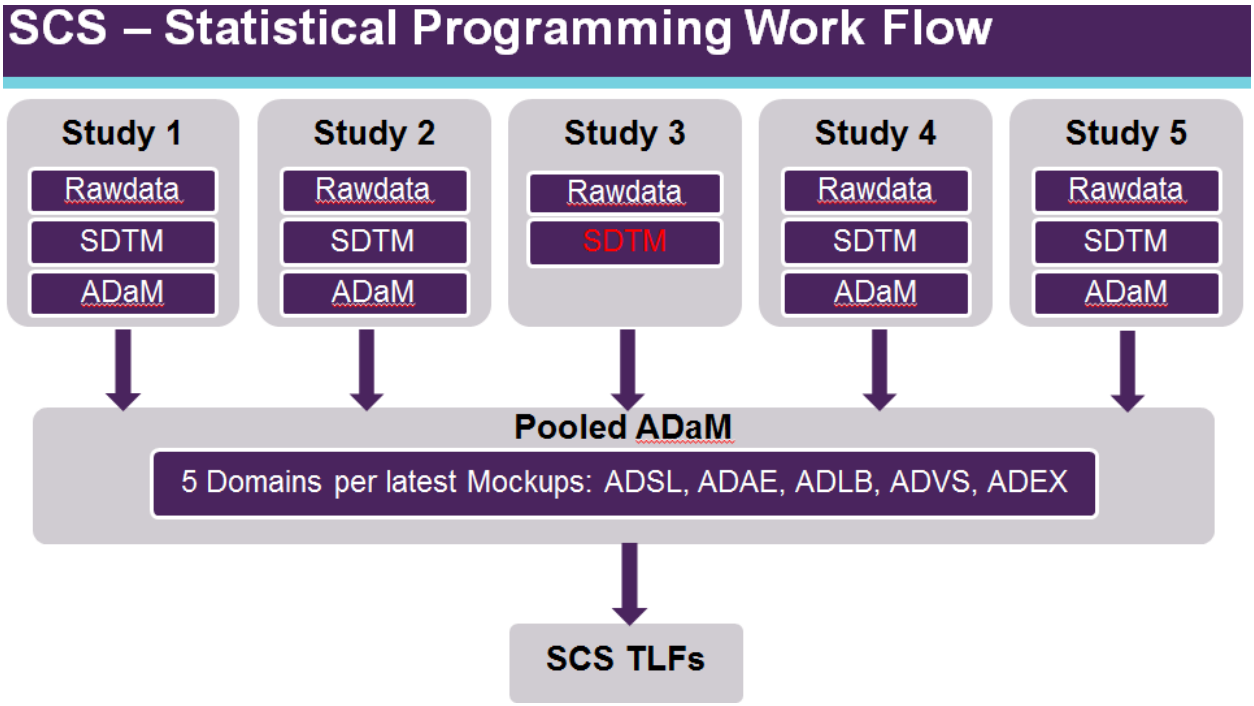

Figure 1. Work Flow Diagram Outlining Studies to be Integrated and the Five ADaM Domains Required for the Proposed Safety Summaries

The TLF mock-ups for the SCS required five pooled ADaM domains: ADSL, ADAE, ADEX, ADLB and ADVS. Based on the mock-ups, the ADaM data sets needed to support both 'unique patient' and 'patient experience' analysis. Each analysis will be described below, along with how ADaM data was structured in order to achieve programming simplicity.

CONTRASTING ANALYSIS TYPES: UNIQUE PATIENT VS PATIENT EXPERIENCE

UNIQUE PATIENT ANALYSIS

The unique patient analysis is one that you are probably most familiar with, where each patient in each study is considered unique. In the subject level analysis data set, ADSL, each patient has a unique identifier, USUBJID, for each study they participated in. In Figure 2, you can see that each study has a separate column in the demographic and baseline characteristics table. Patients who enrolled in the open-label extension study (Study 5) after they completed the originating study (Study 3 or 4) were summarized in separate table columns.

Patient Demographic and Baseline Characteristics
(Integrated Safety Population)

Baseline Characteristics	--------Randomized Study------ ----------Study 1---------- 5mg/day (N=xxx)	xxx (N=xxx)	----------Open-Label Uncontrolled Studies---- Study 2 5mg/day (N=83)	Study 3 5mg/day (N=77)	Study 4 5mg/day (N=xxx)	Open-Label Ext. [1] Study 5 5mg/day (N=XXX)	All Studies Total [2] 5mg/day (N=XXX)
Age							
n	xx	xx	xx	xx	xx	xx	xx
Mean (SD)	xx.x (xx.xx)	xx.x (xx.xx)	xx.x (xx.xx)	xx.x (xx.xx)	xx.x (xx.xx)	xx.x (xx.xx)	xx.x (xx.xx)
Median	xx.x	xx.x	xx.x	xx.x	xx.x	xx.x	xx.x
Min, Max	xx.x, xx.x	xx.x, xx.x	xx.x, xx.x	xx.x, xx.x	xx.x, xx.x	xx.x, xx.x	xx.x, xx.x

[1] Includes all patients who completed studies Study 3 and Study 4, subsequently enrolled in the open-label extension study (Study 5) and initiated treatment with 5mg/day in either the originating or extension study.
[2] Does not include the xxx arm of Study 1. Patients who started Studies 3 or 4 at 5 mg/day and continued in the extension study (Study 5) are counted only once in total number of patients for 'All Studies'.

Figure 2. Table Mock-up Showing 'Unique Patient' and 'Patient Experience' Analysis

PATIENT EXPERIENCE

When a patient enrolls in more than one study, we might ask, "How did the patient respond to treatment during both studies collectively?" The last column ('All Studies') in the table mock-up above (Figure 2) reflects the patient experience analysis.

Under the direction of the study biostatistician, it was decided that baseline values for a patient enrolled in multiple studies would be taken from the originating study rather than from the extension study. For tables that summarized post-baseline results, both the originating and extension study results were evaluated. This detail becomes important when you are deriving exposure parameters such as treatment duration and number of dose reductions/interruptions as well as lab shift parameters. For these derivations, we needed to add a variable, UNIQUEID, to ADSL to identify the same patient in each of the originating and extension studies. From there, we could use UNIQUEID instead of USUBJID to derive the patient experience parameters.

INCORPORATING BOTH PATIENT EXPERIENCE AND UNIQUE PATIENT ANALYSES IN ADAM

ADSL Variables

According to CDER (Center for Drug Evaluation and Research) outlined in the Common Data Standards Issues Document, the DM domain for integrated summaries may contain more than one record per unique patient in the case that an individual patient was enrolled in more than one study. It's important to note however that the document does not give guidance for structuring ADSL in this case.

After reviewing the CDISC ADaM Implementation Guide V1.1 and the TFL mock-ups, we decided to structure ADSL such that each patient in each study would have a separate record. The main reason to structure ADSL in this way was because the TFL mock-ups summarized each study individually, with a separate column for each study, and also included an "All Studies" column. As shown in Table 1, two additional variables were added to ADSL, UNIQUEID and INDEXFL.

Variable Name	Variable Label	Type	Codelist / Controlled Terms	Core	Notes
Study Identifiers					
STUDYID	Study Identifier	Char		Req	Must be identical to the SDTM variables DM.STUDYID, DM.USUBJID, DM.SUBJID and DM.SITEID
USUBJID	Unique Subject Identifier	Char		Req	
SUBJID	Subject Identifier for the Study	Char		Req	
SITEID	Study Site Identifier	Char		Req	
UNIQUEID	Unique Duplicate Subject Identifier	Char			Subject-level identifier that represents the SUBJID from the originating study.
INDEXFL	Index Subject Flag	Char	Y		A character indicator variable to flag which study was the originating or index study for each patient.

Table 1. ADSL Variables

The analysis specified that baseline for patients enrolled in multiple studies would be taken from the originating study. The variable INDEXFL was set to 'Y' on the study that the patient originated in. For demography and baseline tables, this variable was then used to select the appropriate baseline weight, age and lab values for each patient in each study as well as select the originating baseline value to be used in the 'All Studies' column (Figure 2).

UNIQUEID was set to the SUBJID from the originating study (INDEXFL='Y'). As shown below in Figure 3, if patients were enrolled in more than one study, then UNIQUEID has the same value for each study. For UNIQUEID='000201' highlighted in the screenshot of ADSL below, we use baseline values of weight (WEIGHTBL), Body Mass Index (BMIBL), and Aspartate Aminotransferase (ASTBL) from Study 4, the originating study.

STUDYID	USUBJID	SUBJID	UNIQUEID	INDEXFL	AGE	WEIGHTBL	BMIBL	ASTBL
Study 4	Study 4-000201	000201	000201	Y	57	76.00	32.6794	41
Study 5	Study 5-US000201	US000201	000201		57	73.02	31.3980	52
Study 4	Study 4-000203	000203	000203	Y	65	71.80	25.5609	25
Study 4	Study 4-000205	000205	000205	Y	47	55.20	20.2509	32
Study 4	Study 4-000206	000206	000206	Y	73	47.70	15.9931	19
Study 4	Study 4-000208	000208	000208	Y	52	72.30	24.4389	37
Study 1	Study 1-0127-0002	0127-0002	0127-0002	Y	58	46.50	18.1641	13
Study 1	Study 1-0127-0003	0127-0003	0127-0003	Y	46	81.60	29.9362	30
Study 1	Study 1-0127-0004	0127-0004	0127-0004	Y	53	67.50	27.3845	35
Study 1	Study 1-0127-0005	0127-0005	0127-0005	Y	54	89.80	31.9689	21
Study 1	Study 1-0127-0006	0127-0006	0127-0006	Y	49	66.40	24.2422	32
Study 2	Study 2-0127-2002	0127-2002	0127-2002	Y	52	52.60	19.9195	61
Study 2	Study 2-0151-2001	0151-2001	0151-2001	Y	67	58.10	23.0396	47
Study 2	Study 2-0157-2001	0157-2001	0157-2001	Y	43	64.50	27.3743	31
Study 2	Study 2-0157-2002	0157-2002	0157-2002	Y	34	57.70	20.2024	16
Study 2	Study 2-0157-2003	0157-2003	0157-2003	Y	54	57.90	19.8010	26
Study 3	Study 3-0807	0807	0807	Y	60	57.20	22.7687	22
Study 3	Study 3-0901	0901	0901	Y	45	58.80	22.9687	46
Study 3	Study 3-0902	0902	0902	Y	29	59.40	19.1761	41
Study 3	Study 3-0903	0903	0903	Y	67	55.60	18.5773	21
Study 5	Study 5-0216	0216	0903		72	56.50	18.8780	16
Study 3	Study 3-0904	0904	0904	Y	19	53.10	19.9857	21
Study 5	Study 5-0127	0127	1011		48	79.70	30.1822	23

Figure 3. ADSL Variables UNIQUEID and INDEXFL for a Patient Enrolled in Multiple Studies

Basic Data Structure (BDS) Data Sets

This is where it gets interesting. Our table programs had already been written to provide only the unique patient approach, which counted each patient per study only once in any table column. With the additional complexity of adding the patient experience, where the patient is counted once if they were enrolled in two studies (originating and the open-label extension study), we needed to adjust ADaM. In the lab shift table mock-up below, Figure 4, we had separate rows for the 'All Studies' patient experience analyses. If a patient was enrolled in more than one study, the patient was only counted once. For example, for the lab shift table, if the worst toxicity grade (ATOXGR) in the originating study was a 3 and in the open-label study it was a 4, then we needed to count the worst grade of 4 for that patient.

Summary of Laboratory Assessment by Worst Postbaseline Toxicity – Hematology (Hemoglobin)
(Integrated Safety Population)

Treatment Arm: Baseline Toxicity Grade	Worst Postbaseline Toxicity Grade [3]						
	Grade 0	Grade 1	Grade 2	Grade3	Grade 4	Missing	Total
Parameter: Hemoglobin (g/L) [Low]							
Study 1: 5mg/day							
Grade 0	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)
Grade 1	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)
Grade 2	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)
Grade 3	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)
Grade 4	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)
Missing	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)
Total	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)
Etc.							
All Studies							
Grade 0	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)
Grade 1	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)
Grade 2	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)
Grade 3	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)
Grade 4	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)
Missing	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)
Total	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)

Figure 4. Lab Shift Table Mock-up for Each Study and the 'All Studies' (Patient Experience Analysis)

After much discussion, we decided to structure our BDS data sets (ADLB and ADVS) by adding columns instead of parameters (Table 2).

Variable Name	Variable Label	Type	Codelist / Controlled Terms	Core	Notes
Analysis Parameter Variables					
PARAM	Parameter	Char		Req	The description of the analysis parameter.
PARAMCD	Parameter Code	Char		Req	The short name of the analysis parameter in PARAM.
AVAL	Analysis Value	Num		Cond	Numeric analysis value described by PARAM.
BASE	Baseline Value	Num		Cond	The patient's baseline analysis value for a parameter and baseline definition (i.e. BASETYPE) if present. BASE contains the AVAL copied from a record within the parameter on which ABLFL="Y".
CHG	Change from Baseline	Num		Perm	Change from baseline value. Equal to AVAL-BASE.
PCHG	Percent Change from Baseline	Num		Perm	Percent change from baseline analysis value. Equal to ((AVAL-BASE)/BASE)*100.
BASEALL	Baseline Value All Studies	Num		Perm	The patient's baseline analysis value for a parameter. BASEALL contains the AVAL copied from a record within the parameter on which ABLTFL="Y".

Variable Name	Variable Label	Type	Codelist / Controlled Terms	Core	Notes
Analysis Parameter Variables					
CHGALL	Change from Baseline All Studies	Num		Perm	Change from baseline value. Equal to AVAL-BASEALL.
PCHGALL	Percent Change from Baseline All Studies	Num		Perm	Percent change from baseline analysis value. Equal to ((AVAL-BASEALL)/BASEALL)*100.

Table 2. ADLB Analysis Parameter Variables

We added variables BASEALL, CHGALL and PCHGALL to incorporate the new definition of baseline for those patients enrolled in multiple studies who were then summarized in only one column in each table. The new variables added a few lines of code to each table program, to replace the BASE with BASEALL, CHG with CHGALL, etc. for the 'All Studies' column. Unfortunately, this is not in alignment with Rule 1 of the ADaM Implementation Guideline V1.1. It states that 'A parameter-invariant (does not vary at the parameter level) function of AVAL and BASE on the same row ***that does not involve a transform of BASE*** should be added as a new column.' In our case, although it was parameter-invariant, BASE was 'transformed' such that baseline was taken from the originating study if the patient enrolled more than once.

We also discussed Rule 6 of the ADaM Implementation Guideline V1.1: "***When there is more than one definition of baseline***, each additional baseline definition requires the creation of its own set of rows. BASETYPE is used to differentiate the definition of baseline." The IG also describes that addition of BASETYPE with different periods in a study. We did not have different periods and had the same definition of baseline for each of the five studies; however, for one column in the tables we are producing, we needed a separate definition. The questions we had were as follows:

1. Does the definition of baseline need to be similar at the study level?
2. Does 'more than one definition of baseline' relate to a separate column in a given output table such as the 'All Studies' column?

We had 132,541 records in ADLB. If we added BASETYPE='Originating study' records, it would have doubled the number of records and increased our SAS processing time.

For record-level clarity, flag variables were added to ADLB as well (Table 3). These variables were the following:

Variable Name	Variable Label	Type	Codelist / Controlled Terms	Core	Notes
Flag Variables					
ABLFL	Baseline Record Flag	Char	Y	Cond	Character indicator to identify the baseline record for each patient, parameter, and baseline type (BASETYPE) combination. ABLFL is required if BASE is present.

Variable Name	Variable Label	Type	Codelist / Controlled Terms	Core	Notes
Flag Variables					
ABLAFL	Baseline Record Flag for All Studies	Char	Y	Cond	Character indicator to identify the baseline record for each patient, parameter, and originating study (if applicable). ABLAFL is required if BASEALL is present.
ANL01FL	Analysis Record Flag 01	Char	Y	Cond	Conditionally required flag used to identify which records will be selected for AVISIT, if multiple records per PARAMCD and AVISIT.
ANTOTFL	Analysis Flag All Studies	Char	Y	Cond	Per parameter and UNIQUEID, set to 'Y' for the maximum grade value of SHIFT1L or SHIFT1H. If multiple records have the same worst postbaseline grade then assign flag to earliest record.
SHIFT1L	Worst Post-Baseline Toxicity Grade Low	Char	Y		Set the earliest worst postbaseline result of ATOXGRL within the same USUBJID and PARAMCD to 'Y' for all non-missing values in the treatment emergent period.
SHIFT1H	Worst Post-Baseline Toxicity Grade High	Char	Y		Set the earliest worst postbaseline result of ATOXGRH within the same USUBJID and PARAMCD to 'Y' for all non-missing values in the treatment emergent period.
SHIFTT1L	Worst Post-Baseline Tx. Gr. Low for All	Char	Y		Set the earliest worst postbaseline result of ATOXGRL within the same UNIQUEID and PARAMCD to 'Y' for all non-missing values in the treatment emergent period.
SHIFTT1H	Worst Post-Baseline Tx. Gr. High for All	Char	Y		Set the earliest worst postbaseline result of ATOXGRH within the same UNIQUEID and PARAMCD to 'Y' for all non-missing values in the treatment emergent period.

Table 3. ADLB Flag Variables

ADVS, the second BDS data set that was required for the SCS, was structured similarly to ADLB. We added the BASEALL, CHGALL and PCHGALL variables to incorporate the new definition of baseline for those patients enrolled in multiple studies. As we only had one vital sign table to produce, 'Potentially Clinically Significant Changes in Vital Signs and Weight', there were no additional shift or analysis variables required in ADVS.

Occurrence Data Structure Data Set

One Occurrence ADaM data set, ADAE, was required for TLF programming. Common AE tables such as Treatment-Emergent Adverse Events by System Organ Preferred Term were programmed as shown in Figure 5. Similar to the Demographic and Baseline Characteristics table, the events were compared for each study and for "All Studies' where each patient was counted only once if they were enrolled in multiple studies.

Merging in ADSL.UNIQUEID into ADAE by USUBJID and replacing the use of USUBJID with UNIQUEID in your code, will enable you to count events for a given preferred term only once by study and for 'All Studies' (patient experience analysis).

Treatment-Emergent Adverse Events by System Organ Class and Preferred Term
(Integrated Safety Population)

	---------Randomized Study----------		Open-Label Uncontrolled Studies--------			Open-Label Ext. [1]	All Studies
	---------Study 1----------		Study 2	Study 3	Study 4	Study 5	Total [2]
MedDRA System Organ Class Preferred Term	1mg/day (N=xxx)	xxx (N=xxx)	1mg/day (N=83)	1mg/day (N=77)	1mg/day (N=xxx)	1mg/day (N=XXX)	1mg/day (N=XXX)
No. of Patients with at Least 1 TEAE	xxx (xx.x%)	xxx (xx.x%)	xxx (xx.x%)	xxx (xx.x%)	xxx (xx.x%)	xxx (xx.x%)	xxx (xx.x%)
SOC1	xxx (xx.x%)	xxx (xx.x%)	xxx (xx.x%)	xxx (xx.x%)	xxx (xx.x%)	xxx (xx.x%)	xxx (xx.x%)
PT1	xxx (xx.x%)	xxx (xx.x%)	xxx (xx.x%)	xxx (xx.x%)	xxx (xx.x%)	xxx (xx.x%)	xxx (xx.x%)
PT2	xxx (xx.x%)	xxx (xx.x%)	xxx (xx.x%)	xxx (xx.x%)	xxx (xx.x%)	xxx (xx.x%)	xxx (xx.x%)
Etc.	xxx (xx.x%)	xxx (xx.x%)	xxx (xx.x%)	xxx (xx.x%)	xxx (xx.x%)	xxx (xx.x%)	xxx (xx.x%)
SOC2	xxx (xx.x%)	xxx (xx.x%)	xxx (xx.x%)	xxx (xx.x%)	xxx (xx.x%)	xxx (xx.x%)	xxx (xx.x%)
PT1	xxx (xx.x%)	xxx (xx.x%)	xxx (xx.x%)	xxx (xx.x%)	xxx (xx.x%)	xxx (xx.x%)	xxx (xx.x%)
PT2	xxx (xx.x%)	xxx (xx.x%)	xxx (xx.x%)	xxx (xx.x%)	xxx (xx.x%)	xxx (xx.x%)	xxx (xx.x%)
Etc.	xxx (xx.x%)	xxx (xx.x%)	xxx (xx.x%)	xxx (xx.x%)	xxx (xx.x%)	xxx (xx.x%)	xxx (xx.x%)

[1] Includes all patients who completed studies Study 3 and Study 4, subsequently enrolled in the open-label extension study (Study 5) and initiated treatment with 1mg/day in either the originating or extension study.
[2] Does not include the xxx arm of Study 1. Patients who started Studies 3 or 4 at 1 mg/day and continued in the extension study (Study 5) are counted only once in total number of patients for 'All Studies'.
Patients with multiple events for a given preferred term are counted once only for each preferred term.
Events are sorted by decreasing frequency of preferred term in the All Studies column.

Figure 5. Adverse Event Table Mock-up for each Study and the 'All Studies' (Patient Experience Analysis)

Other ADaM Data Sets

The Exposure Analysis Data Set, ADEX, was also needed to support table programming. Due to the structure of ADEX, new parameters (PARAM/PARAMCDs) for each existing parameter in ADEX were added to generate results for the 'All Studies' column in the exposure tables (see Table 4).

To calculate 'All Studies Treatment Duration' for patient who enrolled in multiple studies, we needed data points from both studies to summarize this for the patient experience. The date of first dose from the originating study and the last dose date from the extension study were needed and therefore it made sense to add additional parameters.

Similarly, an additional parameter was added for 'All Studies Cumulative number of doses received'. For patients that have the same ADSL.UNIQUEID (enrolled in multiple studies), the number of doses from each study where PARAMCD='CUMNDOS' were summed. For the rest of patients, the value from the record having PARAMCD='CUMNDOS', taken from study-level ADaM, was used.

PARAMCD	PARAM	Description
TRTDURM	Treatment Duration (months)	Date of last dose – date of first dose + 1)/30.4375. ((ADSL.TRTEDT-ADSL.TRTSDT+1)/30.4375.)
TRTDURMB	All Studies Treatment Duration (months)	For patients that have the same ADSL.UNIQUEID (enrolled in multiple studies), sum the treatment duration from each study where PARAMCD='TRTDURM'. Use core identifier variables (USUBJID, SITE, RACE, etc.) from the originating study (INDEXFL='Y') For rest of patients, use AVAL from PARAMCD='TRTDURM'.
CUMNDOS	Cumulative Number of Doses Received	Sum of the number of doses received per patient.
CUMNDOSB	All Studies Cumulative Number of Doses Received	For patients that have the same ADSL.UNIQUEID (enrolled in multiple studies), sum the number of doses from each study where PARAMCD='CUMNDOS'. Use core identifier variables (USUBJID, SITE, RACE, etc.) from the originating study (INDEXFL='Y'). For rest of patients, use AVAL from PARAMCD='CUMNDOS'.

Table 4. ADEX Parameters and Descriptions

CONCLUSION

When integrating ADaM data sets across multiple studies there are many factors to be considered, and these factors increase further when you have patients that have enrolled in multiple studies. Clearly, data will have to be combined in order to facilitate integrated analysis, either on the same record or by dataset concatenation. If each study is analyzed as if the patient were a 'unique' patient, then just stacking the individual ADaM data sets is sufficient; however, if only one value per patient will be analyzed based on the time of the first dose of active treatment (which may be the first or last study a patient was enrolled in), then careful consideration must be given as to which variables should be created to capture the values needed for analyses.

We strongly recommend a statistical analysis plan be provided for an integrated Summary of Clinical Safety if it is not already supplied with the TLF specifications. Programmers need to know whether patients will be analysed uniquely at the study level or the patient level or both. In the SAP and the TLF specifications, it should be clearly noted how patient uniqueness is defined in each column in the summary tables. These distinctions can add clarity to how ADaM datasets should be structured.

In our case, the SCS analysis summarized patients uniquely at the study level and also included a column in the tables for the patient experience. We showed that adding the variables UNIQUEID and INDEXFL to ADSL (which can be merged into any ADaM data set by USUBJID) can assist in both the patient experience and the unique patient analyses. INDEXFL was set to 'Y' on the records to be used for baseline values. UNIQUEID was set to the USUBJID from the originating study if the patient was enrolled in multiple studies or set to USUBJID if only one study was involved.

Safety BDS datasets were structured with additional variables added for the patient experience analysis where Other ADaM such as ADEX had additional parameters derived.

We are looking forward to new developments in the standards on how ADaM should be structured for these kinds of analyses. Submitting data in standard form is a key part in the advancement of review efficiency and quality.

REFERENCES

"Analysis Data Model (ADaM)". http://www.cdisc.org/adam. The current version is downloadable from the web page and available to CDISC members and non-members.

"Analysis Data Model (ADaM) Implementation Guide". http://www.cdisc.org/adam. The current version is downloadable from the web page and available to CDISC members and non-members.

"CDER Common Data Standards Issues Document". https://www.fda.gov/downloads/Drugs/DevelopmentApprovalProcess/FormsSubmissionRequirements/ElectronicSubmissions/UCM254113.pdf

ACKNOWLEDGMENTS

We would like to thank Ganesh Gopal from Ephicacy Consulting Group and Shuxuan Zhao from Pfizer Inc. for their continued support and encouragement in conference attendance, as well as all our family, friends and colleagues.

CONTACT INFORMATION

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

Name: Tracy Sherman
Enterprise: Ephicacy Consulting Group, Inc.
E-mail: shermantracy@gmail.com

Name: Aakar Shah
Enterprise: Pfizer, Inc.
E-mail: Aakar.Shah@pfizer.com