PharmaSUG 2025 - Paper SS-258

The Show Must Go On: Best Practices for Submitting SDTM Data for Ongoing Studies

Kristin Kelly, Pinnacle 21 by Certara, Radnor, PA, USA

ABSTRACT

Though the CDISC SDTM Implementation Guide provides advice on how to prepare SDTM datasets for completed studies, there is little guidance on what to do when the study is ongoing, leading to varied implementation practices across the industry. At times, it may be difficult for a regulatory reviewer to readily determine that a study is still in progress without looking in the Clinical Study Data Reviewer's Guide (cSDRG). The recent addition of the ONGOSIND (Ongoing Study Indicator) parameter in the FDA Study Data Technical Conformance Guide (sdTCG) allows sponsors to clearly specify within the data whether a study is ongoing. In this paper, some considerations for preparing domains such as Demographics (DM), Disposition (DS), and Trial Summary (TS) for an ongoing study as well as strategies to ensure data transparency across the SDTM submission package will be discussed.

INTRODUCTION

The CDISC SDTM Implementation Guide (SDTMIG) guidance assumes that a clinical trial has been completed. However, an increasing number of submissions now include data from studies still in progress. Due to limited guidance on handling ongoing study data, sponsors face challenges in implementing SDTM standards consistently. The varied approaches taken with SDTM data and other submission deliverables across sponsors result in a lack of standardization and reduced transparency for regulatory reviewers in determining study status. The recent inclusion of the ONGOSIND (Ongoing Study Indicator) Trial Summary parameter in the FDA Study Data Technical Conformance Guide (sdTCG) offers a standardized method for sponsors to indicate the ongoing status of a study within the dataset.

This paper explores key considerations for structuring SDTM domains, including Demographics (DM), Disposition (DS), and Trial Summary (TS) for ongoing studies, while ensuring transparency across the SDTM submission package. Also, ways in which the ONGOSIND flag in TS can be leveraged in validation of study data will be discussed.

BEST PRACTICES FOR SDTM DOMAINS

One of the ways to ensure transparency in the data that a study is ongoing would be to represent it in its current state rather than presenting it as a completed study. Subjects that are still on-study should be readily discernable when looking at the SDTM data. The following sections will discuss how this can be done in the SDTM datasets themselves.

DEMOGRAPHICS (DM)

The Demographics (DM) domain is a Special-Purpose dataset in SDTM that contains subject-level information related to sex, race, age, treatment arm assignment, and date/times of treatment. Other information included are informed consent, last date of participation, and death information, if applicable. The reference date/times such as RFSTDTC, RFENDTC, RFXSTDTC, RFXENDTC, etc, are pulled in from other domains in order to make DM a 'one-stop shop' when starting a review.

When a sponsor chooses to populate all the reference date/time variables in DM as if all subjects have reached the milestones needed to complete the study, it is difficult for a reviewer to determine if the study is still ongoing. The example DM dataset below shows data for four subjects that have been randomized, treated and seemingly completed the study. Both ARMCD/ARM and ACTARMCD/ACTARM variable pairs are populated as well as the date/times for the study reference period (RFSTDTC and RFENDTC), date/times for first and last treatment (RFXSTDTC and RFXENDTC), as well as the date/time of end of participation (RFPENDTC). Please note that in this example, USUBJID = 'CWCW-645-0006' (highlighted in orange), is a screen failure so all the variables mentioned are null with ARMNRS set to 'SCREEN FAILURE'.

STUDYID	USUBJID	RFSTDTC	RFENDTC	RFXSTDTC	RFXENDTC
CWCW-645	CWCW-645-0005	2023-07-15T08:32	2024-02-24T09:34	2023-07-15T08:32	2024-02-24T09:34
CWCW-645	CWCW-645-0006				
CWCW-645	CWCW-645-1007	2023-06-10T10:30	2024-08-10T10:30	2023-06-10T10:30	2024-08-10T10:30
CWCW-645	CWCW-645-1008	2023-08-01T08:23	2024-09-15T11:34	2023-08-01T08:23	2024-09-15T11:34
CWCW-645	CWCW-645-1111	2023-07-12T15:00	2024-05-12T13:00	2023-07-12T15:00	2024-05-12T13:00

RFICDTC	RFPENDTC	>>>	ARMCD	ARM	ACTARMCD	ACTARM	ARMNRS
2023-06-30	2024-04-15		CWCW15	Drug CWCW645 15 mg	CWCW15	Drug CWCW645 15 mg	
2023-07-01	2023-08-15						SCREEN FAILURE
2023-05-17	2024-09-10		PLACEBO	Placebo	PLACEBO	Placebo	
2023-06-28	2024-09-24		CWCW10	Drug CWCW645 10 mg	CWCW10	Drug CWCW645 10 mg	
2023-06-29	2024-06-30		CWCW05	Drug CWCW645 5 mg	CWCW05	Drug CWCW645 5 mg	

In this example ongoing study, USUBJID = ‘CWCW-645-1007’ and ‘CWCW-645-1111’ (highlighted rows in pink and purple; respectively) have not yet completed treatment and are still participating in the study. But it seems that they have left the study because all the reference dates/times are populated in DM. It is not determined until further investigation of the submission package that these subjects are still on-study.

In the define.xml file for this study, the following Methods for assigning the reference dates/times were provided:

DM Variable	Method
RFSTDTC	Date/Time of First Dose
RFENDTC	Date/Time of Last Dose
RFXSTDTC	Date/Time of First Dose
RFXENDTC	Date/Time of Last Dose
RFPENDTC	Latest Date available for subject/Last Contact Date

A better representation in DM for these subjects would be to only populate variables for which the subject met that milestone. Since both subjects have been treated at least once and still in the treatment phase of the study, their records in DM would look like this:

STUDYID	USUBJID	RFSTDTC	RFENDTC	RFXSTDTC	RFXENDTC
CWCW-645	CWCW-645-0005	2023-07-15T08:32	2024-02-24T09:34	2023-07-15T08:32	2024-02-24T09:34
CWCW-645	CWCW-645-0006				
CWCW-645	CWCW-645-1007	2023-06-10T10:30		2023-06-10T10:30	
CWCW-645	CWCW-645-1008	2023-08-01T08:23	2024-09-15T11:34	2023-08-01T08:23	2024-09-15T11:34
CWCW-645	CWCW-645-1111	2023-07-12T15:00		2023-07-12T15:00	

RFICDTC	RFPENDTC	>>>	ARMCD	ARM	ACTARMCD	ACTARM	ARMNRS
2023-06-30	2024-04-15		CWCW15	Drug CWCW645 15 mg	CWCW15	Drug CWCW645 15 mg	
2023-07-01	2023-08-15						SCREEN FAILURE
2023-05-17			PLACEBO	Placebo	PLACEBO	Placebo	
2023-06-28	2024-09-24		CWCW10	Drug CWCW645 10 mg	CWCW10	Drug CWCW645 10 mg	
2023-06-29			CWCW05	Drug CWCW645 5 mg	CWCW05	Drug CWCW645 5 mg	

For USUBJID = ‘CWCW-645-1007’ and ‘CWCW-645-1111’, RFENDTC, RFXENDTC, and RFPENDTC are null to indicate that they have not received their final dose and have not completed/discontinued from the study. This representation of the data helps a reviewer readily determine which subjects are still on-study after only having examined the DM dataset.

DISPOSITION (DS)

The Disposition (DS) domain is an Events domain that provides an accounting for all subjects who entered the study and may include protocol milestones, such as randomization, as well as the subject's completion status or reason for discontinuation for the entire study or each phase or segment of the study, including screening and post-treatment follow-up. Sponsors may choose which disposition events and milestones to collect for a study.

For subjects that are on-study, it is recommended to only create records in DS that have been collected thus far. If the subject has not completed the study and does not have a record where DSDECOD = ‘COMPLETED’ (or a term from the SDTM NCOMPLT codelist) and DSCAT = ‘DISPOSITION EVENT’, then there should not be a record created for the subject. This is similar to the recommendation given for the DM domain.

STUDYID	USUBJID	DSTERM	DSDECOD	DSCAT	EPOCH	DSSTDTC
CWCW-645	CWCW-645-0005	INFORMED CONSENT OBTAINED	INFORMED CONSENT OBTAINED	PROTOCOL MILESTONE	SCREENING	2023-06-30
CWCW-645	CWCW-645-0005	COMPLETED	COMPLETED	DISPOSITION EVENT	TREATMENT	2024-02-24
CWCW-645	CWCW-645-0005	COMPLETED	COMPLETED	DISPOSITION EVENT	FOLLOW-UP	2024-04-15
CWCW-645	CWCW-645-0006	INFORMED CONSENT OBTAINED	INFORMED CONSENT OBTAINED	PROTOCOL MILESTONE	SCREENING	2023-07-01
CWCW-645	CWCW-645-0006	SCREEN FAILURE	SCREEN FAILURE	DISPOSITION EVENT	SCREENING	2023-08-15
CWCW-645	CWCW-645-1007	INFORMED CONSENT OBTAINED	INFORMED CONSENT OBTAINED	PROTOCOL MILESTONE	SCREENING	2023-05-17
CWCW-645	CWCW-645-1111	INFORMED CONSENT OBTAINED	INFORMED CONSENT OBTAINED	PROTOCOL MILESTONE	SCREENING	2023-06-29

The first subject, USUBJID = 'CWCW-645-0005', completed all parts of the study with records where DSDECOD = 'COMPLETED' and DSCAT = 'DISPOSITION EVENT' for EPOCH = 'TREATMENT' and 'FOLLOW-UP' (blue rows). The second subject, USUBJID = 'CWCW-645-0006', is a screen failure and only has records where DSDECOD = 'INFORMED CONSENT OBTAINED' and 'SCREEN FAILURE' (orange rows). The next two subjects, USUBJIDs = 'CWCW-645-0007' and 'CWCW-645-1111', are still ongoing in the study and thus only have one record each in DS for DSDECOD = 'INFORMED CONSENT OBTAINED' (pink and purple rows).

TRIAL SUMMARY (TS)

The Trial Summary (TS) domain is a Trial Design domain. This dataset allows the sponsor to submit a summary of the trial in a structured format. Each record in the TS dataset contains the value of a parameter, a characteristic of the trial. For example, TS is used to record basic information about the study such as trial phase, protocol title, and trial objectives. The TS dataset contains information about the planned and actual trial characteristics. It is also a required dataset for submissions to the FDA.

The parameters that should be included in TS are listed in Appendix B in the FDA Study Data Technical Conformance Guide (sdTCG). Each parameter that should be included will have a value of 'Y' in the 'FDA Desired' column. For ongoing studies, there are several parameters that can be used to denote that the study has not completed, and they are all 'FDA Desired' = 'Y' for clinical trials.

FDA Desired - Clinical	TSPARMCD	TSPARM
Y	DCUTDESC	Data Cutoff Description
Y	DCUTDTC	Data Cutoff Date
Y	ONGOSIND	Ongoing Study Indicator
Y	SENDTC	Study End Date

The most recent addition to this list is TSPARMCD/TSPARM = 'ONGOSIND/Ongoing Study Indicator' that first appeared in the March 2024 version of the sdTCG. Since it is an indicator, TSVAL would be populated with either 'N' (No) or 'Y' (Yes). Prior to this, there was no indication in the data that the study was ongoing and because the CDISC standards assume a completed study, there was no mechanism in the data that could be used to perform cross-checks with TS in validation. Now that this parameter should be included in TS, this cross-checking can be done.

In keeping with the current ongoing study example, the TS snippet below shows how these four parameters can be used to indicate that the study is ongoing.

STUDYID	TSSEQ	TSGRPID	TSPARMCD	TSPARM	TSVAL	TSVALNF	TSVALCD
CWCW-645	1	IA	DCUTDESC	Data Cutoff Description	INTERIM ANALYSIS		
CWCW-645	1	IA	DCUTDTC	Data Cutoff Date	2024-09-30		
CWCW-645	1		ONGOSIND	Ongoing Study Indicator	Y		C49488
CWCW-645	1		SENDTC	Study End Date		NAV	

The DCUTDTC and DCUTDESC parameters indicate when the data was cutoff for submission and for what reason. In this example, the data submitted was cut on 2024-09-30 for an 'INTERIM ANALYSIS'. Then the TSGRPID variable is used to tie the two parameters together. The ONGOSIND parameter is set to 'Y' to indicate that the study is ongoing with TSVALCD populated with the C-code for this term from the CDISC NY codelist. Because the study is still in progress, TSVAL for SENDTC (Study End Date) is set to null and TSVALNF is set to 'NAV' for 'Not Available' from the ISO21090 standard. When these four parameters are configured properly, it should be easy to determine that the study is not yet complete when the TS dataset is reviewed.

UPDATES TO P21 VALIDATION RULES

Another component that needs to be taken into account is running the CDISC and FDA validation rules against the soon to be submitted data package. Historically, there are several rules that are triggered for ongoing study data that preparers of the SDTM data have explained in the cSDRG as 'false positives'. This is not an accurate explanation. A 'false positive' in validation is when a rule is not working as it should. The rules that typically surface in the validation for ongoing studies are working as intended. They are flagging exactly what the rule was designed to flag under the assumption that the data is from a completed study. That being said, it has always been desired to find some way to 'turn off' specific rules in this instance, but without some indication present in the data, it was almost impossible to do so.

With the addition of the TSPARMCD/TSPARM = 'ONGOSIND'/'Ongoing Study Indicator' in TS, a mechanism in the data that a study is ongoing can now be leveraged in validation. By cross-checking subject-level datasets with TSVAL = 'Y' for the ONGOSIND parameter in TS, certain rules will not be triggered in the validation, and they will not have to be explained in the cSDRG. The following sections below discuss some of these rule updates.

NEW RULE FOR 'ONGOSIND' IN TS

In addition to updated rules, there is one new rule that checks for the presence of the ONGOSIND parameter because it is now an 'FDA Desired' parameter listed in the sdTCG. If this rule fires in validation, it is because the parameter is missing and should be added to TS. It should NOT be explained in the cSDRG! It is simple enough to update TS per FDA requirements so….JUST DO IT!

Rule ID	Message	Description	Message Type	Review Impact
SD2287	Missing ONGOSIND Trial Summary Parameter	Ongoing Study Indicator' (ONGOSIND) record must be populated in Trial Summary (TS) domain. It is an 'FDA Desired' parameter listed in the Technical Conformance Guide (TCG).	Error	High

RULES THAT CHECK THE DISPOSITION (DS) DOMAIN

As mentioned earlier, subjects that are still on-study will typically not yet have a record in DS where DSCAT = 'DISPOSITION EVENT'. Before the ONGOSIND flag, this instance would be flagged by the rules shown below.

SD0069 checks for the presence of any record in the DS domain for SDTMIG v3.2 (and earlier). SD2288 checks that there is a record where DSCAT = 'DISPOSITION EVENT' for subjects that are not screen failures, not assigned or where ARMCD is not null. This rule applies for all versions of the SDTMIG. These two rules will no longer be triggered in validation as long as DS and TS are run together and ONGOSIND = 'Y'.

Rule ID	Message	Description	Message Type	Review Impact
SD0069	No Disposition record found for subject	All Demographics (DM) subjects (USUBJID) should have at least one record in the Disposition (DS) domain. The only possible exceptions for SDTMIG v3.2 (and earlier) are Screen Failures (ARMCD=SCRNFAIL) and Not Assigned Treatment (ARMCD=NOTASSGN) subjects, or subjects with ARMCD=null. If the study is ongoing, Parameter Value (TSVAL) for the 'ONGOSIND/Ongoing Study Indicator' parameter should be set to 'Y' in the Trial Summary (TS) domain.	Warning	Medium
SD2288	No DISPOSITION EVENT record found for subject	All subjects (USUBJID) should have at least one record in the Disposition (DS) domain with Disposition Category (DSCAT) equal to DISPOSITION EVENT. Subjects with ARMCD='SCRNFAIL', ARMCD = null, or ARMCD='NOTASSGN' are excluded from this check. If the study is ongoing, Parameter Value (TSVAL) for the 'ONGOSIND/Ongoing Study Indicator' parameter should be set to 'Y' in the Trial Summary (TS) domain.	Warning	High

Another set of rules that typically surface in validation due to the absence of the final DS record for a subject are shown below. SD0080 compares AE.AESTDTC to the latest DS.DSSTDTC. SD0082 compares the EX.EXENDTC for a subject to the latest DS.DSSTDTC. SD1446 checks EX.EXSTDTC against the latest DS.DSSTDTC for a subject. Again, if TS is run in the validation and has TSVAL = 'Y' for the ONGOSIND parameter, these rules will no longer fire and will not have to be explained in the cSDRG.

Rule ID	Message	Description	Message Type	Review Impact
SD0080	AE start date is after the latest Disposition date	Start Date/Time of Adverse Event (AESTDTC) should be less than or equal to the Start Date/Time of the latest Disposition Event (DSSTDTC). Investigator should provide a date of study discontinuation for each subject in Disposition (DS) domain. If the study is ongoing, Parameter Value (TSVAL) for the 'ONGOSIND/Ongoing Study Indicator' parameter should be set to 'Y' in the Trial Summary (TS) domain.	Warning	High
SD0082	Exposure end date is after the latest Disposition date	End Date/Time of Treatment (EXENDTC) should be less than or equal to the Start Date/Time of the latest Disposition Event (DSSTDTC). Investigator should provide a date of study discontinuation for each subject in Disposition (DS) domain. If the study is ongoing, Parameter Value (TSVAL) for the 'ONGOSIND/Ongoing Study Indicator' parameter should be set to 'Y' in the Trial Summary (TS) domain.	Warning	High
SD1446	Exposure start date is after the latest Disposition date	Start Date/Time of Treatment (EXSTDTC) should be less than or equal to the Start Date/Time of the latest Disposition Event (DSSTDTC). Investigator should provide a date of study discontinuation for each subject in Disposition (DS) domain. If the study is ongoing, Parameter Value (TSVAL) for the 'ONGOSIND/Ongoing Study Indicator' parameter should be set to 'Y' in the Trial Summary (TS) domain.	Warning	High

RULES THAT CHECK AGAINST RFPENDTC

SD1202, SD1203 and SD1204 check –STDTC, --DTC, and –ENDTC in subject-level domains against the date/time in RFPENDTC in DM. These rules will fire when RFPENDTC is before the date in the subject-level domain. Please note that if the sponsor chooses to leave RFPENDTC null for subjects that are ongoing, these rules will not appear in the validation report. For an ongoing study and RFPENDTC is populated, these rules will no longer be triggered if TS is included in the validation and ONGOSIND = 'Y'.

Rule ID	Message	Description	Message Type	Review Impact
SD1202	--STDTC date is after RFPENDTC	Start Date/Time (--STDTC) variable value must be less than or equal to Date/Time of End of Participation (RFPENDTC). If the study is ongoing, Parameter Value (TSVAL) for the 'ONGOSIND/Ongoing Study Indicator' parameter should be set to 'Y' in the Trial Summary (TS) domain.	Error	Medium
SD1203	--DTC date is after RFPENDTC	Date/Time of Collection (--DTC) variable value must be less than or equal to Date/Time of End of Participation (RFPENDTC). If the study is ongoing, Parameter Value (TSVAL) for the 'ONGOSIND/Ongoing Study Indicator' parameter should be set to 'Y' in the Trial Summary (TS) domain.	Error	Medium
SD1204	--ENDTC date is after RFPENDTC	End Date/Time (--ENDTC) variable value must be less than or equal to Date/Time of End of Participation (RFPENDTC). If the study is ongoing, Parameter Value (TSVAL) for the 'ONGOSIND/Ongoing Study Indicator' parameter should be set to 'Y' in the Trial Summary (TS) domain.	Error	Medium

RULES THAT CHECK AGAINST RFENDTC

SD0088 will be triggered if RFENDTC is not provided for a subject when ARMCD is not null and not 'SCRNFAIL' or 'NOTASSGN'. SD1376 fires when RFENDTC is null and ACTARMCD is not 'SCRNFAIL', 'NOTASSGN', 'NOTTRT' or null. This can occur for an ongoing study when the sponsor chooses to leave RFENDTC null for subjects that are still in the study. These rules will no longer fire in this case if TS is included in the validation and ONGOSIND = 'Y'.

Rule ID	Message	Description	Message Type	Review Impact
SD0088	RFENDTC is not provided for a randomized subject	Subject Reference End Date/Time (RFENDTC) should be populated for all randomized subjects, those where Planned Arm Code (ARMCD) is not equal to 'SCRNFAIL' or 'NOTASSGN' or null. If the study is ongoing, Parameter Value (TSVAL) for the 'ONGOSIND/Ongoing Study Indicator' parameter should be set to 'Y' in the Trial Summary (TS) domain.	Warning	High
SD1376	RFENDTC is not provided for a treated subject	Subject Reference End Date/Time (RFENDTC) should be populated for all treated subjects, those where Actual Arm Code (ACTARMCD) is not equal to 'SCRNFAIL', 'NOTASSGN' or 'NOTTRT' (or null). If the study is ongoing, Parameter Value (TSVAL) for the 'ONGOSIND/Ongoing Study Indicator' parameter should be set to 'Y' in the Trial Summary (TS) domain.	Warning	Medium

BEST PRACTICES FOR VALIDATION

The previous sections outlined how validation should be run for ongoing studies so that certain rules that assume a completed study can be 'turned off'. These methods should NOT be used for a completed, locked study in order to avoid having to provide explanations in the cSDRG! When these rules are triggered for subjects in a completed study, it is important to investigate why the rules might be firing. If there is an outage in the data, then it is important to explain these issues to ensure transparency to the reviewer.

Further, it may make more sense to use these methods in the last validations prior to submission if perhaps, the preparer of the SDTM data uses the validation report for QC purposes to find programming issues. If this is the case that the user would like to have the rules fire, there are a few ways to keep the rules 'turned on'.

1. TS is not included in the validation
2. TS is included but the ONGOSIND parameter is not present
3. TS is included and ONGOSIND = 'N'
4. TS is included and ONGOSIND = NULL

Please note that even when ONGOSIND = 'Y' and TS is included in the validation, if there are instances in the data that fail the criteria in the rule logic, the rules will still fire and will need to be explained in the cSDRG.

BEST PRACTICES FOR THE CSDRG

Lastly, please do note in the cSDRG that the study is ongoing. There is a question in the cSDRG template that asks whether the study is ongoing or not (Section 3.1). Be sure that this section is completed properly and don't forget to update this section when the study is complete.

3. Subject Data Description

3.1 Overview

Are the submitted data taken from an ongoing study? <Yes> <No>

If yes, describe the data cut or database status:

(Text here)

CONCLUSION

Though there is little guidance from CDISC on what to do for ongoing studies, there are ways that the SDTM data and submission deliverables can be compiled to ensure transparency downstream to a reviewer. With the introduction of the ONGOSIND parameter in TS, there is finally a mechanism present in the data that indicates that the study is ongoing. It can be leveraged in validation for rules that are based on the assumptions for a completed study. Also, the recommendations for structuring certain datasets in this case, are simply that. It is always a sponsor decision on how to implement the standards. But hopefully, these best practices will be adopted by industry in order to ensure speed of review in an effort to always get drugs and vaccines to patients faster.

REFERENCES

Study Data Tabulation Model Implementation Guide: Human Clinical Trials. Clinical Data Interchange Standards Consortium (CDISC) Submission Data Standards (SDS) Team. Version 3.4. November 2021

FDA. Study Data Technical Conformance Guide. Version 5.9. November 2024. https://www.fda.gov/media/153632/download

CONTACT INFORMATION

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

Kristin Kelly
Pinnacle 21 by Certara
kristin.kelly@certara.com