PharmaSUG 2025 - Paper PO-235

Guidance from the FDA authored "Submitting Patient-Reported Outcome Data in Cancer Clinical Trials" and recommendations for their implementation in study CDISC COA data

Charity Quick, Emergent BioSolutions

ABSTRACT

In November of 2023, the U.S. Food & Drug Administration (FDA) released a 48-page document titled "Submitting Patient-Reported Outcome Data in Cancer Clinical Trials". This document provides technical specifications for submitting patient-reported outcome (PRO) data collected in cancer clinical trials to support oncology studies. Given the breadth and specificity of the guidance, the increasing prevalence of Clinical Outcome Assessments (COAs) in the work we do, and the direct link of the content to the FDA publications such as the "Study Data Technical Conformance Guide", adoption of the standards contained in "Submitting Patient-Reported Outcome Data in Cancer Clinical Trials" should be considered regardless of the COA therapeutic area. This paper and the accompanying poster seek to provide a summary of these FDA recommendations as they relate to study data collection, tabulation, analysis, and submission.

INTRODUCTION

The FDA defines a COA as “an assessment of a clinical outcome that describes or reflects how an individual feels, functions or survives. The assessment can be made through report by a clinician, a patient, a non-clinical observer, or through a performance-based assessment” (FDA CDER SBIA, 2017). The 4 types of COAs outlined in this definition are:

1. Patient-Reported Outcome (PRO)
2. Clinician-Reported Outcome (ClinRO)
3. Observer-Reported Outcome (ObsRO)
4. Performance Outcome (PerfO)

For the most part, this topics and discussion in this paper will focus on PRO COA data. In clinical trials, PRO data is used to provide data pertaining to the subject experience, help regulatory authorities make decisions, and ensure patient needs are being met. Where PRO data is included in a study, it’s common to see it included as a primary or secondary endpoint. In many cases, the PRO data are captured outside of the EDC system which can add complexity to statistical programming and additional manual work to integrate it into other study data streams and submission documents such as the acrf.pdf.

Acknowledging the importance of these data, the FDA, PMDA, NMPA, and EMA have all published recent guidance related to patient experience data. Specifically, over the past decade, the FDA has published numerous guidance documents including final Patient-Focused Drug Delivery (PFDD) guidance in 2020 (FDA, 2020) that includes the “Submitting Patient-Reported Outcome Data in Cancer Clinical Trials” (PROCCT) in 2023 (FDA, 2023a).

As the title suggests, the PROCCT non-binding guidance is recommended for Cancer studies, but the data examples and discussion pertaining to CDISC implementation are largely therapeutic area independent. Additionally, while many of the CDISC data mapping solutions included in the PROCCT are not currently presently accounted for in the CDISC Implementation Guides (IGs) the mappings are consistent with the evolution that has been occurring over the past decade in the FDA Study Data Technical Conformance Guide (sdTCG), Therapeutic Area Guides (TAUGs), and CDISC IGs for PRO based domains – particularly when it comes to missing and skipped item tabulation, analysis, and documentation. For example, the PROCCT encourages the use of “phantom records” to show missing question response rows in ADaM data – a concept that has been included in the ADaM IGs since version

1.0 but is neither encouraged not discouraged by CDISC. In 2020, “PHANTOM” was added to the ADaM DTYPE codelist and in 2023 the FDA published PROCCT relies heavily on PHANTOM records to facilitate analysis surrounding missing PRO data. Given this change and others like it over time, it is the author’s opinion that the use of PHANTOM records in PRO data has evolved from optional in most cases to essential in many cases.

The author believes that the guidance exemplified in the PROCCT is what the FDA will eventually prefer to see implemented for PRO data tabulation and analysis in agency submissions regardless of therapeutic area. Given that the PROCCT guidance is non-binding and many of its SDTM and ADaM recommendations are not currently required by CDISC, this paper aims to provide the reader with up to date and practical recommendations for PRO vendor selection, data capture, and CDISC data mapping or at least help facilitate discussions about the evolution of PRO data expectations and requirements.

OVERVIEW OF THE “SUBMITTING PATIENT-REPORTED OUTCOME DATA IN CANCER CLINICAL TRIALS” GUIDANCE

The PROCCT is part of the FDAs Patient Focused Drug Development (the PFDD) initiative series. The goal of the PFDD initiative is to address how stakeholders can effectively collect and submit patient experience data for regulatory review. The initiative demonstrates how central patient experience data is to the FDAs commitment to encouraging better understanding of the patient experience and patient-centered drug development. To supplement the PFDD Guidance Series, FDA issued two technical specifications guidance documents, “Submitting Clinical Trial Datasets and Documentation for Clinical Outcome Assessments Using Item Response Theory” (FDA, 2023b) and “Submitting Patient-Reported Outcome Data in Cancer Clinical Trials”.

At 48 pages, the PROCCT is a comprehensive and highly detailed document that standardizes the collection and reporting of patient-reported outcome data. It provides:

- Numerous SDTM data examples with practical implementations
- SDTM variable specifications for Non-Standard Variables (NSVs)
- SDTM trial design recommendations specifically for the TS domain
- ADaM examples built from the sample SDTM source data, including specifications and guidance
- Instructions for handling missing data through "phantom records" imputation (varying based on whether the PRO measures efficacy or safety/tolerability)
- New ADaM variables to enhance data analysis capabilities

While analysis tables and figures aren't the focus of this paper, it's worth noting that the PROCCT includes 17 table and figure examples that demonstrate the practical application of its SDTM and ADaM recommendations. Ultimately, the PROCCT serves as a comprehensive blueprint for end-to-end processing of PRO data in clinical trials.

The author has 2 additional notes: First, the descriptions of CDISC tabulation and analysis as well as examples in the PROCCT are given in terms of the SDTM QS domain and variables and the ADaM ADQS domain and variables. Although SDTM QS and ADaM ADQS are the most common recipients of PRO data, other domains may be used where appropriate. For simplicity and to align with the PROCCT, this paper will refer to PRO SDTM data in QS and PRO ADaM data in ADQS. Second, it is advised that your Study Data Standardization Plan (SDSP) include the PROCCT if you plan on using the guidance. In the sdTCG the FDA states “Technical specification documents provide detailed information for content on specific topics, where applicable, submitted to FDA for an application. Sponsors should consult with the review division early in the process to discuss issues with trial design or conduct that may affect the content of the study data being submitted”.

SDTM PROCCT GUIDANCE AND CURRENT RECOMMENDATIONS

The PROCCT contains recommendations for SDTM that range from simple tabulation updates to broader implementation proposals that may be challenging to incorporate depending on how the PRO data were

collected. In this section, we will outline the PROCCT recommendations for SDTM and then consider the potential value they may bring to a clinical study as well as PRO recommendations from other industry publications.

SDTM PROCCT GUIDANCE

SDTM PROCCT guidance can be summarized in 4 topics.

1. **Adherence to CDISC Standards**: Follow current CDISC guidance (SDTM Implementation Guide, applicable Questionnaire Supplements, and CDISC Controlled Terminology) as closely as possible, particularly regarding naming conventions for QSTEST and QSTESTCD variables.
2. **Supplemental Variables**: Include additional Non-Standard Variables (NSVs) in SUPPQS when relevant data has been collected, such as collection mode and language information.
3. **Missing Data Handling**: Implement specific approaches for handling missing data and logically skipped questions within the SDTM QS domain, ensuring comprehensive documentation of data completeness.
4. **Trial Summary Documentation**: Add specific TSPARM/TSPARMCD records to the TS domain when implementing PROCCT guidance to inform regulatory reviewers of the methodology used.

Adherence to CDISC Standards

Before moving into specific recommendations, the PROCCT states:

> *“Sponsors should implement the CDISC SDTM standard when submitting clinical tabulation data and the CDISC ADaM standard when submitting analysis data”. Correct usage of CDISC CT is also mentioned prior to moving into specific recommendations. Further, the PROCCT provides variable specifications for key Questionnaire variables such as QSCAT, QSTEST, QSTESTCD, etc. and notes “PRO data should be tabulated using the SDTM QS domain specifications. For select measures, CDISC publishes QRS Supplements that provide guidance on representing named COA measures in the SDTM. CDISC also provides submission values within the controlled terminology related to named COA measures. These resources should be consulted for guidance on a PRO measure, in addition to the guidance provided in the SDTM and SDTMIG.”*

Additional NSVs

If other information pertaining to the PRO were collected and could be useful for regulatory review, these should be included in the SUPPQS data. Suggestions from the PROCCT are “Data Collection Mode”, “Data Collector” and “Language”. These NSVs should not be tabulated into SUPPQS if they weren’t collected and their inclusion in ADaM is only recommended if they are needed for analysis.

Handling Missing Data and Logically Skipped Questions in the SDTM QS Domain

Missing Data in SDTM QS

The most significant SDTM recommendation in the PROCCT concerns how to handle missing responses in the QS domain. The guidance emphasizes:

> *"Understanding the reasons for and prevalence of missing PRO data are critical to support FDA review and regulatory decision-making."*

To achieve this understanding, the PROCCT recommends:

1. All missing PRO data should be explicitly represented within the QS dataset
2. The reason for missingness should be documented using the 'Reason Not Performed' (QSREASND) variable

3. The QS dataset should include one record per item per PRO measure per patient per assessment timepoint—even when responses are missing
4. When summary scores are included in source data, the QS dataset should also include one row per summary score per PRO measure per patient per assessment timepoint—regardless of whether that summary score is missing

This approach ensures regulatory reviewers have complete visibility into data completeness patterns and reasons for missing information.

The following table is provided in the PROCCT to illustrate 4 missing data scenarios and how they should be handled in the SDTM QS domain using QSSTAT and QSREASND.

Scenario	Recommended Representation in QS Dataset
The patient did not respond to an item administered within a PRO measure.	The row for the missing item response should include: • QSSTAT = 'NOT DONE' • QSREASND contains the reason the patient did not respond if known/collected. Otherwise, QSREASND is empty/null.
A source data summary score cannot be calculated per the scoring algorithm based on the available item responses (e.g., due to insufficient item response data).	The row for the missing source data summary score should include: • QSSTAT = 'NOT DONE' • QSREASND is populated if known/collected (e.g., QSREASND = 'NOT CALCULABLE'). Otherwise, QSREASND is empty/null.
The patient was not administered the PRO measure either at an onsite visit attended by the patient or at a planned (per protocol) offsite PRO assessment timepoint.	The row for each missing item response and source data summary score within the measure should include: • QSSTAT = 'NOT DONE' • QSREASND contains the reason the measure was not administered if known/collected. Examples include, but are not limited to, patient was physically unable to complete the PRO measure due to adverse event, patient refusal, patient did not provide, study site failed to administer or other site staff error, or technological problems with a PRO administered electronically.
The patient did not attend an onsite visit and the PRO measure is only administered onsite.	The row for each missing item response and source data summary score within the measure should include: • QSSTAT = 'NOT DONE' • QSREASND contains the reason the patient did not attend the visit if known/collected. Examples include, but are not limited to, patient was unable to attend a scheduled trial visit due to hospitalization.

Display 1. PROCCT QS missing response table

Logically Skipped Questions in SDTM QS

In addition to the guidance for missing responses above, the PROCCT states that items missing due to logical skips in the instrument should be handled consistent with the FDA sdTCG. These instructions have been present in the sdTCG since 2017. The instructions are outlined below but the reader should refer to the sdTCG for additional information and context if needed.

> instructions on how to record and/or score responses to logically skipped items are available from the instrument developer, then records for logically skipped items should be included in the submission dataset with the following:
>
> • QSSTAT = 'NOT DONE';
> • QSREASND = 'LOGICALLY SKIPPED ITEM'; and
> • QSORRES, QSSTRESC, and QSSTRESN would be assigned according to the instrument's instructions.
>
> If instructions on how to record and/or score responses to logically skipped items are not available from the instrument developer, then records for logically skipped items should be included in the submission dataset with the following:
>
> • QSSTAT = 'NOT DONE';
> • QSREASND = 'LOGICALLY SKIPPED ITEM'; and
> • QSORRES, QSSTRESC, and QSSTRESN all set to null.

NOTE - If the PRO instrument utilizes Computerized Adaptive Testing (CAT), rows for unanswered questions should not be included in the QS dataset.

Trial Summary (TS) Data

If guidance from the PROCCT is used to tabulate SDTM for your study, the recommendation is to add a row to the TS domain to notify regulatory reviewers. Specifications for variable values in the additional TS row are as follows:

- TSPARMCD = 'FDATCHSP'
- TSPARM = 'FDA Technical Specification'
- TSVAL = 'Oncology PROs Technical Specifications Guidance v1.0'

Note that the PROCCT contains a few typos relating to the TSPARM/TSPARMCD that the user should be aware of. First, the variables are incorrectly named TSPARAM and TSPARAMCD. Second, the correct TSPARM in the latest CDISC CT is "FDA Technical Specification" but the PROCCT states to use the value "FDA Tech Spec". To avoid conformance issues in your TS domain, use the TSPARM and TSPARMCD that matches the CDISC CT version for the codelists in your study.

SDTM PROCCT EXAMPLE

The PROCCT provides example data pertaining to SDTM QS tabulation of missing rows. The example data represents responses for a 2-item questionnaire for 2 subjects that should be completed during screening and 3 treatment cycles. Explanation of rows added and QSSTATQSREASND values adhere to the guidance in table XX above and are as follows:

1. Row 3 shows QSSTAT = "NOT DONE" and a missing value for QSREASND since the reason not done was not captured in the PRO instrument collection for the first patient's Cycle 1, Day 1 visit.
2. Rows 7 and 8 show QSSTAT = "NOT DONE" but the QSREASND = "PATIENT REFUSAL" in this instance since the PRO collected data captured the reason not done. The patient did attend the visit so rows for both questionnaire items are created and the QSDTC is populated since the date of the visit and intended questionnaire completion is known.
3. Rows 11 and 12 show QSSTAT = "NOT DONE" and QSREASND = "HOSPITALIZATION". In this instance the patient didn't attend the visit at all so rows for both questionnaire items are created and the QSDTC is not populated since the subject did not attend the visit.
4. Patient 2 only has records for Screening and Cycle 1, Day 1. The patient died after the Cycle 1, Day 1 visit. Since the PRO data in this example are used to evaluate Clinical Benefit, the missing records are derived in ADaM ADQS as phantom records. See ADaM PROCCT recommendations and examples below for additional details.

Row	USUBJID	VISIT	QSCAT	QSTEST	QSTESTCD	QSORRES	QSSTAT	QSREASND	QSDTC
1	A_100_1	SCREENING	Measure Name and Version	I01-Item 1	I01	3			2022-02-01
2	A_100_1	SCREENING	Measure Name and Version	I01-Item 2	I02	5			2022-02-01
3	A_100_1	CYCLE 1 DAY 1	Measure Name and Version	I01-Item 1	I01		NOT DONE		2022-02-22
4	A_100_1	CYCLE 1 DAY 1	Measure Name and Version	I01-Item 2	I02	4			2022-02-22
5	A_100_1	CYCLE 2 DAY 1	Measure Name and Version	I01-Item 1	I01	2			2022-03-15
6	A_100_1	CYCLE 2 DAY 1	Measure Name and Version	I01-Item 2	I02	4			2022-03-15
7	A_100_1	CYCLE 3 DAY 1	Measure Name and Version	I01-Item 1	I01		NOT DONE	PATIENT REFUSAL	2022-04-05
8	A_100_1	CYCLE 3 DAY 1	Measure Name and Version	I01-Item 2	I02		NOT DONE	PATIENT REFUSAL	2022-04-05
9	A_100_2	SCREENING	Measure Name and Version	I01-Item 1	I01	4			2022-03-14
10	A_100_2	SCREENING	Measure Name and Version	I01-Item 2	I02	5			2022-03-14
11	A_100_2	CYCLE 1 DAY 1	Measure Name and Version	I01-Item 1	I01		NOT DONE	HOSPITALIZATION	
12	A_100_2	CYCLE 1 DAY 1	Measure Name and Version	I01-Item 2	I02		NOT DONE	HOSPITALIZATION	

Display 2. QS example data in PROCCT

RECOMMENDATIONS FOR SDTM PRO TABULATION BASED ON PROCCT GUIDANCE

As already noted above, the PROCCT states “Understanding the reasons for and prevalence of missing PRO data are critical to support FDA review and regulatory decision-making.” While this statement is contained in a non-binding document for cancer studies, the sentence indicates a broader intent that should be considered when tabulating collected PRO data into SDTM. Additionally, the PROCCT SDTM recommendations use CDISC SDTM standards as a foundation so the following the PROCCT isn’t inherently breaking CDISC SDTM rules. The following author recommendations are based on guidance from the PROCCT, the FDA sdTCG, and CDISC advice from the SDTM IG and Metadata Submissions Guidelines (MSG) and are relevant for all therapeutic areas.

1. Use the CDISC SDTM model and IG as well as relevant QRS and TAUG guidance as a foundation for PRO data tabulation.
2. Unless a CDISC TAUG or QRS being implemented for your study contradict the PROCCT, your organization should consult with their review division and consider tabulating the SDTM dataset that contains your PRO in a way that will include one record per item per PRO measure per patient per assessment timepoint, regardless of whether an item response is missing. Depending on how your PRO data were collected, this implementation could mean manual addition of rows and additional data sources to be merged to populate variables resulting in additional time and risk. Until further guidance is given by CDISC or regulatory authorities, it’s the author’s recommendation that following PROCCT SDTM guidance for missing items should be utilized if the trial is/will be pivotal to your submission program or if the PRO data will be used in a primary or secondary endpoint for your study. If SDTM recommendations for the PROCCT are followed correctly, there should be no data conformance issues to explain because of using PROCCT SDTM recommendations.
3. Since guidelines for logically skipped items has been included in the FDA sdTCG for several years, logically skipped items should be included in your PRO SDTM data if the instrument contains logical skips.
4. When possible, ensure that the collected PRO data will already contain a solid framework for missing and logically skipped items as source data so that additional manipulation to tabulate the data into SDTM can be minimized as much as possible. Work with your organization to ensure that ePRO vendor selection or the CRF build for PRO data will do as much of the heavy lifting as possible when it comes to ensuring that data collection aligns as closely as possible with the desired SDTM output.
5. If additional information is collected pertaining to the PRO data, ensure that those details are captured in your supplemental (SUPP) SDTM domain if they would be helpful for a reviewer.
6. If guidance from the PROCCT was implemented in SDTM data tabulation for the PRO, include a row in the SDTM TS domain to inform reviewers. Given that most of the SDTM advice in the PROCCT is consistent with current CDISC standards and established recommendations in the FDA sdTCG, the addition of this TS row should probably only be needed if missing response rows are added as described in item 2 above.
7. In addition to #6 above, ensure that the csdrg.pdf for the study cites the use of the PROCCT. The author suggests adding details in the QS dataset portion of the document and in section 1.3 “Study Data Standards and Dictionary Inventory” table.
8. Although not mentioned in the PROCCT, if your PRO data are not part of the acrf.pdf those pages should be added to the acrf.pdf consistent with instructions in the CDISC MSG 2.0. This is particularly important for trials that use PRO data in a primary or secondary endpoint. Not including PRO instrument pages in your acrf.pdf could result in a request from your reviewer to add them post-submission.
9. If total scores for the PRO instrument are already included in the collected data, these should be included in SDTM QS. Otherwise, do not derive total scores in the SDTM data.

ADAM PROCCT GUIDANCE AND CURRENT RECOMMENDATIONS

The PROCCT provides extensive ADaM recommendations with two primary goals:

1. **Streamlining Analysis**: Preparing data that enables "one proc away" table and figure programming—meaning statistical displays can be created with minimal programming steps from the analysis datasets
2. **Ensuring Data Completeness**: Creating datasets that present a comprehensive picture of PRO data completeness by explicitly representing missing items and scores

The ADaM examples in the PROCCT build directly upon the SDTM QS examples, providing a clear illustration of how to implement the missing data guidance across both standards.

An important distinction in the ADaM guidance is that the approach to handling missing data depends on the intended use of the PRO measure:

- Different handling for PROs used to assess clinical benefit (efficacy)
- Different handling for PROs used to assess safety/tolerability

These distinctions and their implementation details are explained in the following sections.

ADAM PROCCT GUIDANCE

ADaM PROCCT guidance can be summarized in 5 topics.

1. Follow current CDISC guidance (ADaM IG and CDISC CT) as closely as possible
2. Ensure all items needs for analysis and traceability from SDTM QS and ADSL are included in ADQS
3. Correct usage of ADaM variables PARCATx and DTYPE
4. Missing data represented by PHANTOM records are included in ADQS as needed based on intended use of PRO data (safety/tolerability vs clinical benefit)
5. Additional ADaM variables included to aid reviewers and allow for efficient “one proc away” analysis

Follow current ADaM CDISC guidance

Before moving into specific recommendations, the PROCCT states “Sponsors should implement the CDISC SDTM standard when submitting clinical tabulation data and the CDISC ADaM standard when submitting analysis data”. Further, the PROCCT provides variable specifications for key Questionnaire variables such as QSCAT, QSTEST, QSTESTCD, etc. and notes “PRO data should be tabulated using the SDTM QS domain specifications. For select measures, CDISC publishes QRS Supplements that provide guidance on representing named COA measures in the SDTM. CDISC also provides submission values within the controlled terminology related to named COA measures. These resources should be consulted for guidance on a PRO measure, in addition to the guidance provided in the SDTM and SDTMIG.”

Ensure ADaM Traceability and Analysis Readiness

The PROCCT emphasizes two critical aspects of ADaM ADQS datasets:

1. Traceability Back to Source Data The ADQS dataset must maintain clear traceability back to both SDTM QS and ADSL. To achieve this traceability, ADQS should retain key SDTM QS variables including:

- Question identifiers (QSTEST/QSTESTCD)
- Result values (QSORRES/QSSTRESC/QSSTRESN)
- Key metadata variables (QSSEQ, QSSTAT, QSREASND)
- Visit information (VISIT, VISITNUM)

Retaining these variables enables reviewers to trace each ADQS record back to its original SDTM source data or verify that it was derived appropriately.

2. Analysis-Ready Structure To support efficient analysis without redundant programming, ADQS should incorporate all relevant variables from ADSL that might be needed for subgroup analysis, including:

- Treatment variables
- Stratification factors
- Population flags (SAFFL, ITTFL)
- Key dates (RANDDT, DTHDT, EOSDT)
- Status indicators (EOTSTT, RANDFL)

This comprehensive approach to dataset construction applies universally across therapeutic areas, not just for oncology studies.

Proper use of PARCATx and DTYPE variables in ADQS

PARCATx

One of the most significant contributions in the PROCCT ADaM guidance is its detailed explanation of PARCATx variable implementation. These variables are especially important when individual item scores are used to calculate summary scores—a common pattern in PRO instruments.

The PARCATx guidance stands out for two reasons:

1. **Universal Applicability**: This section is notably not cancer-specific, making it immediately relevant to all therapeutic areas.
2. **Comprehensive Examples**: The document provides five detailed examples with matching scenarios showing exactly how to map PARCATx variables. This level of detail helps ensure accurate identification of each measure, score, and individual item—a critical factor for regulatory reviewers who need to understand the relationships between items and derived scores.

The example below from the PROCCT demonstrates the recommended approach:

Scenario 2 within Table 5 represents a scenario where multiple scores are calculated, and each scale score is calculated from distinct, mutually exclusive item score(s). PARCAT3 reports the scale to which each item contributes.

Table 5. ADQS Dataset Structure for Scenario 2

PARCAT1	PARCAT2	PARCAT3	PARAM
Measure Name and Version	ITEM	Scale Score 1	Item 1
Measure Name and Version	ITEM	Scale Score 1	Item 2
Measure Name and Version	ITEM	Scale Score 1	Item 3
Measure Name and Version	ITEM	Scale Score 2	Item 4
Measure Name and Version	ITEM	Scale Score 2	Item 5
Measure Name and Version	ITEM	Scale Score 3	Item 6
Measure Name and Version	SCALE SCORE	Scale Score 1	Scale Score 1
Measure Name and Version	SCALE SCORE	Scale Score 2	Scale Score 2
Measure Name and Version	SCALE SCORE	Scale Score 3	Scale Score 3

Display 3. PROCCT PARCATx example

DTYPE

The concept of "phantom records" represents a significant shift in how missing data should be handled in PRO datasets:

Background in ADaM Standards ADaM Implementation Guide version 1.3 (section 4.10.1) describes a method for explicitly representing missing items in ADaM datasets rather than implicitly indicating missingness through the absence of records. According to this guidance:

- These explicit records would have missing values for AVAL and AVALC
- They would be identified by DTYPE = "PHANTOM"
- The ADaM IG took a neutral stance, neither encouraging nor discouraging this practice

PROCCT's Strong Recommendation In contrast to ADaM's neutral position, the PROCCT explicitly advocates for the use of PHANTOM records, stating they are essential because they:

- Allow reviewers to clearly understand the prevalence of missing data
- Document the specific reasons for missingness
- Provide transparency about how missing data impact study endpoints
- Enable more thorough validation of the statistical approaches used

The PROCCT provides detailed specifications for implementing PHANTOM records in the sample ADaM datasets and refers users to the ADaM IG for additional technical details about DTYPE implementation.

This shift from "optional practice" to "strongly recommended approach" represents one of the most significant practical implications of the PROCCT guidance.

Implications for Missing Data PHANTOM records based on intended use of PRO data

A key insight from the PROCCT is that handling missing data and intercurrent events should differ based on the PRO's primary objective. This distinction is critical for implementing the guidance correctly.

What are Intercurrent Events? Intercurrent events are unplanned occurrences that can prevent subjects from completing assessments according to protocol. Examples include:

- Death
- Disease progression
- Treatment discontinuation
- Adverse events requiring hospitalization

Core Principle for All PRO Data The fundamental requirement is that ADQS should contain a record for every item and subscore at every protocol-specified timepoint for each patient. However, how extensively this principle is applied varies by PRO objective.

Specific Implementation Rules:

PRO Objective	Implementation Requirements
Clinical Benefit (Efficacy)	• Include records for all possible timepoints for every randomized subject, even if not treated • Continue including records after intercurrent events • Add PHANTOM records during treatment pauses
Safety/Tolerability	• Include records only for treated subjects• Include records only up to intercurrent events • Add PHANTOM records during treatment pauses • No records needed for randomized-but-not-treated subjects

This differentiated approach aligns with the statistical principles for handling missing data in efficacy versus safety analyses, providing the appropriate context for each type of endpoint.

Additional ADaM variables to use when records are added to ADQS

The PROCCT provides details about the use of additional variables to be used in ADaM ADQS to aid reviewer's understanding of records added for analysis of PRO data. Some of these variables are already part of the ADaM model while others are newly introduced in the PROCCT. Some of these variables are outlined below but refer to the PROCCT for additional information.

1. ONTRTFL and ONTRxxFL – a flag indicating that the observation occurred while the patient was on treatment (or on treatment during a particular period xx). The use of this variable is strongly recommended by the FDA in the PROCCT.
2. SCBLFL - used on the ABLFL = "Y" record to indicate that baseline is sourced from a Screening assessment timepoint(s) rather than from a prespecified baseline assessment timepoint.
3. PROEXPFL - indicator variable to specify whether the PRO parameter (item or summary score) corresponds to a per protocol planned PRO assessment timepoint. Use of PROEXPFL will be based on the intended PRO use (clinical benefit vs safety/tolerability). If PRO use is both, use PROEX1FL and PROEX2FL.
4. PROSCMFL - indicator variable to specify whether the PRO item score or summary score is populated at a planned (per protocol) PRO assessment timepoint. Empty if AVAL/AVALC are missing and response not provided by patient.
5. AREASND – use when the item or summary score is missing. Populated from SDTM QS data QSREASND if the record comes from SDTM. Otherwise, AREASND is populated by another source, when available. For example, if the PRO data are used to evaluate clinical benefit, AREASND may be populated for phantom records using SDTM DS.DSDECOD or ADSL DCTREASP for a patient who died or discontinued from treatment.

ADAM PROCCT EXAMPLE

The PROCCT provides example data for ADaM ADQS pertaining to the same SDTM QS tabulation example shown above. Again, the example data represents responses for a 2-item questionnaire for 2 subjects that should be completed during screening and 3 treatment cycles. Total score is calculated in ADQS for each patient and analysis visit (i.e., where PARAM = 'Total Score') since scores are not captured in SDTM QS. While not all variables expected in ADQS are represented, the example includes an explanation of rows added and variables used as follows:

1. Row 6 shows AREASND = "NOT CALCULABLE" since the scoring manual for this example PRO assessment states that the Total Score can't be calculated if only one item is completed. Note that QSSTAT and QSREASND since the record for PARAM = "Total Score" is created in ADQS and not present in SDTM QS.
2. Rows 4 through 6 provide examples of ONTRTFL, PROEXPFL, and PROSCMFL. ONTRTFL = "Y" for all 3 records since the patient was on treatment for this visit. PROEXPFL = "Y" since the patient was expected to complete the PRO at the visit. PROSCMFL = "Y" only in row 5 since it is the only record that had a response by the patient.
3. Row 12 represents the Total Score which has AREASND = "PATIENT REFUSAL" as reported in QS QSREASND for the individual items in rows 10 and 11.
4. Rows 10 through 12 provide examples of ONTRTFL, PROEXPFL, and PROSCMFL. ONTRTFL = "Y" for all 3 records since the patient was on treatment for this visit. PROEXPFL = "Y" since the patient was expected to complete the PRO at the visit. PROSCMFL is empty for all rows since the measure wasn't completed.
5. Row 18 shows the Total Score AREASND = "HOSPITALIZATION" as is shown in QS QSREASND for individual items in rows 16 and 17.
6. Rows 16 through 18 provide examples of ONTRTFL, PROEXPFL, and PROSCMFL. ONTRTFL is empty since the patient was not on treatment but PROEXPFL = "Y" since the subject was

expected to complete the PRO at the visit. PROSCMFL is empty for all rows since the measure was not completed.

7. In this ADaM example, the PRO is used to evaluate clinical benefit so phantom records are derived in the ADQS dataset for all items and summary scores for the missing records after the patient died. DTYPE = “PHANTOM” is used with AREASND = “DEATH” (populated from ADSL DCTREAS). QS variables QSSTAT and QSREASND are empty since these records were not present in SDTM. Since the patient died, ONTRTFL, PROEXPFL, and PROSCMFL are all empty.

Row	USUBJID	VISIT	AVISIT	PARCAT1	PARAM	PARAMCD	AVAL	QSSTAT
1	A_100_1	SCREENING	SCREENING	Measure Name and Version	I01-Item 1	I01	3	
2	A_100_1	SCREENING	SCREENING	Measure Name and Version	I01-Item 2	I02	5	
3	A_100_1	SCREENING	SCREENING	Measure Name and Version	Total Score	TS	8	
4	A_100_1	CYCLE 1 DAY 1	BASELINE	Measure Name and Version	I01-Item 1	I01		NOT DONE
5	A_100_1	CYCLE 1 DAY 1	BASELINE	Measure Name and Version	I01-Item 2	I02	4	
6	A_100_1	CYCLE 1 DAY 1	BASELINE	Measure Name and Version	Total Score	TS		
7	A_100_1	CYCLE 2 DAY 1	CYCLE 2 DAY 1	Measure Name and Version	I01-Item 1	I01	2	
8	A_100_1	CYCLE 2 DAY 1	CYCLE 2 DAY 1	Measure Name and Version	I01-Item 2	I02	4	
9	A_100_1	CYCLE 2 DAY 1	CYCLE 2 DAY 1	Measure Name and Version	Total Score	TS	6	
10	A_100_1	CYCLE 3 DAY 1	CYCLE 3 DAY 1	Measure Name and Version	I01-Item 1	I01		NOT DONE
11	A_100_1	CYCLE 3 DAY 1	CYCLE 3 DAY 1	Measure Name and Version	I01-Item 2	I02		NOT DONE
12	A_100_1	CYCLE 3 DAY 1	CYCLE 3 DAY 1	Measure Name and Version	Total Score	TS		
13	A_100_2	SCREENING	SCREENING	Measure Name and Version	I01-Item 1	I01	4	
14	A_100_2	SCREENING	SCREENING	Measure Name and Version	I01-Item 2	I02	5	
15	A_100_2	SCREENING	SCREENING	Measure Name and Version	Total Score	TS	9	
16	A_100_2	CYCLE 1 DAY 1	BASELINE	Measure Name and Version	I01-Item 1	I01		NOT DONE
17	A_100_2	CYCLE 1 DAY 1	BASELINE	Measure Name and Version	I01-Item 2	I02		NOT DONE
18	A_100_2	CYCLE 1 DAY 1	BASELINE	Measure Name and Version	Total Score	TS		
19	A_100_2	CYCLE 2 DAY 1	CYCLE 2 DAY 1	Measure Name and Version	I01-Item 1	I01		
20	A_100_2	CYCLE 2 DAY 1	CYCLE 2 DAY 1	Measure Name and Version	I01-Item 2	I02		
21	A_100_2	CYCLE 2 DAY 1	CYCLE 2 DAY 1	Measure Name and Version	Total Score	TS		
22	A_100_2	CYCLE 3 DAY 1	CYCLE 3 DAY 1	Measure Name and Version	I01-Item 1	I01		
23	A_100_2	CYCLE 3 DAY 1	CYCLE 3 DAY 1	Measure Name and Version	I01-Item 2	I02		
24	A_100_2	CYCLE 3 DAY 1	CYCLE 3 DAY 1	Measure Name and Version	Total Score	TS		

Row	QSREASND	DTYPE	AREASND	DCTREAS	PROEXPFL	PROSCMFL	ONTRTFL
1					Y	Y	
2					Y	Y	
3					Y	Y	
4					Y		Y
5					Y	Y	Y
6			NOT CALCULABLE		Y		Y
7					Y	Y	Y
8					Y	Y	Y
9					Y	Y	Y
10	PATIENT REFUSAL		PATIENT REFUSAL		Y		Y
11	PATIENT REFUSAL		PATIENT REFUSAL		Y		Y
12			PATIENT REFUSAL		Y		Y
13					Y	Y	
14					Y	Y	
15					Y	Y	
16	HOSPITALIZATION		HOSPITALIZATION		Y		
17	HOSPITALIZATION		HOSPITALIZATION		Y		
18			HOSPITALIZATION		Y		
19		PHANTOM	DEATH	DEATH			

20		PHANTOM	DEATH	DEATH			
21		PHANTOM	DEATH	DEATH			
22		PHANTOM	DEATH	DEATH			
23		PHANTOM	DEATH	DEATH			
24		PHANTOM	DEATH	DEATH			

Display 4. ADQS example data in PROCCT

RECOMMENDATIONS FOR ADAM PRO TABULATION BASED ON PROCCT GUIDANCE

The FDA has clearly stated that understanding both the quantity and reasons for missing PRO data is critical to their regulatory decision-making process. Given this emphasis, organizations should strongly consider consulting their review division and incorporating the PROCCT guidance into their PRO ADaM data strategy—even for non-oncology studies.

It's important to note that the PROCCT builds upon existing CDISC ADaM standards rather than conflicting with them. Following these recommendations doesn't violate CDISC compliance; rather, it implements CDISC standards in a way that specifically addresses FDA concerns about PRO data.

The following author recommendations are based on guidance from the PROCCT, the FDA sdTCG, and CDISC advice from the ADaM IG. While developed for oncology studies, these recommendations have practical value across all therapeutic areas that collect PRO data.

1. Use the CDISC ADaM model and IG as well as relevant QRS and TAUG guidance as a foundation for PRO data tabulation.
2. Ensure that the PRO ADaM ADQS data use SDTM QS and ADSL as a source and that all variables needed from both SDTM QS and ADSL to support both traceability and analysis are included in ADQS. ADQS should be able to directly facilitate “one proc away” display programming and provide reviewers with flag variables and other helpful indicators as suggested in the PROCCT. For example, the addition of the ONTRTFL to indicate that the record occurred while the subject was on study treatment could be of use to reviewers and is a simple variable to add to the ADQS data. If ADaM recommendations for the PROCCT are followed correctly, there should be no data conformance issues to explain because of using PROCCT ADaM recommendations.
3. As shown in the PROCCT examples, ensure that the PARCATx variable is used correctly to allow reviewers to identify each measure, score, item, total score, and sub score. Given the number of examples in the PROCCT to illustrate correct usage of PARCATx, it’s clear that it’s important to the FDA that information is captured accurately there.
4. Unless a CDISC TAUG or QRS being implemented for your study contradict the PROCCT, your organization should consider using DTYPE = “PHANTOM” records with the accompanying AREASND to create the PRO ADaM dataset in a way that will include one record per item per PRO measure per patient per assessment timepoint with sub-scores and total scores for all expected records for a subject. If the SDTM tabulation already followed the PROCCT and added records for individual missing item responses, some of the work to achieve this has already been implemented in the SDTM QS domain that will be used as a basis for the ADaM ADQS.
5. The PROCCT offers guidance pertaining to adding rows for items, scores, sub-scores, and total scores in ADaM after intercurrent events and during treatment pauses that varies based on the trial objective (clinical benefit vs safety/tolerability). It’s the authors recommendation at this time that the PROCCT ADaM advice should be implemented for cancer studies and seriously considered for other therapeutic areas. The decision for organizations should include a discussion of the following:
    a. Study Phase. Is the study a pivotal phase III? Will the study be a key component of a regulatory submission? How likely is the study to receive detailed regulatory scrutiny?

b. Is the PRO data part of the primary or secondary endpoint or will the PRO be used in labeling claims?

c. Complexity of the PRO instrument. Could the complexity introduce data quality risks that outweigh the benefits? Would the additional work required create undue burden on programming resources?

6. Ensure that the adrg.pdf for the study cites the use of the PROCCT. The author suggests adding details in the ADQS dataset portion of the document and in section 1.3 “Study Data Standards and Dictionary Inventory” table.
7. PROCCT guidance shows missing item responses added to SDTM coupled with the addition of PHANTOM records in ADaM. The author advises against combining these activities into only one of the models. For example, it isn’t recommended to use SDTM to add missing item responses and to add PHANTOM records to simplify the process.

CONCLUSION

The FDA guidance document "Submitting Patient-Reported Outcome Data in Cancer Clinical Trials" (PROCCT) establishes a critical principle for PRO data: "Understanding the reasons for and prevalence of missing PRO data are critical to support FDA review and regulatory decision-making."

This principle has far-reaching implications beyond oncology studies. By building upon existing CDISC standards, the PROCCT provides a comprehensive framework that enhances SDTM and ADaM data in ways that give regulatory reviewers deeper insight into:

1. The completeness of PRO assessments
2. Specific reasons for missing data
3. The relationship between missing data and key intercurrent events
4. How missing data were handled in analysis datasets

While the PROCCT is technically a non-binding document focused on cancer studies, its guidance has clear relevance for all therapeutic areas that utilize PRO endpoints—particularly when those endpoints support primary or secondary analyses. This broader applicability is supported by several trends:

- Both CDISC and the FDA have increasingly emphasized Patient Experience data
- PRO standards have evolved rapidly, as seen in CDISC QRS and FDA Patient-Focused Drug Delivery guidance
- The PROCCT's recommendations build upon, rather than contradict, existing CDISC standards

Looking forward, the author anticipates that additional guidance documents resembling the PROCCT will emerge in the coming years, further standardizing how PRO data should be handled across therapeutic areas. Since data collected today may not be submitted for regulatory review for several years, implementing PROCCT recommendations in critical program studies now represents a proactive approach that will likely benefit organizations in future submissions.

REFERENCES

Food and Drug Administration CDER SBIA. (2017). *Study Data Technical Conformance Guide (TCG) Webinar Series 2017* [PowerPoint Slides]. Retrieved from https://www.fda.gov/media/109146/download

Food and Drug Administration. (2020). *Patient-Focused Drug Development: Collecting Comprehensive and Representative Input* [PDF]. Retrieved from https://www.fda.gov/media/139088/download

Food and Drug Administration. (2023). *Submitting Patient-Reported Outcome Data in Cancer Clinical Trials* [PDF]. Retrieved from https://www.fda.gov/media/173581/download

Food and Drug Administration. (2023). *Submitting Clinical Trial Datasets and Documentation for Clinical Outcome Assessments Using Item Response Theory* [PDF]. Retrieved from https://www.fda.gov/media/173587/download

RECOMMENDED READING

- Drach, A. 2023. "The Phantom of the ADaM: Adding Missing Records to BDS Datasets." *Proceedings of the PharmaSUG 2023 Conference*.
- CDISC. (2022). *Study Data Tabulation Model Implementation Guide for Human Clinical Trials, version 3.4.* https://www.cdisc.org/system/files/members/standard/foundational/SDTMIG%20v3.4-FINAL_2022-07-21.pdf
- CDISC. (2021). *Analysis Data Model Implementation Guide, version 1.3.* https://www.cdisc.org/system/files/members/standard/foundational/ADaMIG_v1.3.pdf
- Food and Drug Administration. (2024). *Study Data Technical Conformance Guide, version Oct 2024*. https://www.fda.gov/media/153632/download
- CDISC. Questionnaires, Ratings, and Scales. https://www.cdisc.org/standards/foundational/qrs
- CDISC. Therapeutic Area Guides. https://www.cdisc.org/standards/therapeutic-areas
- CDISC. SDTM Metadata Submission Guidelines 2.0. https://www.cdisc.org/standards/foundational/sdtm/sdtm-metadata-submission-guidelines-v2-0

CONTACT INFORMATION

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

Charity Quick
Emergent BioSolutions
quickc@ebsi.com
https://www.linkedin.com/in/charity-quick-5896488/