TABLE OF CONTENTS

TABLE OF TABLES

1.2. Scope

This document provides specifications for the submission of Clinical Data Interchange Standards Consortium (CDISC) Study Data Tabulation Model (SDTM) and Analysis Data Model (ADaM) datasets and specifications for recommended tables and figures. These technical specifications aim to provide general guidelines for (1) standardized dataset content and structure and (2) recommended tables and figures to facilitate FDA review of the marketing application that the submitted data and analysis outputs are intended to support. The SDTM and ADaM specifications outlined in section 3.0 Overview of Dataset Specifications are not prescriptive and do not include an exhaustive list of all datasets, variables, and controlled terminologies to be submitted for FDA review. Further, the recommended tables and figures do not comprise all information needed to support FDA review of a marketing application. The dataset specifications and specifications for tables and figures are pursuant to discussions with FDA and may vary by clinical drug development program and clinical trial therein. These specifications are intended to be applicable to any PRO data used to inform the evaluation of (1) safety and tolerability or (2) clinical benefit in randomized studies (i.e., improvement in disease symptoms) within a cancer clinical trial.

This document does not pertain to submissions needed to support FDA review of the PRO measure itself or the proposed interpretation and use of scores generated by the PRO measure within the context of a specific clinical trial. Agreement on the PRO measure(s) used to collect study data and analyses of the resulting PRO data should be discussed with FDA as early as possible in a medical product development program, for example, prior to trial initiation. Sponsors are strongly encouraged to use the resources described in section 1.3 Relationships to Other Documents and to seek Agency input for confirmation and clarification as needed. Sponsors should consult with the Agency to determine which requested displays defined in section 4.0 Specifications for Tables and Figures apply to the PRO measure used within the marketing application. Lastly, although this guidance focuses on PRO measures, some of these recommendations may be relevant to other COAs (i.e., clinician-reported, observer-reported, and performance outcome measures) in cancer clinical trials.

1.3. Relationship to Other Documents

These technical specifications have been drafted in accordance with the business rules and assumptions outlined in the CDISC SDTM,[6] the SDTM Implementation Guide (SDTMIG),[7] the ADaM[8], and the ADaM Implementation Guide (ADaMIG). As new versions of the models and implementation guides become available, these technical specifications may be updated accordingly to maintain alignment. In addition, the FDA Study Data Technical Conformance Guide (sdTCG)[9] provides general specifications and recommendations for submitting datasets

[6] More information is available at CDISC's SDTM web page: https://cdisc.org/standards/foundational/sdtm.
[7] More information is available at CDISC's SDTMIG web page: https://cdisc.org/standards/foundational/sdtmig.
[8] More information is available at CDISC's ADaM web page: https://cdisc.org/standards/foundational/adam.
[9] More information is available at FDA's Study Data Standards Resources web page: https://www.fda.gov/industry/fda-resources-data-standards/study-data-standards-resources.

Additional Considerations:

Additional source data captured should be submitted either as additional rows with relevant QSTEST and QSTESTCD values within the QS dataset or within a Supplemental Questionnaires (SUPPQS) dataset. Additional content is dependent on the PRO measure administered, and if needed for analysis, should be copied into the applicable ADaM dataset. Example content captured within SUPPQS for each patient and assessment timepoint includes:

- **Data Collection Mode:** The mode of data collection used in the administration of the PRO measure, if differing from the protocol and/or varying across patients, assessment timepoints, or sites (e.g., clinical trial site, home). Examples of data collection mode may include paper-based administration, handheld electronic device, computer web-based application, or telephone-based administration.

- **Data Collector:** In cases where the measure is not self-administered (i.e., not independently completed by the patient without any assistance), the individual administering the PRO measure to the patient (e.g., caregiver, study staff member) by reading items to the patient and/or recording the patient's responses.

- **Language:** The language in which the measure was administered to the patient.

3.1.1.2 Handling of Missing PRO Data

Understanding the reasons for and prevalence of missing PRO data are critical to support FDA review and regulatory decision-making. Missing PRO data should be represented within the QS dataset with the reason for missingness captured under 'Reason Not Performed' (QSREASND). The QS dataset should include one record per item per PRO measure per patient per assessment timepoint, regardless of whether an item response is missing. When applicable, the QS dataset should also include one row per source data summary score per PRO measure per patient per assessment timepoint, regardless of whether the source data summary score is missing. Table 2 provides scenario-specific recommendations for displaying PRO data that are missing at a planned (i.e., per protocol) assessment timepoint. Appendix 5.1 demonstrates scenarios for representing missing data within the QS dataset. CDISC QRS Supplements[19] provide additional guidance on modeling missing data for named COA measures in SDTM datasets, including the modeling of timing variables.

Table 2. Recommended QS Representation of Missing PRO Data

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	The row for the missing source data summary score should include: • QSSTAT = 'NOT DONE'

[19] More information is available at CDISC's QRS web page: https://cdisc.org/standards/foundational/qrs.

Scenario	Recommended Representation in QS Dataset
responses (e.g., due to insufficient item response data).	• 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.

3.1.1.3 Handling of PRO Data Not Collected due to Use of Skip Logic or Computerized Adaptive Testing

Separate from missing data, PRO data may not be collected from the patient due to the use of skip logic or computerized adaptive testing to administer PRO items. When implemented, skip logic may be created based on the response to certain items. When the patient is not administered an item within a PRO measure due to the use of skip logic, the representation in the QS dataset should follow the guidance provided in the sdTCG.

Computerized adaptive testing (CAT) refers to a sequential form of individual testing administered by a computer in which successive items in the measure are selected for administration based primarily on the item's psychometric properties and content in relation to the patient's responses to previous items.[20] When the patient was not administered an item from an item bank (i.e., the total set of items from which a subset is selected for the patient during adaptive testing) for a PRO measure due to the use of CAT, a row for each remaining unadministered item within the item bank should not be included within the QS dataset. Rather, only the administered items for CAT-administered measures should be submitted within the QS dataset.

3.1.2 *Trial Summary Dataset*

Data related to the trial summary should be stored in the TS dataset. Of particular interest to FDA is the frequency with which these technical specifications are used to create and submit

[20] American Educational Research Association, American Psychological Association, National Council on Measurement in Education, 2014, The Standards for Educational and Psychological Testing, Washington (DC): American Educational Research Association Publications.

trial data. Per the FDA sdTCG, sponsors may include the following parameter and associated value in the TS dataset to indicate that these technical specifications were used for the study:

- TSPARAMCD = 'FDATCHSP'
- TSPARAM = 'FDA Tech Spec'
- TSVAL = 'Oncology PROs Technical Specifications Guidance v1.0'

3.2 ADaM Specifications

This section provides specifications for the ADaM dataset containing analysis-ready PRO data (referenced in this document as the 'ADQS' dataset), which are derived from PRO data in the SDTM QS dataset discussed in section 3.1.1 Questionnaires (QS) Dataset in conjunction with other SDTM and ADaM data.

3.2.1 *General Considerations*

The ADQS dataset described in this section follows the ADaM Basic Data Structure (BDS). In addition to variables for treatment assignment, stratification, subgrouping, and other covariates needed for analysis, the ADQS dataset should contain all individual item scores and summary scores (e.g., subscale scores, total scores, other composite or index scores). Table 3 contains specifications for a subset of ADQS variables, some of which are standard variables (included here to clarify how they should be completed for PRO measures used in oncology studies to foster consistency and standardization across industry as well as traceability) and some of which are newly defined. Table 3 does not include all ADQS variables to be submitted, such as timing and treatment variables. An example ADQS dataset is provided in Appendix 5.2.

Table 3. Specifications for a Subset of ADQS Variables

Variable Name	Variable Label	Type	Comments
PARCATy	Parameter Category y	Char	**PARCAT1:** The measure name(s) and version(s) should be provided within PARCAT1 for each item and summary score provided in PARAM to differentiate between PRO measures administered during the study. The measure name may match the value stored in the variable QS.QSCAT from the input SDTM QS dataset. **Additional PARCATy variables:** As demonstrated within section 3.2.3 ADQS Dataset Structure, the recommended number of PARCATy variables and their corresponding values differ based on the PRO measure and the number of summary scores calculated.
PARAM	Parameter	Char	The description of the analysis parameter (e.g., individual item or summary score). The value of PARAM may match the value stored in QS.QSTEST for parameters existing in the input SDTM QS dataset. Individual parameters are needed for each summary score. Documentation for derived parameters should be provided in submitted study metadata (e.g., the Define-XML file and the ADRG).

Variable Name	Variable Label	Type	Comments
EOTSTT	End of Treatment Status	Char	The patient's status as of the end of treatment or data cutoff. Examples: COMPLETED, DISCONTINUED, ONGOING.
DCTREAS	Reason for Discontinuation of Treatment	Char	Reason for a patient's discontinuation of treatment, if applicable. This variable represents discontinuation of treatment in the overall study and not discontinuation of treatment within individual treatment periods; reference the ADaMIG for period-specific discontinuation variables.
TRTDURD	Total Treatment Duration (Days)	Num	Total treatment duration as measured in days.
TRxxDURD	Treatment Duration in Period xx (Days)	Num	For multi-period studies, treatment duration for period xx as measured in days.
RANDDT	Date of Randomization	Num	Date of patient's randomization.
RANDFL	Randomized Population Flag	Char	Indicates whether the patient is included in the randomized population.
SAFFL	Safety Population Flag	Char	Indicates whether the patient is included in the safety population.
ITTFL	Intent-To-Treat Population Flag	Char	Indicates whether the patient is included in the intent-to-treat population.
Variables in ADQS copied from input SDTM QS Dataset			
QSSEQ	Sequence Number	Num	Sponsors should include any SDTM variables in the ADQS dataset needed to provide traceability to the source SDTM QS dataset.
VISIT	Visit Name	Char	
VISITNUM	Visit Number	Num	
QSSTAT	Completion Status	Char	Sponsors should include SDTM variables that provide explanations for missing item scores or missing source data summary scores. See Comments for QSSTAT and QSREASND provided in Table 1. Specifications for a Subset of QS Variables.
QSREASND	Reason Not Performed	Char	

3.2.2 *Handling of Missing PRO Data and Intercurrent Events*

As discussed in SDTM section 3.1.1.2 Handling of Missing PRO Data and section 3.1.1.3 Handling of PRO Data Not Collected due to Skip Logic or Computerized Adaptive Testing, understanding the reasons for and prevalence of missing PRO data or PRO data not collected, as well as intercurrent events[21] occurring during the study, are critical to support FDA review and regulatory decision-making. Approaches to represent missing data and intercurrent events within the ADQS dataset are provided in section 3.2.2.1 Approaches to Represent Missing Data and Intercurrent Events and scenarios to represent missing PRO data and intercurrent events within the ADQS dataset are provided in section 3.2.2.2 Scenarios to Represent Missing Data and Intercurrent Events which further depend on the PRO objective (i.e., clinical benefit vs. safety and tolerability). When provided, the reason for missing PRO data should be distributed across all rows for each item score and summary score within the missing PRO measure at each planned (per protocol) PRO assessment timepoint. The example provided in Appendix 5.2 illustrates the

[21] See the guidance for industry *E9(R1) Statistical Principles for Clinical Trials Addendum: Estimands and Sensitivity Analysis in Clinical Trials* (May 2021). For the most recent version of a guidance, check the FDA guidance web page at https://www.fda.gov/RegulatoryInformation/Guidances/default.htm.

When the PRO objective is to inform the evaluation of safety and tolerability, rows are only represented in the ADQS dataset for each item score and summary score within the PRO measure for patients where the PRO is expected to be completed at the planned (per protocol) PRO assessment timepoints, including rows for patients who are on therapy throughout the study and rows for patients at assessment timepoints prior to treatment discontinuation or death. The number of PRO assessments to be administered after a patient discontinues treatment should be carefully considered and minimized to reduce patient burden. Rationale regarding the number of PRO assessments to be administered should be provided, given that the administration of PRO measures beyond treatment discontinuation can be challenging in cancer clinical trials. Sponsors are strongly encouraged to consult with the Agency to make appropriate determinations. Rows would not be created in the ADQS dataset for assessment timepoints after a patient's death or for any PRO assessment timepoints for randomized but not treated patients. When rows are created in the ADQS dataset for the applicable patients and assessment timepoints when informing the evaluation of safety and tolerability, rows should be created for each item score and summary score within the PRO measure, regardless of whether the item score or summary score has a value populated.

Regardless of PRO objective, if a patient pauses treatment and does not have records included in the QS dataset, these records should be derived as phantom records in the ADQS dataset. In addition, there are certain scenarios regardless of PRO objective where the PRO measure is not administered to a patient, and missing data rows do not need to be created for the patient in the ADQS dataset such as when the PRO measure is not available in the patient's language.

3.2.3 *ADQS Dataset Structure*

Example dataset structures are described below for PRO measures where individual item scores are used to calculate at least one summary score and where item scores are analyzed individually.

3.2.3.1 Dataset Structure for PRO Measures Where Summary Scores are Calculated

The preferred dataset structure for PRO measures where individual item scores are used to calculate summary scores includes the consistent use of parameter category variables (e.g., PARCAT1, PARCAT2) to allow for the easy identification of each measure, score, and item. The scenarios below within Tables 4-7 illustrate how categorical variables are recommended for use based on the complexity of relationships between items, subscales, and higher-level scales within the PRO measure.

As described within Table 3: Specifications for a Subset of ADQS Variables, PARCAT1 reports the PRO measure name and version. PARCAT2 is created within the ADQS dataset for PRO measures where summary scores are calculated to indicate whether PARAM represents an item or a summary score, where the summary score calculated is dependent on the instrument scoring manual. Example terminology values for PARCAT2 include, but are not limited to, 'ITEM', 'SUBSCALE SCORE', 'SCALE SCORE', 'RAW SCORE', 'TOTAL SCORE', 'COMPOSITE SCORE', and 'INDEX SCORE'. The number of additional PARCATy variables within the

ADQS dataset depends on the summary scores to be calculated based on the intended use of the PRO measure as described in the protocol, SAP, and instrument scoring manual.

Scenario 1 within Table 4 represents a simple scenario where items within a two-item measure are used to compute a single total score.

Table 4. ADQS Dataset Structure for Scenario 1

PARCAT1	PARCAT2	PARAM
Measure Name and Version	ITEM	Item 1
Measure Name and Version	ITEM	Item 2
Measure Name and Version	TOTAL SCORE	Total Score

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

Scenario 3 within Table 6 represents a scenario where (1) multiple subscale scores are calculated, and each subscale score is calculated from distinct, mutually exclusive item score(s), and (2) multiple scale scores are calculated, and each scale score is calculated from distinct, mutually exclusive subscale score(s). PARCAT3 reports the subscale to which an item contributes, and PARCAT4 reports the scale to which a subscale contributes.

Table 6. ADQS Dataset Structure for Scenario 3

PARCAT1	PARCAT2	PARCAT3	PARCAT4	PARAM
Measure Name and Version	ITEM	Subscale Score 1	Scale Score A	Item 1
Measure Name and Version	ITEM	Subscale Score 1	Scale Score A	Item 2
Measure Name and Version	ITEM	Subscale Score 1	Scale Score A	Item 3
Measure Name and Version	ITEM	Subscale Score 2	Scale Score A	Item 4
Measure Name and Version	ITEM	Subscale Score 2	Scale Score A	Item 5
Measure Name and Version	ITEM	Subscale Score 3	Scale Score B	Item 6
Measure Name and Version	SUBSCALE SCORE	Subscale Score 1	Scale Score A	Subscale Score 1
Measure Name and Version	SUBSCALE SCORE	Subscale Score 2	Scale Score A	Subscale Score 2
Measure Name and Version	SUBSCALE SCORE	Subscale Score 3	Scale Score B	Subscale Score 3
Measure Name and Version	SCALE SCORE		Scale Score A	Scale Score A
Measure Name and Version	SCALE SCORE		Scale Score B	Scale Score B

Scenario 4 within Table 7 represents a scenario where (1) multiple subscale scores are calculated, and each subscale is calculated from distinct, mutually exclusive item score(s), and (2) multiple scale scores are calculated. However, in this scenario, a single subscale score can contribute to multiple scale scores. PARCAT3 reports the subscale to which an item contributes and separate categorical variables (i.e., PARCAT4 and PARCAT5) report the individual scale to which a subscale contributes. Similarly, if an item can contribute to multiple subscales, a separate PARCATy is needed for each subscale (not pictured).

Table 7. ADQS Dataset Structure for Scenario 4

PARCAT1	PARCAT2	PARCAT3	PARCAT4	PARCAT5	PARAM
Measure Name and Version	ITEM	Subscale Score 1	Scale Score A	Scale Score B	Item 1
Measure Name and Version	ITEM	Subscale Score 1	Scale Score A	Scale Score B	Item 2
Measure Name and Version	ITEM	Subscale Score 1	Scale Score A	Scale Score B	Item 3
Measure Name and Version	ITEM	Subscale Score 2	Scale Score A	Scale Score B	Item 4
Measure Name and Version	ITEM	Subscale Score 2	Scale Score A	Scale Score B	Item 5
Measure Name and Version	ITEM	Subscale Score 3	Scale Score A		Item 6
Measure Name and Version	SUBSCALE SCORE	Subscale Score 1	Scale Score A	Scale Score B	Subscale Score 1
Measure Name and Version	SUBSCALE SCORE	Subscale Score 2	Scale Score A	Scale Score B	Subscale Score 2
Measure Name and Version	SUBSCALE SCORE	Subscale Score 3	Scale Score A		Subscale Score 3
Measure Name and Version	SCALE SCORE		Scale Score A		Scale Score A
Measure Name and Version	SCALE SCORE			Scale Score B	Scale Score B

3.2.3.2 Dataset Structure for PRO Measures Where Summary Scores are Not Calculated

The preferred dataset structure when summary scores are not calculated and item scores are analyzed individually include PRO measures that analyze attributes of symptomatic AEs, where AEs are selected from an item library prior to trial initiation. Scenario 5 within Table 8 represents the consistent use of parameter category variables to allow for the easy identification of the measure, symptoms (e.g., rash, headache), and attributes (e.g., presence, severity, frequency) within an example PRO measure. PARCAT2 reports the stand-alone symptom and PARCAT3 reports the stand-alone attribute. PARAM is a compliant, stand-alone analysis variable. CDISC Controlled Terminology is implemented when available for the PRO measure.

Disposition for Safety and Tolerability (i.e., where PRO Expected Flag (PROEXPFL) equals 'Y' as defined in Table 3: Specifications for a Subset of ADQS Variables). Thus, the denominator can decrease throughout an oncology trial based on attrition over time. The table and data visualization should be provided at the PRO instrument level based on completion rate. The sponsor should consult with the Agency to determine additional tables and data visualizations to provide for individual items (including patient-reported symptomatic adverse events) and/or summary scores. An example table and bar chart illustrating the structure and inputs are provided in Appendix 5.3.4, with percentages provided for each category within the bar chart. Reasons for treatment discontinuation may be excluded as reasons for PRO noncompletion in the table and figure as described in the table footnotes.

4.3 Distributions

Provide a table and data visualization summarizing the distribution of responses and the distribution of change in responses from baseline. When the PRO measure is used to inform the evaluation of safety and tolerability, the sponsor should consult with the Agency to determine the tables and data visualizations to provide for individual items (including patient-reported symptomatic adverse events) and/or summary scores. When the PRO measure is to evaluate clinical benefit, the table and data visualization should be provided for the concept(s) evaluated by the PRO measure (e.g., single item score, subscale score, total score). When the concept measured is a summary score (e.g., physical functioning), the table and data visualization should be provided for both the summary score and for each individual item score that contributes to the summary score.

Within the table, counts and percentages should be provided for PRO Completed and PRO Not Completed. Additional columns within the table depends on PRO objective; when the PRO objective is safety and tolerability, the count and percentage for PRO Expected should be provided. When the PRO objective is clinical benefit, the count and percentage for the randomized population should be provided. When the PRO objective is safety and tolerability, the denominator used to determine the percentage for PRO Completed and PRO Not Completed is the number of patients expected to complete the PRO measure at the designated PRO assessment timepoint as described in section 4.1.2 Patient Disposition for Safety and Tolerability (i.e., where PRO Expected Flag (PROEXPFL) equals 'Y' as defined in Table 3: Specifications for a Subset of ADQS Variables). When the PRO objective is clinical benefit, the denominator used to determine the percentage for PRO Completed and PRO Not Completed is the randomized population.

When the concept measured has categorical response options with binary or ordinal outcomes, counts and percentages within the table for the response categories or change in response categories should be provided where the denominator used to determine the percentage for each category is PRO Completed (i.e., where PRO Score Completed Flag (PROSCMFL) equals 'Y' as defined in Table 3: Specifications for a Subset of ADQS Variables). When the concept measured has continuous response options, summary statistics should be provided and are calculated based on PRO Completed.

Within the data visualization, counts for PRO Completed and PRO Not Completed should be provided below the figure as shown in the examples within the appendices. In addition, if the PRO objective is safety and tolerability, counts for PRO Expected should be provided; if the PRO objective is clinical benefit, counts for the randomized population should be provided. The type of data visualization provided further depends on the response options for the PRO measure. A bar chart should be provided when the concept measured has categorical response options with binary or ordinal outcomes. When the concept measured has continuous response options, a line graph should be provided. For example, a line graph with descriptive means may be provided when the PRO objective is safety and tolerability objective. The sponsor may consult with the Agency to discuss alternative data visualizations (e.g., density curves, box plots) to provide when continuous data are captured. Within the line graph, standard error bars with jittering should be provided as well as labels to indicate improving/worsening or higher/lower functioning, depending on the concept measured.

4.3.1 *Distribution of Responses*

When the concept measured has categorical response options with a binary or ordinal outcome, the distribution of responses in the table and bar chart should include all possible response options for the item or summary score with the percentage provided for each category within the bar chart. Within Appendix 5.3.5, an example table is provided in Table A8 and a bar chart is provided in Figure A5 for a single item for a safety and tolerability PRO objective.

When the concept measured has continuous response options, the distribution of responses in the table and line graph should represent a summary statistic (e.g., the mean score) over time. An example table is provided in Table A9, and a line graph is provided in Figure A6 for a single item showing descriptive means for a safety and tolerability PRO objective.

4.3.2 *Distribution of Change in Responses from Baseline*

When the concept measured includes categorical response options with a binary or ordinal outcome, the distribution of change in response categories for the table and bar chart should include all possible scenarios (e.g., improving, no change, worsening) based on the number of response categories for the item or summary score with the percentage provided for each change category within the bar chart. Within Appendix 5.3.6, an example table is provided in Table A10 and a bar chart is provided in Figure A7 for a single item for a safety and tolerability PRO objective. Labels for 'No Change or Improving' and 'Worsening' should be provided in the bar chart as shown in the example.

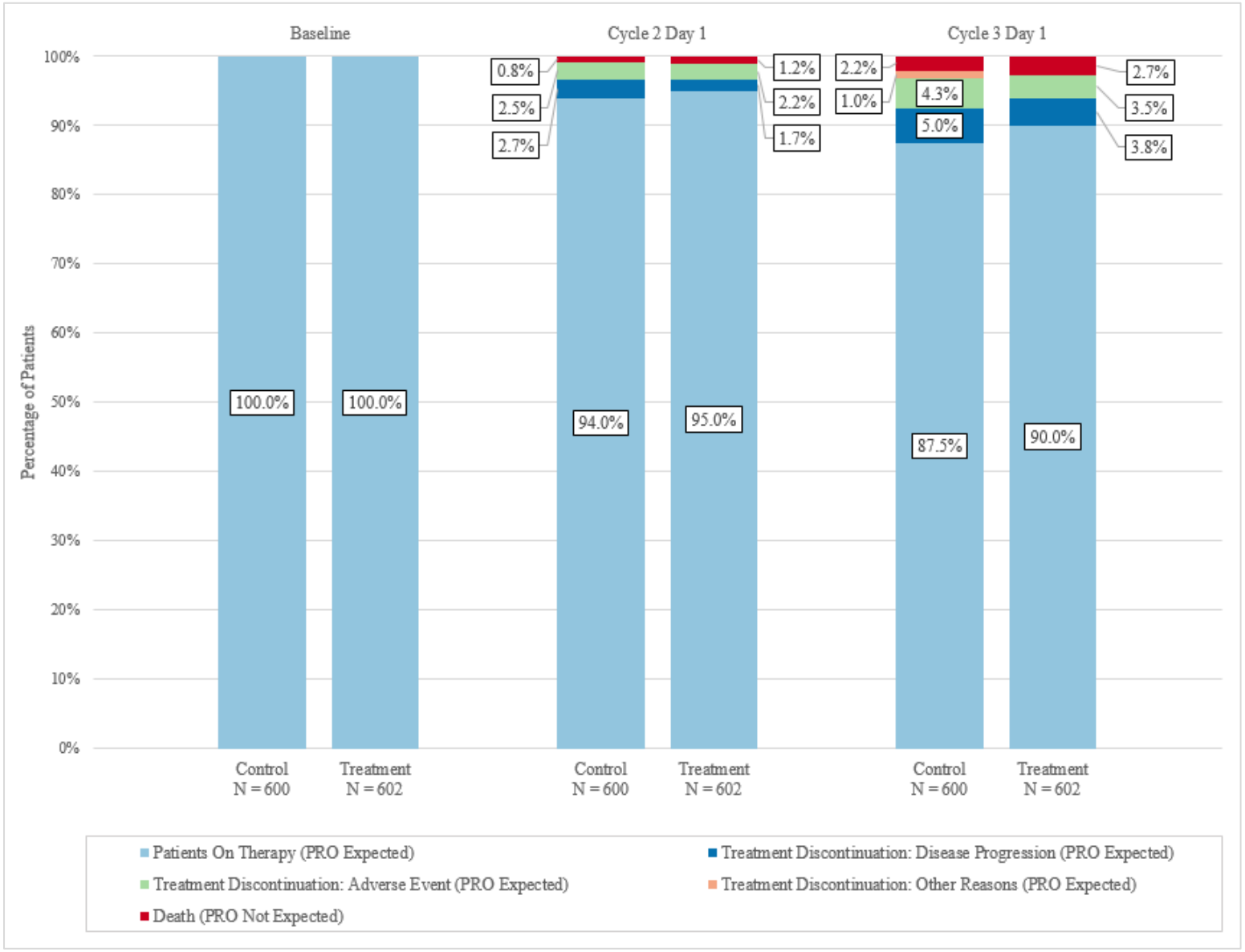
When the concept measured has continuous response options, the distribution of change in responses from baseline in the table and line graph should represent a summary statistic (e.g., mean) change from baseline over time. An example table is provided in Table A11 and a line graph is provided in Figure A8 for a single item showing descriptive mean change from baseline for a safety and tolerability PRO objective.

4.4 Incidence of Healthcare Utilization

Provide a table summarizing the incidence of healthcare utilization, including emergency department visits, hospitalizations, supportive care medications, supportive care procedures, and other relevant interventions depending on the study. Each supportive care medication and supportive care procedure should be represented as a separate column within the table. The denominator used to determine the percentage for each intervention depends on PRO objective. When the PRO objective is safety and tolerability, the denominator is the number of patients expected to complete the PRO measure at the designated PRO assessment timepoint as described in section 4.1.2 Patient Disposition for Safety and Tolerability (i.e., where PRO Expected Flag (PROEXPFL) equals 'Y' as defined in Table 3: Specifications for a Subset of ADQS Variables). When the PRO objective is clinical benefit, the denominator is the randomized population. In addition to columns for healthcare utilization interventions, columns are provided for Randomized Patients and PRO Expected.

An example table illustrating the structure and inputs is provided in Appendix 5.3.7. The table contains example healthcare utilization values for supportive procedures and supportive medications, but the table should be modified to represent the intervention within the study. The sponsor may consult with the Agency to determine the appropriate supportive care medications and supportive care procedures to provide (e.g., growth factors, steroids, and transfusions depend on the cancer type and may not always be relevant).

Figure A1. Patient Disposition when Evaluating Clinical Benefit (Denominator = Randomized Population)



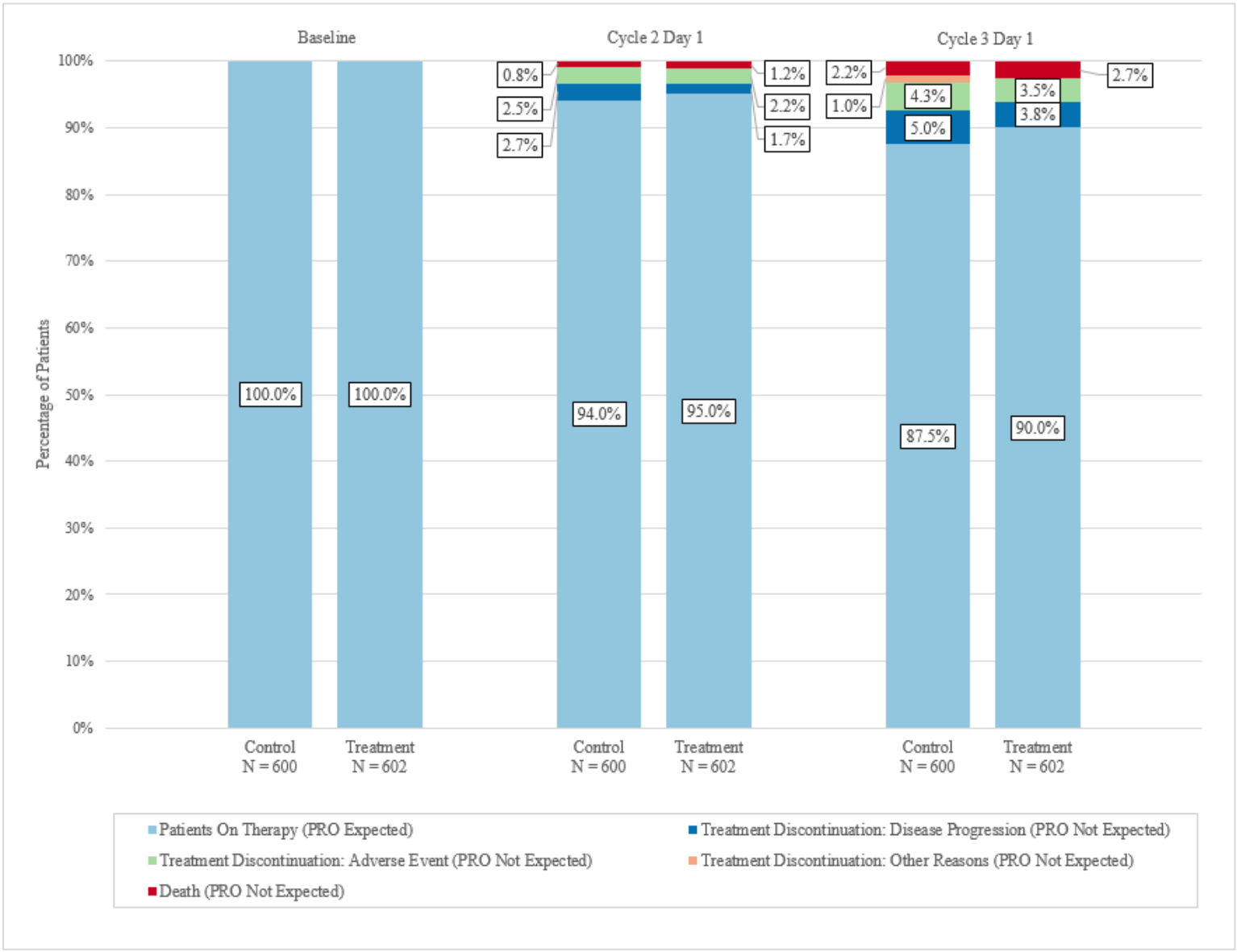
5.3.2 *Patient Disposition when Informing Safety and Tolerability*

Table A5. Patient Disposition when Informing the Evaluation of Safety and Tolerability (Denominator = Safety Population)[4]

Analysis Visit	Treatment Arm	Randomized Population (N)	Safety Population (N)	PRO Expected,[5] n (%)	PRO Not Expected				
					Death, n (%)	Treatment Discontinuation: Disease Progression, n (%)	Treatment Discontinuation: Adverse Event, n (%)	Treatment Discontinuation: Other Reasons, n (%)	Other,[6] n (%)
Baseline	Control	600	600	600 (100.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
	Treatment	602	602	602 (100.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Cycle 2 Day 1	Control	600	600	564 (94.0%)	5 (0.8%)	16 (2.7%)	15 (2.5%)	0 (0.0%)	0 (0.0%)
	Treatment	602	602	572 (95.0%)	7 (1.2%)	10 (1.7%)	13 (2.2%)	0 (0.0%)	0 (0.0%)
Cycle 3 Day 1	Control	600	600	525 (87.5%)	13 (2.2%)	30 (5.0%)	26 (4.3%)	6 (1.0%)	0 (0.0%)
	Treatment	602	602	542 (90.0%)	16 (2.7%)	23 (3.8%)	21 (3.5%)	0 (0.0%)	0 (0.0%)

[4] Denominator used to calculate percentages is the number of patients in the safety population.
[5] When PRO data are used to inform the evaluation of safety and tolerability, the PRO measure may not be expected to be completed after a patient discontinues from treatment as shown in the table. Thus, the PRO Expected column excludes patients who discontinued from treatment. The PRO Expected column is determined where PRO Expected Flag (PROEXPFL) equals 'Y' and includes patients who both completed and did not complete the PRO measure (e.g., the patient did not attend an onsite visit, the patient did not complete the PRO measure at the attended onsite visit or at a prespecified offsite assessment timepoint, the patient partially completed the PRO measure resulting in incalculable summary scores).
[6] The Other column groups patients who were not expected to complete the PRO measure at a designated assessment timepoint for reasons other than treatment discontinuation or patient death (e.g., the translation of the PRO measure is not available in the patient's language).

Figure A2. Patient Disposition when Informing the Evaluation of Safety and Tolerability (Denominator = Safety Population)



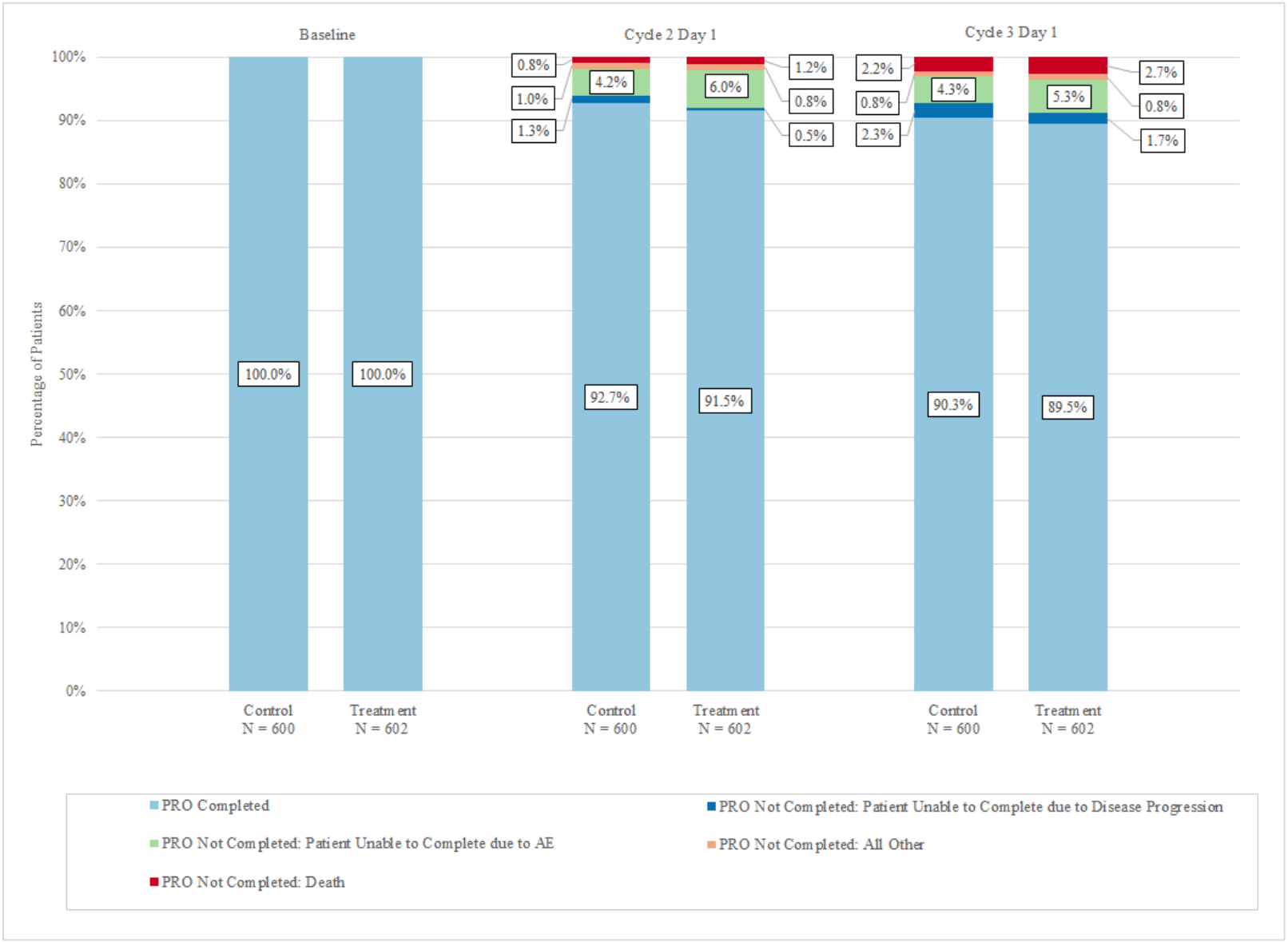
5.3.3 *Available Data Rate for Clinical Benefit*

Table A6. Available Data Rate for Clinical Benefit (Denominator = Randomized Population)[7]

Analysis Visit	Treatment Arm	Randomized Patients (N)	PRO Completed, n (%)	PRO Not Completed[8] (excluding Death), n (%)	Reason for PRO Not Completed,[9] n (%)					Death, n (%)
					Patient Unable to Complete due to Disease Progression	Patient Unable to Complete due to Adverse Event (AE)	Patient Refusal	Device Failure	Reason Unknown,[10] n (%)	
Baseline	Control	600	600 (100.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
	Treatment	602	602 (100.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Cycle 2 Day 1	Control	600	556 (92.7%)	39 (6.5%)	8 (1.3%)	25 (4.2%)	6 (1.0%)	0 (0.0%)	0 (0.0%)	5 (0.8%)
	Treatment	602	551 (91.5%)	44 (7.3%)	3 (0.5%)	36 (6.0%)	5 (0.8%)	0 (0.0%)	0 (0.0%)	7 (1.2%)
Cycle 3 Day 1	Control	600	542 (90.3%)	45 (7.5%)	14 (2.3%)	26 (4.3%)	0 (0.0%)	5 (0.8%)	0 (0.0%)	13 (2.2%)
	Treatment	602	539 (89.5%)	47 (7.8%)	10 (1.7%)	32 (5.3%)	5 (0.8%)	0 (0.0%)	0 (0.0%)	16 (2.7%)

[7] Denominator used to calculate percentages is the number of randomized patients.
[8] The PRO Not Completed column is calculated where PROEXPFL = 'Y' and PROSCMFL is null; thus, death is excluded from the counts and percentages and is provided as a standalone column.
[9] When PRO data are used to evaluate clinical benefit, reasons for PRO Not Completed are based on collected data from QS.QSREASND within the study. All reasons for noncompletion collected during the study should be included. Counts can include patients who were on therapy and who discontinued; thus, counts in Table A6 may be larger than counts in Table A7 given that patients complete the PRO measure after treatment discontinuation when evaluating clinical benefit.
[10] Unknown reasons, if present, should be tabulated within the separate 'Reason Unknown' column.

Figure A3. Available Data Rate for Clinical Benefit (Denominator = Randomized Population)



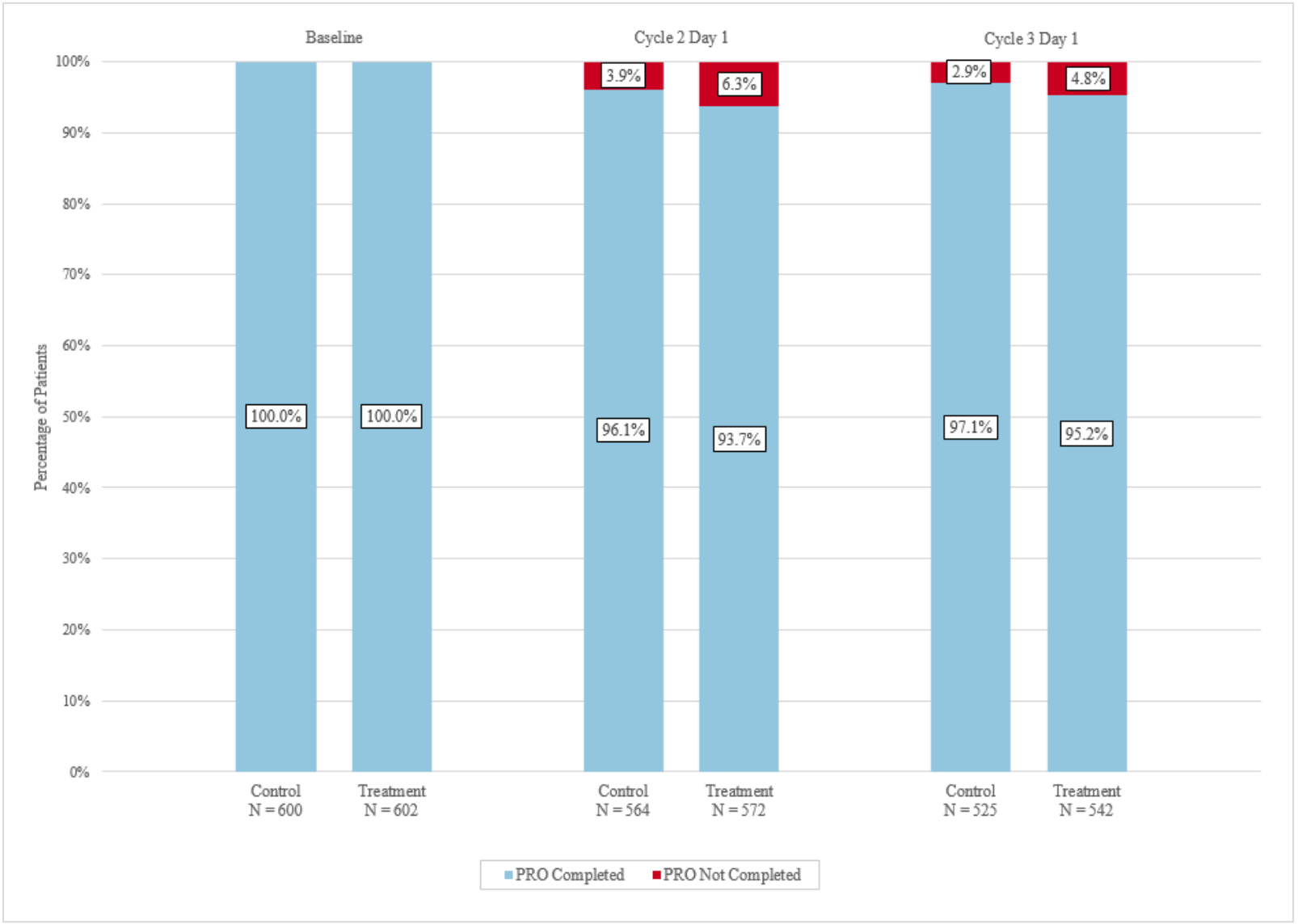
5.3.4 *Completion Rate for Safety and Tolerability*

Table A7. Completion Rate for Safety and Tolerability (Denominator = PRO Expected Population)[11]

Analysis Visit	Treatment Arm	PRO Expected[12] (N)	PRO Completed, n (%)	PRO Not Completed, n (%)	Reason for PRO Not Completed,[13] n (%)			
					Patient Refusal	Patient Unable to Complete due to AE	Device Failure	Reason Unknown,[14] n (%)
Baseline	Control	600	600 (100.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
	Treatment	602	602 (100.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Cycle 2 Day 1	Control	564	542 (96.1%)	22 (3.9%)	6 (1.1%)	16 (2.8%)	0 (0.0%)	0 (0.0%)
	Treatment	572	536 (93.7%)	36 (6.3%)	5 (0.9%)	31 (5.4%)	0 (0.0%)	0 (0.0%)
Cycle 3 Day 1	Control	525	510 (97.1%)	15 (2.9%)	0 (0.0%)	10 (1.9%)	5 (1.0%)	0 (0.0%)
	Treatment	542	516 (95.2%)	26 (4.8%)	5 (0.9%)	21 (3.9%)	0 (0.0%)	0 (0.0%)

[11] Denominator used to calculate percentages is PRO Expected.
[12] When PRO data are used to inform the evaluation of safety and tolerability, the PRO measure may not be expected to be completed after a patient discontinues from treatment as shown in the table. The PRO Expected column excludes patients who discontinued from treatment and is determined where PRO Expected Flag (PROEXPFL) equals 'Y'.
[13] Reasons for PRO Not Completed are based on collected data within the study and represent reasons the PRO measure was not completed when the patient did not discontinue from treatment. All reasons for noncompletion collected during the study should be included.
[14] Unknown reasons, if present, should be tabulated within the separate 'Reason Unknown' column.

Figure A4. Completion Rate for Safety and Tolerability (Denominator = PRO Expected Population)



5.3.5 *Distribution of Responses*

Table A8. Distribution of Categorical Responses for Item 1 (Safety and Tolerability Example)[15]

Analysis Visit	Treatment Arm	PRO Expected[16]	PRO Completed, n (%)	PRO Not Completed, n (%)	Response Categories,[17] n (%)			
					Not at all	A little	Quite a bit	Very much
Baseline	Control	600	600 (100.0%)	0 (0.0%)	332 (55.3%)	220 (36.7%)	31 (5.2%)	17 (2.8%)
	Treatment	602	602 (100.0%)	0 (0.0%)	313 (52.0%)	228 (37.9%)	38 (6.3%)	23 (3.8%)
Cycle 2 Day 1	Control	564	542 (96.1%)	22 (3.9%)	299 (55.2%)	188 (34.7%)	34 (6.3%)	21 (3.9%)
	Treatment	572	536 (93.7%)	36 (6.3%)	268 (50.0%)	199 (37.1%)	41 (7.6%)	28 (5.2%)
Cycle 3 Day 1	Control	525	510 (97.1%)	15 (2.9%)	225 (44.1%)	189 (37.1%)	63 (12.4%)	33 (6.5%)
	Treatment	542	516 (95.2%)	26 (4.8%)	203 (39.3%)	193 (37.4%)	71 (13.8%)	49 (9.5%)

[15] Denominator used to calculate percentages for PRO Completed and PRO Not Completed is PRO Expected. Denominator used to calculate percentages for each response category is PRO Completed.

[16] When PRO data are used to inform the evaluation of safety and tolerability, the PRO measure may not be expected to be completed after a patient discontinues from treatment as shown in the table. The PRO Expected column excludes patients who discontinued from treatment and is determined where PRO Expected Flag (PROEXPFL) equals 'Y'.

[17] The example response categories represent the response options for the item.

Figure A5. Distribution of Categorical Responses for Item 1 (Safety and Tolerability Example where Denominator = PRO Completed)

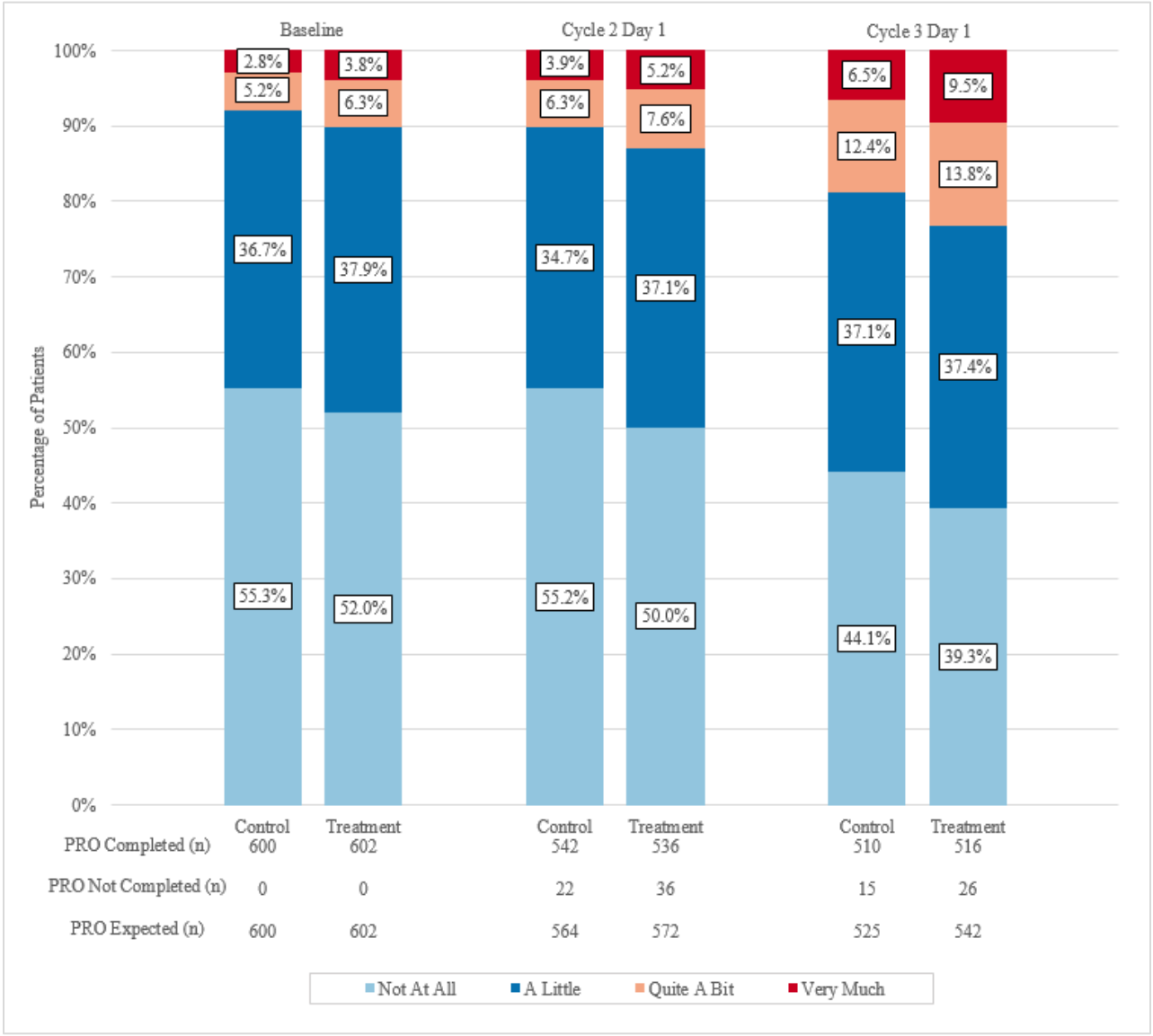


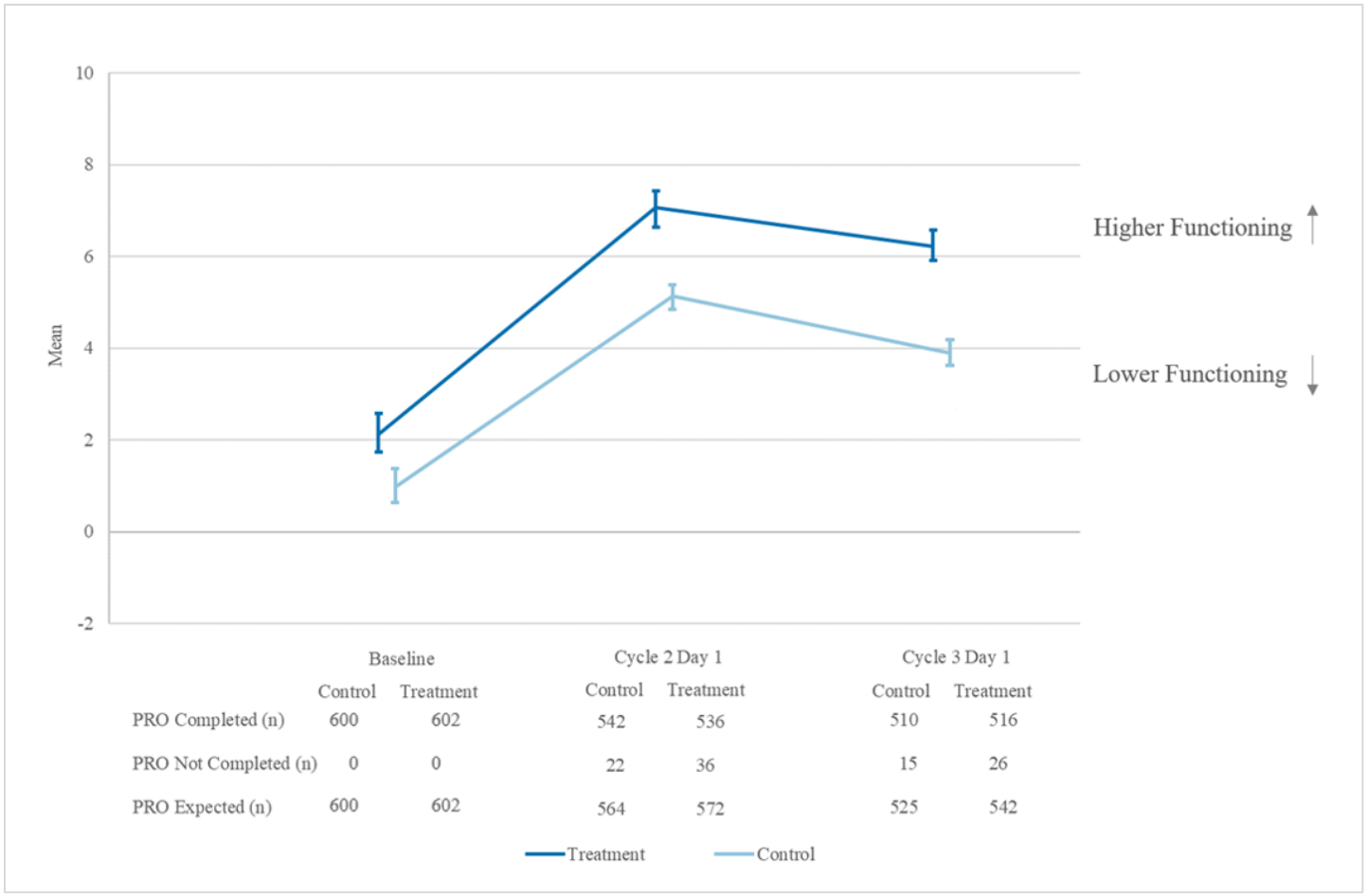
Table A9. Summary Statistics for Item 2 with Continuous Response Options (Safety and Tolerability Example)[18]

Analysis Visit		Control	Treatment
Baseline	PRO Expected[19] (N)	600	602
	PRO Not Completed, n (%)	0 (0.0%)	0 (0.0%)
	PRO Completed, n (%)	600 (100.0%)	602 (100.0%)
	Summary Statistics[20]		
	Mean	2.1	1.0
	Standard Deviation	1.8	0.9
	Standard Error	0.07	0.04
	Median	2.1	1.0
	Minimum	0.0	0.0
	Maximum	4.1	2.0
Cycle 2 Day 1	PRO Expected (N)	564	572
	PRO Not Completed, n (%)	22 (3.9%)	36 (6.3%)
	PRO Completed, n (%)	542 (96.1%)	536 (93.7%)
	Summary Statistics		
	Mean	7.1	5.1
	Standard Deviation	4.6	3.7
	Standard Error	0.19	0.15
	Median	7.2	5.1
	Minimum	0.3	0.2
	Maximum	11.8	9.8
Cycle 3 Day 1	PRO Expected (N)	525	542
	PRO Not Completed, n (%)	15 (2.9%)	26 (4.8%)
	PRO Completed, n (%)	510 (97.1%)	516 (95.2%)
	Summary Statistics		
	Mean	6.2	3.9
	Standard Deviation	5.2	2.7
	Standard Error	0.23	0.12
	Median	6.6	3.8
	Minimum	0.1	0.0

[18] Denominator used to calculate percentages for PRO Completed and PRO Not Completed is PRO Expected.
[19] When PRO data are used to inform the evaluation of safety and tolerability, the PRO measure may not be expected to be completed after a patient discontinues from treatment as shown in the table. The PRO Expected column excludes patients who discontinued from treatment and is determined where PRO Expected Flag (PROEXPFL) equals 'Y'.
[20] Summary Statistics are calculated based on PRO Completed.

	Maximum	11.4	10.0

Figure A6. Descriptive Means for Item 2 with Continuous Response Options (Safety and Tolerability Example for Physical Functioning)[21,22]



[21] The 'Higher Functioning' and 'Lower Functioning' labels are provided as examples for physical functioning. Labels provided within the figure should accurately represent the concept being measured.
[22] Error bars based on a 95% confidence internal for the mean are represented within the line plot.

5.3.6 *Distribution of Change in Responses from Baseline*

Table A10. Distribution of Change in Response Categories from Baseline for Item 1 (Safety and Tolerability Example)[23]

Analysis Visit	Treatment Arm	PRO Expected[24]	PRO Completed, n (%)	PRO Not Completed, n (%)	Change in Response Categories,[25] n (%)						
					Improving 1	Improving 2	Improving 3	No Change	Worsening 1	Worsening 2	Worsening 3
Cycle 2 Day 1	Control	564	542 (96.1%)	22 (3.9%)	38 (7.0%)	11 (2.0%)	3 (0.6%)	303 (55.9%)	132 (24.4%)	38 (7.0%)	17 (3.1%)
	Treatment	572	536 (93.7%)	36 (6.3%)	33 (6.2%)	14 (2.6%)	6 (1.1%)	296 (55.2%)	141 (26.3%)	32 (6.0%)	14 (2.6%)
Cycle 3 Day 1	Control	525	510 (97.1%)	15 (2.9%)	50 (9.8%)	24 (4.7%)	10 (2.0%)	261 (51.2%)	126 (24.7%)	29 (5.7%)	10 (2.0%)
	Treatment	542	516 (95.2%)	26 (4.8%)	44 (8.5%)	28 (5.4%)	11 (2.1%)	261 (50.6%)	123 (23.8%)	39 (7.6%)	10 (1.9%)

[23] Denominator used to calculate percentages for PRO Completed and PRO Not Completed is PRO Expected. Denominator used to calculate percentages for each change in response category is PRO Completed.
[24] When PRO data are used to inform the evaluation of safety and tolerability, the PRO measure may not be expected to be completed after a patient discontinues from treatment as shown in the table. The PRO Expected column excludes patients who discontinued from treatment and is determined where PRO Expected Flag (PROEXPFL) equals 'Y'.
[25] The example response categories represent the response options for the item.

Figure A7. Distribution of Change in Response Categories from Baseline for Item 1 (Safety and Tolerability Example where Denominator = PRO Completed)

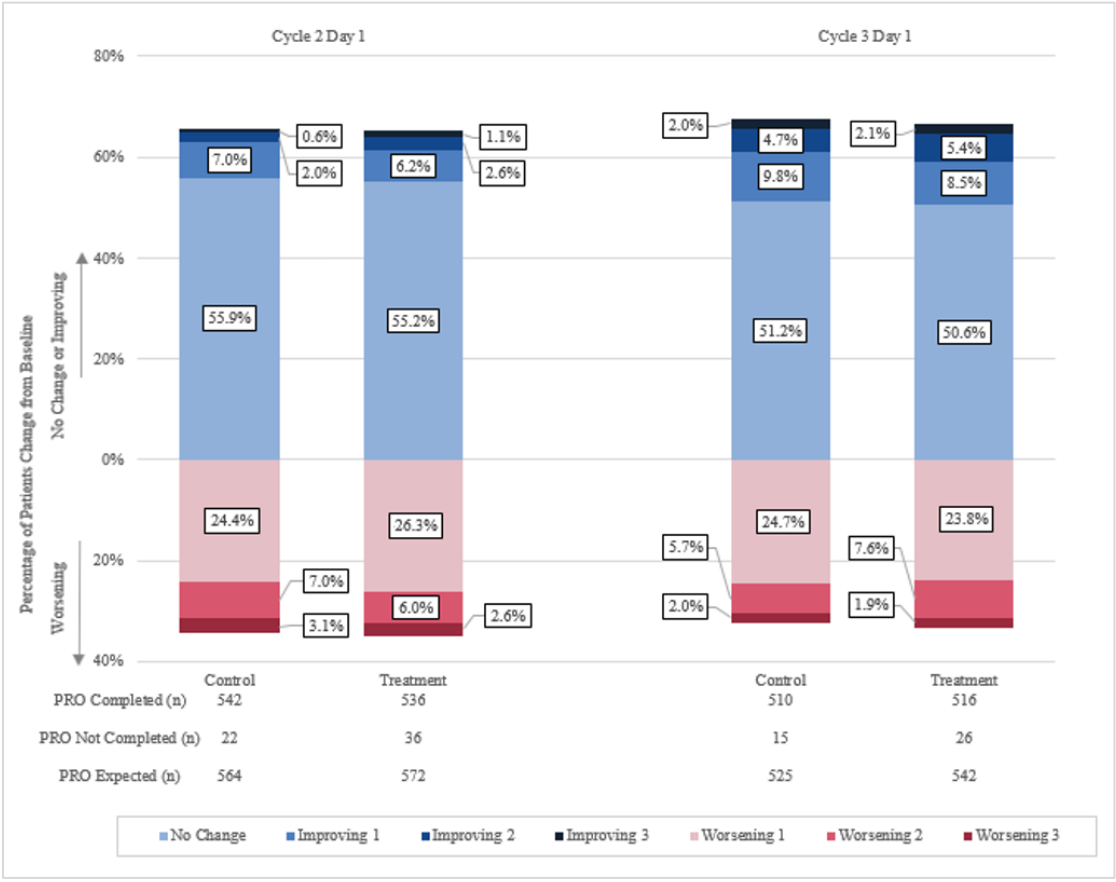


Table A11. Change from Baseline for Item 2 with Continuous Response Options (Safety and Tolerability Example)[26]

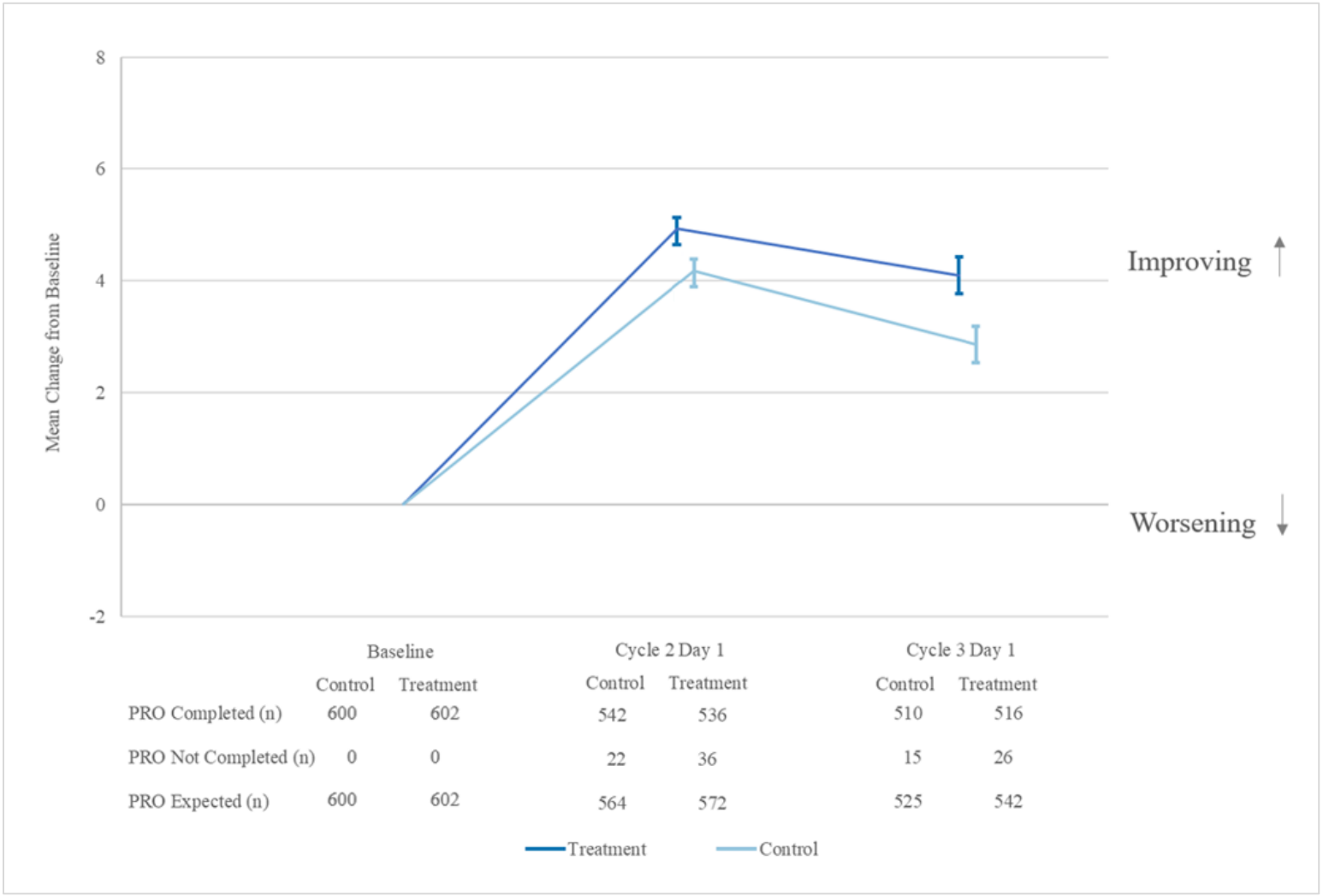
Analysis Visit		Treatment	Control
Cycle 2 Day 1	PRO Expected[27] (N)	564	572
	PRO Not Completed, n (%)	22 (3.9%)	36 (6.3%)
	PRO Completed, n (%)	542 (96.1%)	536 (93.7%)
	Summary Statistics[28]		
	Mean	4.9	4.5
	Standard Deviation	4.0	1.7
	Standard Error	0.17	0.19
	Median	5.0	4.2
	Minimum	-1.1	-1.3
	Maximum	10.3	9.0
Cycle 3 Day 1	PRO Expected[1] (N)	525	542
	PRO Not Completed, n (%)	15 (2.9%)	26 (4.8%)
	PRO Completed, n (%)	510 (97.1%)	516 (95.2%)
	Summary Statistics		
	Mean	4.1	2.9
	Standard Deviation	5.7	5.6
	Standard Error	0.25	0.24
	Median	4.2	2.9
	Minimum	-1.6	-1.4
	Maximum	8.0	8.5

[26] Denominator used to calculate percentages for PRO Completed and PRO Not Completed is PRO Expected.

[27] When PRO data are used to inform the evaluation of safety and tolerability, the PRO measure may not be expected to be completed after a patient discontinues from treatment as shown in the table. The PRO Expected column excludes patients who discontinued from treatment and is determined where PRO Expected Flag (PROEXPFL) equals 'Y'.

[28] Summary Statistics are calculated based on PRO Completed.

Figure A8. Change from Baseline for Item 2 with Continuous Response Options (Safety and Tolerability Example)[29] [30]



[29] The 'Improving' and 'Worsening' labels are provided as examples within the figure. Labels of directionality should align with what is provided in the scoring manual for the PRO measure used.
[30] Error bars based on a 95% confidence internal for the mean are represented within the line plot.

5.3.7 *Incidence of Healthcare Utilization*

Table A12. Incidence of Healthcare Utilization (Safety and Tolerability Example where Denominator = PRO Expected)[31]

Analysis Visit	Treatment Arm	Randomized Patients	PRO Expected[32] (N)	Healthcare Utilization Intervention, n (%)					
				Emergency Department (ED) Visits	Hospitalizations	Opiates	Supportive Care Medications (e.g., Steroids, Transfusions, Growth Factors)	Supportive Care Procedures (e.g., Palliative: Hospice, Nephrostomy)	Other (Describe)
Baseline	Control	600	600	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
	Treatment	602	602	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Cycle 2 Day 1	Control	600	564	5 (0.9%)	2 (0.4%)	0 (0.0%)	0 (0.0%)	3 (0.5%)	0 (0.0%)
	Treatment	602	572	5 (0.9%)	3 (0.5%)	0 (0.0%)	1 (0.2%)	1 (0.2%)	0 (0.0%)
Cycle 3 Day 1	Control	600	525	7 (1.3%)	5 (1.0%)	0 (0.0%)	0 (0.0%)	3 (0.6%)	0 (0.0%)
	Treatment	602	542	7 (1.3%)	3 (0.6%)	0 (0.0%)	2 (0.4%)	5 (0.9%)	0 (0.0%)

[31] Denominator used to calculate percentages is PRO Expected.

[32] When PRO data are used to inform the evaluation of safety and tolerability, the PRO measure may not be expected to be completed after a patient discontinues from treatment as shown in the table. The PRO Expected column excludes patients who discontinued from treatment and is determined where PRO Expected Flag (PROEXPFL) equals 'Y'.