U.S. FOOD & DRUG
ADMINISTRATION

Public Workshop

Patient-Focused Drug Development: Workshop #2 to Discuss Methodologic and Other Challenges Related to Patient Experience Data

September 18-19, 2025

Welcome

Ethan Gabbour, MS

Public Health Analyst, Patient-Focused Drug Development Program
Office of Center Director
Center for Drug Evaluation and Research
US Food and Drug Administration

Background

Prescription Drug User Fee Act VII Commitment:

- FDA will issue a Request for Information (RFI) to elicit public input on methodological issues, including the submission and evaluation of patient experience data in the context of the benefit-risk assessment and product labeling, and other areas of greatest interest or concern to public stakeholders.
- FDA will issue a Federal Register Notice summarizing the input to the RFI and based on the input received in response to the RFI, FDA will plan to conduct at least 2 public workshops focused on methodological issues. Based on the RFI and learnings from the workshops, FDA will produce a written summary with identified priorities for future work.

Summary: https://www.regulations.gov/document/FDA-2023-N-1506-0011

Agenda – Day 1

12:30	Welcome & Opening Remarks
12:40	Understanding the Intersection of Meaningful Aspect(s) of Health and Concept of Interest
1:40	Understanding Patient Baseline Severity and How it May Impact a COA’s Ability to Detect Change
2:40	Break
2:55	Challenges and Opportunities for Deriving and Applying Meaningful Score Regions
4:55	Day 1 Wrap-Up

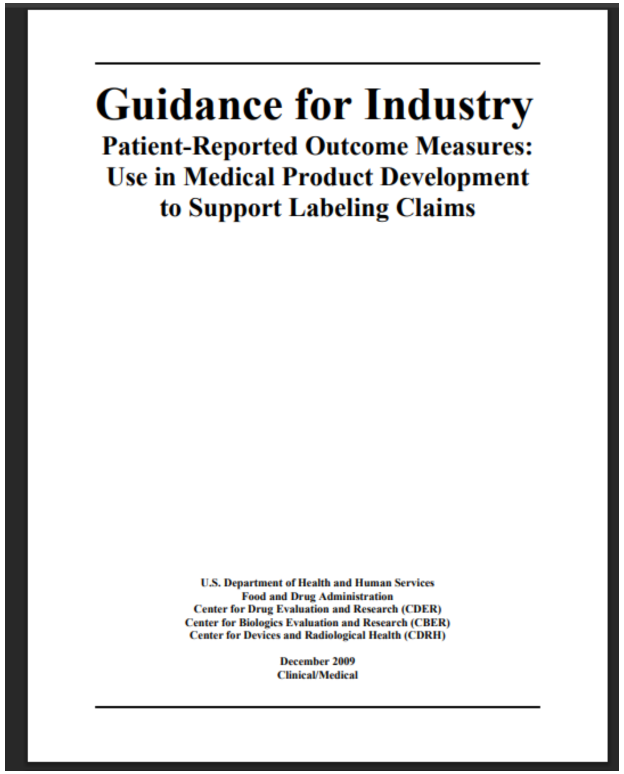
Agenda – Day 2

12:30	Welcome
12:35	Information on a COA Instrument Reviewed by the FDA: Updates to, and Learnings from, the 2009 Outline of the PRO Dossier
2:35	Break
2:50	Clinical Outcome Assessment Data: Interaction of Data Standardization and Statistical Programming
3:50	Closing Remarks

Information on a COA Instrument Reviewed by the FDA: Updates to, and Learnings from, the 2009 Outline of the PRO Dossier

Background

2009 PRO Guidance



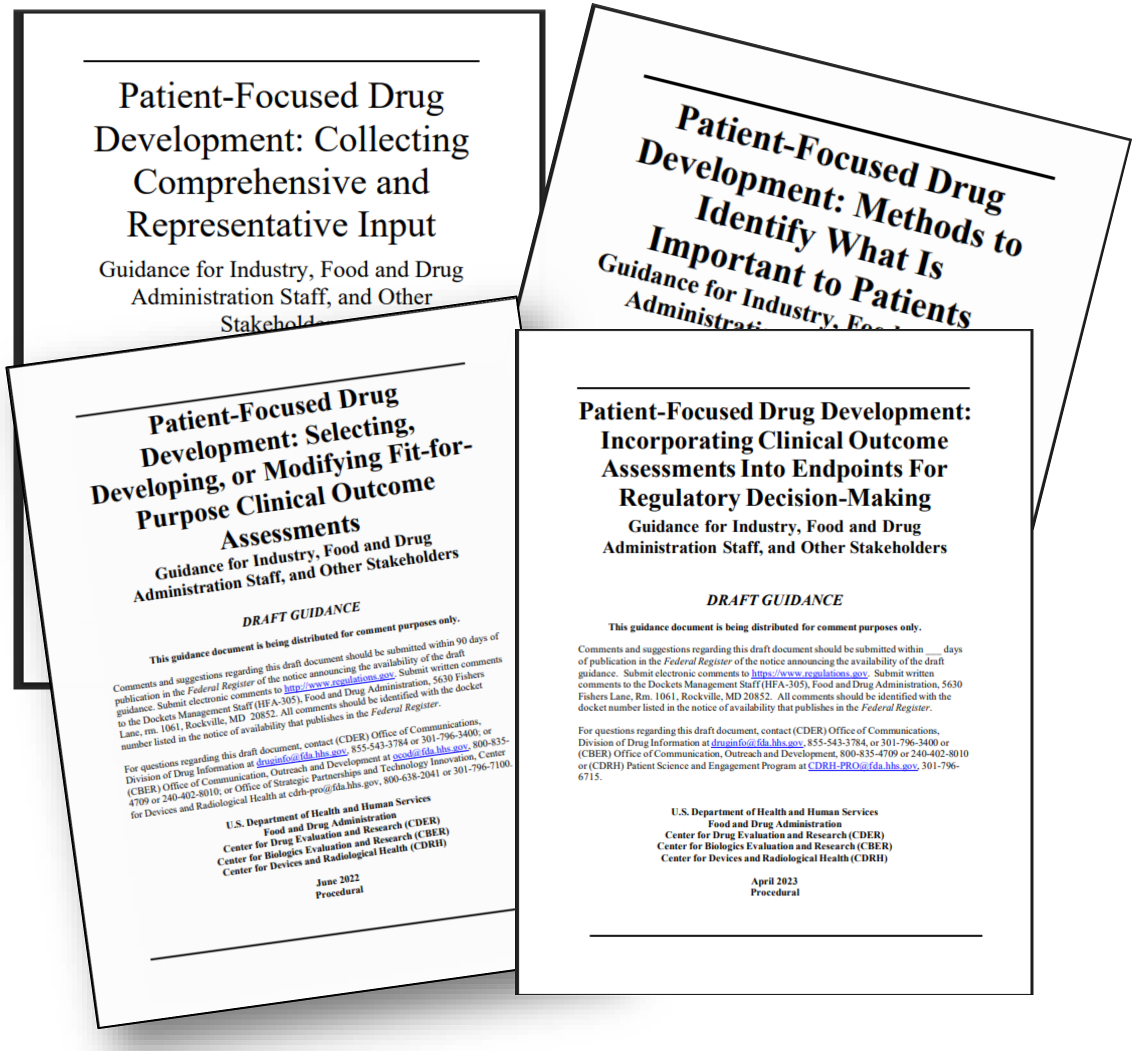
Guidance for Industry

Patient-Reported Outcome Measures:
Use in Medical Product Development
to Support Labeling Claims

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research (CDER)
Center for Biologics Evaluation and Research (CBER)
Center for Devices and Radiological Health (CDRH)

December 2009
Clinical/Medical

PFDD Guidance Series



Patient-Focused Drug
Development: Collecting
Comprehensive and
Representative Input

Guidance for Industry, Food and Drug
Administration Staff, and Other
Stakehold

Patient-Focused Drug
Development: Methods to
Identify What Is
Important to Patients

Guidance for Industry, Fo
Administrati

Patient-Focused Drug
Development: Selecting,
Developing, or Modifying Fit-for-
Purpose Clinical Outcome
Assessments

Guidance for Industry, Food and Drug
Administration Staff, and Other Stakeholders

DRAFT GUIDANCE

This guidance document is being distributed for comment purposes only.

Comments and suggestions regarding this draft document should be submitted within 90 days of publication in the *Federal Register* of the notice announcing the availability of the draft guidance. Submit electronic comments to http://www.regulations.gov. Submit written comments to the Dockets Management Staff (HFA-305), Food and Drug Administration, 5630 Fishers Lane, rm. 1061, Rockville, MD 20852. All comments should be identified with the docket number listed in the notice of availability that publishes in the *Federal Register*.

For questions regarding this draft document, contact (CDER) Office of Communications, Division of Drug Information at druginfo@fda.hhs.gov, 855-543-3784 or 301-796-3400; or (CBER) Office of Communication, Outreach and Development at ocod@fda.hhs.gov, 800-835-4709 or 240-402-8010; or Office of Strategic Partnerships and Technology Innovation, Center for Devices and Radiological Health at cdrh-pro@fda.hhs.gov, 800-638-2041 or 301-796-7100.

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research (CDER)
Center for Biologics Evaluation and Research (CBER)
Center for Devices and Radiological Health (CDRH)

June 2022
Procedural

Patient-Focused Drug Development:
Incorporating Clinical Outcome
Assessments Into Endpoints For
Regulatory Decision-Making

Guidance for Industry, Food and Drug
Administration Staff, and Other Stakeholders

DRAFT GUIDANCE

This guidance document is being distributed for comment purposes only.

Comments and suggestions regarding this draft document should be submitted within ___ days of publication in the *Federal Register* of the notice announcing the availability of the draft guidance. Submit electronic comments to https://www.regulations.gov. Submit written comments to the Dockets Management Staff (HFA-305), Food and Drug Administration, 5630 Fishers Lane, Rm. 1061, Rockville, MD 20852. All comments should be identified with the docket number listed in the notice of availability that publishes in the *Federal Register*.

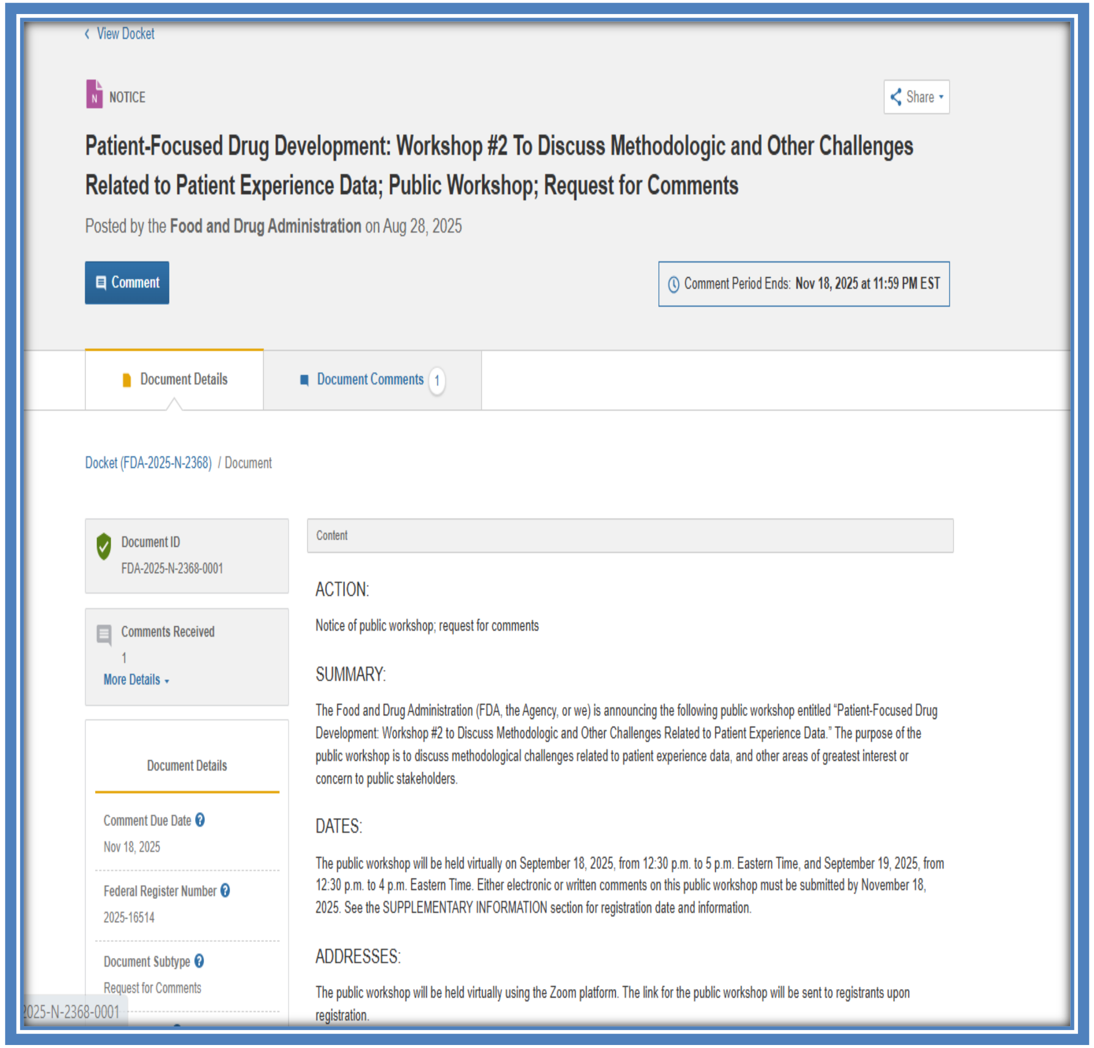
For questions regarding this draft document, contact (CDER) Office of Communications, Division of Drug Information at druginfo@fda.hhs.gov, 855-543-3784, or 301-796-3400 or (CBER) Office of Communication, Outreach and Development, 800-835-4709 or 240-402-8010 or (CDRH) Patient Science and Engagement Program at CDRH-PRO@fda.hhs.gov, 301-796-6715.

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research (CDER)
Center for Biologics Evaluation and Research (CBER)
Center for Devices and Radiological Health (CDRH)

April 2023
Procedural

Background

- Discussion Document:
  - https://www.fda.gov/media/188624/download
- Docket:
  - https://www.regulations.gov/document/FDA-2025-N-2368-0001



Patient-Focused Drug Development: Workshop #2 To Discuss Methodologic and Other Challenges Related to Patient Experience Data; Public Workshop; Request for Comments

Posted by the **Food and Drug Administration** on Aug 28, 2025

Comment Period Ends: **Nov 18, 2025 at 11:59 PM EST**

Docket (FDA-2025-N-2368) / Document

Document ID
FDA-2025-N-2368-0001

Comments Received
1

Document Details

Comment Due Date
Nov 18, 2025

Federal Register Number
2025-16514

Document Subtype
Request for Comments

ACTION:

Notice of public workshop; request for comments

SUMMARY:

The Food and Drug Administration (FDA, the Agency, or we) is announcing the following public workshop entitled "Patient-Focused Drug Development: Workshop #2 to Discuss Methodologic and Other Challenges Related to Patient Experience Data." The purpose of the public workshop is to discuss methodological challenges related to patient experience data, and other areas of greatest interest or concern to public stakeholders.

DATES:

The public workshop will be held virtually on September 18, 2025, from 12:30 p.m. to 5 p.m. Eastern Time, and September 19, 2025, from 12:30 p.m. to 4 p.m. Eastern Time. Either electronic or written comments on this public workshop must be submitted by November 18, 2025. See the SUPPLEMENTARY INFORMATION section for registration date and information.

ADDRESSES:

The public workshop will be held virtually using the Zoom platform. The link for the public workshop will be sent to registrants upon registration.

2025-N-2368-0001

Overview of Session

Presentations (approx. 55 minutes):

Introduction to the Draft Dossier Template:

- Vishal Bhatnagar - Associate Director for Patient Outcomes, Oncology Center of Excellence, US Food and Drug Administration

Walk through Draft Dossier Template:

- Lili Garrard - Master Mathematical Statistician, Division of Biometrics III, Office of Biostatistics, Center for Drug Evaluation and Research, US Food and Drug Administration
- Cynthia Grossman - Division Director, Division of Patient-Centered Development, Center for Devices and Radiological Health, US Food and Drug Administration
- Naomi Knoble - Associate Director, Division of Clinical Outcome Assessment, Center for Drug Evaluation and Research, US Food and Drug Administration
- Weimeng Wang - Staff Fellow, Division of Biometrics III, Office of Biostatistics, Center for Drug Evaluation and Research, US Food and Drug Administration
- Michelle Campbell - Associate Director, Stakeholder Engagement and Clinical Outcomes, Office of Neuroscience, Center for Drug Evaluation and Research, US Food and Drug Administration

Overview of Session (cont.)

Panel Discussion (approx. 45 min):

- Cheryl Coon - Vice President, Clinical Outcome Assessment Program, Critical Path Institute
- Magdalena Harrington - Senior Director, Value & Evidence Patient-Centered Outcomes Assessment Lead, Pfizer
- Ashley Slagle - Clinical Outcome Assessments Scientific and Regulatory Consultant, Aspen Consulting
- Presenters

Audience Questions and Answers (approx. 20 minutes)

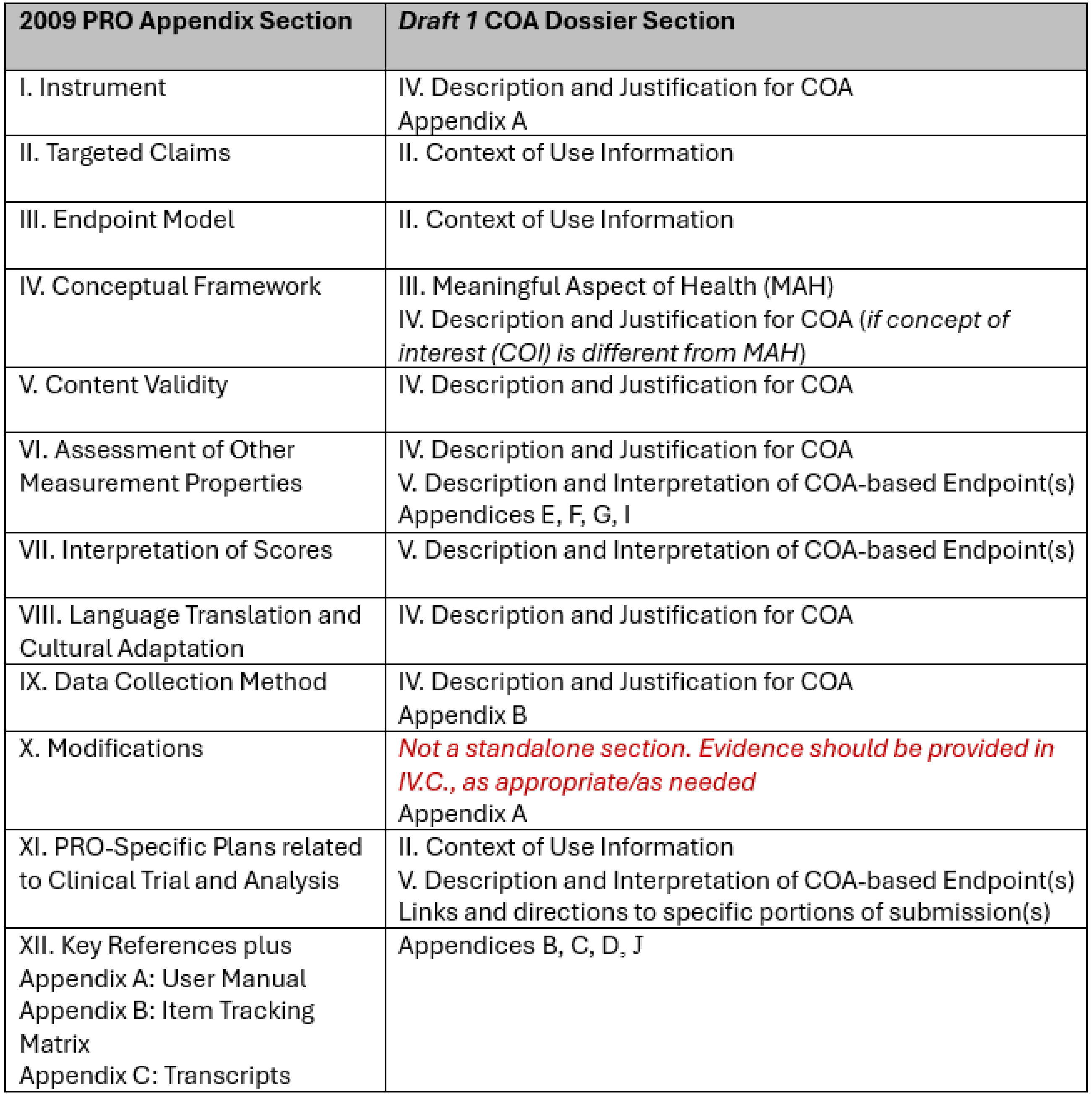
Introduction

- **Purpose**: Support use and interpretation of proposed COA scores and COA-based endpoints
- **Submission Options**: Single comprehensive dossier OR separate dossier per COA
- **Template Compliance:** All sections should be included (use "NA" with explanation if not applicable)
- **Technical Standards**:
  - Active hyperlinks required
  - Full-text searchable PDF
  - Complete table of contents, list of tables, and list of figures
- **Submission Location:** Module 5.3.5.3 of eCTD with reviewer's guide and cross-references
- **Supporting Materials**: Full articles for all cited references in Module 5.4

2009 PRO Appendix to *Draft 1* Dossier Template

- 2009 PRO Guidance Appendix
  - Outline of information on a PRO instrument that should be provided to FDA for review
- *Draft 1* Dossier Template
  - Based on similar information but reorganized to align with PFDD Guidance 3 rationale-based framework
  - Applicable to all four types of COAs
  - Demonstrates that a single source of evidence could support multiple components of a rationale
  - Provides an opportunity to be transparent with information used to support more efficient regulatory review and to better inform regulatory decision-making

Table mapping the 2009 Appendix to the current draft documents



2009 PRO Appendix Section	*Draft 1* COA Dossier Section
I. Instrument	IV. Description and Justification for COA Appendix A
II. Targeted Claims	II. Context of Use Information
III. Endpoint Model	II. Context of Use Information
IV. Conceptual Framework	III. Meaningful Aspect of Health (MAH) IV. Description and Justification for COA (*if concept of interest (COI) is different from MAH*)
V. Content Validity	IV. Description and Justification for COA
VI. Assessment of Other Measurement Properties	IV. Description and Justification for COA V. Description and Interpretation of COA-based Endpoint(s) Appendices E, F, G, I
VII. Interpretation of Scores	V. Description and Interpretation of COA-based Endpoint(s)
VIII. Language Translation and Cultural Adaptation	IV. Description and Justification for COA
IX. Data Collection Method	IV. Description and Justification for COA Appendix B
X. Modifications	*Not a standalone section. Evidence should be provided in IV.C., as appropriate/as needed* Appendix A
XI. PRO-Specific Plans related to Clinical Trial and Analysis	II. Context of Use Information V. Description and Interpretation of COA-based Endpoint(s) Links and directions to specific portions of submission(s)
XII. Key References plus Appendix A: User Manual Appendix B: Item Tracking Matrix Appendix C: Transcripts	Appendices B, C, D, J

Section I. Executive Summary

Section I.A Summary of COA-Based Endpoint Approach

- High-level summary
  - Intended context of use (COU)
  - Description and justification for meaningful aspect(s) of health (MAH)
  - Concept of interest (COI), if not identical to the MAH

Table 1. Summary of COAs by Context of Use

		COAs				
MAH	**COI for Measurement**	Type	Name	Score	**COA-based Endpoint**	**COA-based Endpoint positioning**
Example: Lower Limb-Related Functioning	Walking Capacity	PerfO	Six Minute Walk Test	Six Minute Walk Distance (Meters)	Six Minute Walk Distance (Meters) at Week 24	Multiplicity-controlled Key Secondary

Section I.B Summary of the Rationale to Support the Fitness-for-Purpose of Each COA

- Rationale explaining how and why the COA is intended to work
- **Summary table of components that support the rationale**
  - Rationale and components should be tailored to the COI in a specific COU
  - High-level summary, including statistical results if applicable
  - **Supporting evidence should be linked to the corresponding dossier section where full justification and relevant results can be found**

Section I.B Summary of the Rationale to Support the Fitness-for-Purpose of Each COA

- Rationale explaining how and why the COA is intended to work
- Summary table of components that support the rationale
  - Rationale and components should be tailored to the COI in a specific COU
  - High-level summary, including statistical results if applicable
  - Supporting evidence should be linked to the corresponding dossier section where full justification and relevant results can be found
- Example table provided in *draft* dossier template; Sponsors should **include language and specific components relevant to their development program**
  - Not all example components may be needed
  - Additional components may be needed

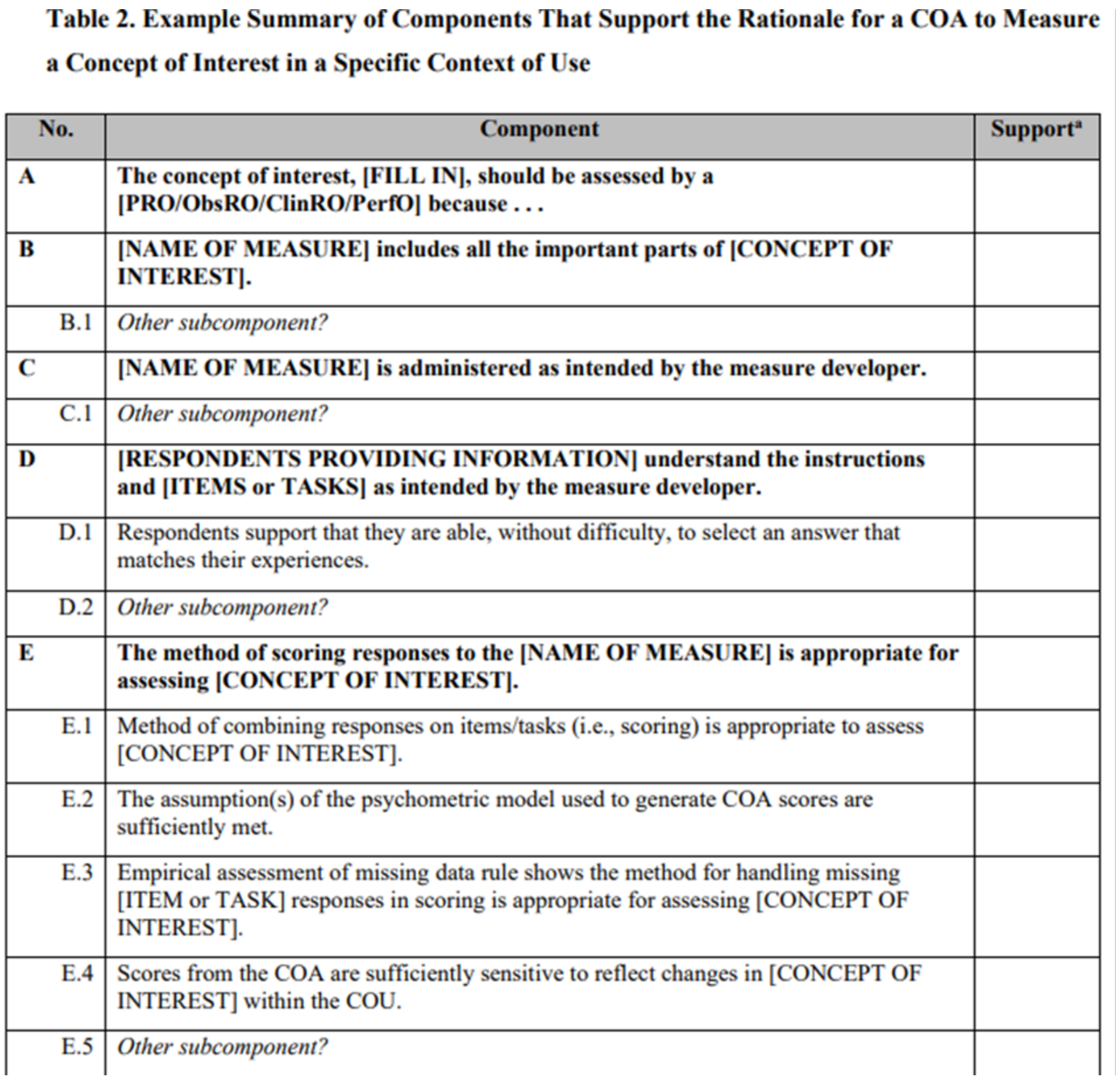
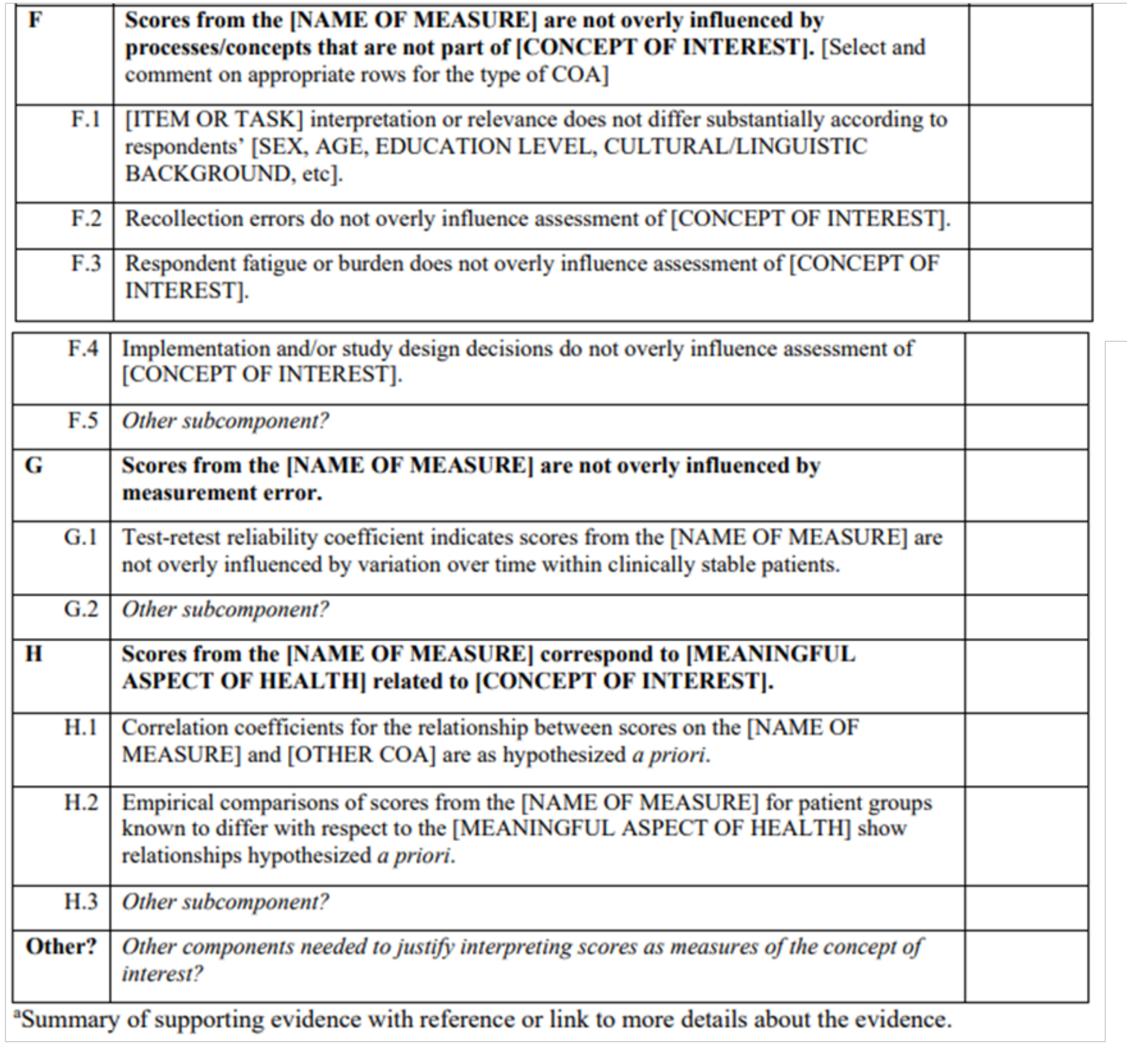


Table 2. Example Summary of Components That Support the Rationale for a COA to Measure a Concept of Interest in a Specific Context of Use

No.	Component	Support[a]
A	**The concept of interest, [FILL IN], should be assessed by a [PRO/ObsRO/ClinRO/PerfO] because . . .**	
B	**[NAME OF MEASURE] includes all the important parts of [CONCEPT OF INTEREST].**	
B.1	*Other subcomponent?*	
C	**[NAME OF MEASURE] is administered as intended by the measure developer.**	
C.1	*Other subcomponent?*	
D	**[RESPONDENTS PROVIDING INFORMATION] understand the instructions and [ITEMS or TASKS] as intended by the measure developer.**	
D.1	Respondents support that they are able, without difficulty, to select an answer that matches their experiences.	
D.2	*Other subcomponent?*	
E	**The method of scoring responses to the [NAME OF MEASURE] is appropriate for assessing [CONCEPT OF INTEREST].**	
E.1	Method of combining responses on items/tasks (i.e., scoring) is appropriate to assess [CONCEPT OF INTEREST].	
E.2	The assumption(s) of the psychometric model used to generate COA scores are sufficiently met.	
E.3	Empirical assessment of missing data rule shows the method for handling missing [ITEM or TASK] responses in scoring is appropriate for assessing [CONCEPT OF INTEREST].	
E.4	Scores from the COA are sufficiently sensitive to reflect changes in [CONCEPT OF INTEREST] within the COU.	
E.5	*Other subcomponent?*	



F	**Scores from the [NAME OF MEASURE] are not overly influenced by processes/concepts that are not part of [CONCEPT OF INTEREST].** [Select and comment on appropriate rows for the type of COA]	
F.1	[ITEM OR TASK] interpretation or relevance does not differ substantially according to respondents' [SEX, AGE, EDUCATION LEVEL, CULTURAL/LINGUISTIC BACKGROUND, etc].	
F.2	Recollection errors do not overly influence assessment of [CONCEPT OF INTEREST].	
F.3	Respondent fatigue or burden does not overly influence assessment of [CONCEPT OF INTEREST].	
F.4	Implementation and/or study design decisions do not overly influence assessment of [CONCEPT OF INTEREST].	
F.5	*Other subcomponent?*	
G	**Scores from the [NAME OF MEASURE] are not overly influenced by measurement error.**	
G.1	Test-retest reliability coefficient indicates scores from the [NAME OF MEASURE] are not overly influenced by variation over time within clinically stable patients.	
G.2	*Other subcomponent?*	
H	**Scores from the [NAME OF MEASURE] correspond to [MEANINGFUL ASPECT OF HEALTH] related to [CONCEPT OF INTEREST].**	
H.1	Correlation coefficients for the relationship between scores on the [NAME OF MEASURE] and [OTHER COA] are as hypothesized *a priori*.	
H.2	Empirical comparisons of scores from the [NAME OF MEASURE] for patient groups known to differ with respect to the [MEANINGFUL ASPECT OF HEALTH] show relationships hypothesized *a priori*.	
H.3	*Other subcomponent?*	
Other?	*Other components needed to justify interpreting scores as measures of the concept of interest?*	

[a] Summary of supporting evidence with reference or link to more details about the evidence.

Section I.C Summary of Qualitative and Quantitative Evidence to Support the Evaluation of the Meaningfulness of COA-Based Endpoint Scores

- **High-level summary of efficacy analyses results of COA-based endpoint(s)**
  - Include links to clinical study report(s) (CSR)
- **Overall conclusion of the clinical meaningfulness** of COA-based endpoint results, as appropriate
  - Provide final range of meaningful score differences (MSDs) or meaningful score regions (MSRs)
  - Provide summary on how MSDs or MSRs were applied to trial results to interpret meaningfulness
  - Provide summary of qualitative evidence, if available
  - Provide links to the corresponding dossier section where full justification and relevant results can be found

Section I.C Summary of Qualitative and Quantitative Evidence to Support the Evaluation of the Meaningfulness of COA-Based Endpoint Scores

- **Section 1.C should only be completed after trial results are available.** For example:
  - Submitting an evidence dossier as part of end-of-phase 2 (EOP2) meeting where blinded data from phase 2 were used to determine MSDs or MSRs then applied to unblinded phase 2 trial results to explore meaningfulness
  - Submitting an evidence dossier as part of NDA/BLA submission where blinded phase 3 data were used to determine MSDs or MSRs then applied to unblinded trial results to evaluate meaningfulness

Section II. Context of Use Information

Section II.A-F: Context of Use

- **Cross-walk to 2009 PRO Appendix**
  - Targeted Claims and Endpoint Model

- **Provides information on -**
  A. Clinical Trial Population
  B. Clinical Trial Design
  C. Clinical Trial Objective/Hypothesis Corresponding to COA
  D. Endpoint Definition(s) and Positioning
  E. Desired Claim Related to COA-Based Endpoint

Section III. Meaningful Aspect of Health

Section III: Meaningful Aspect of Health

- **Cross-walk to 2009 PRO Appendix**
  - Conceptual Framework

- **Meaningful Aspect of Health (MAH) comprises the following -**
  1. Conceptual disease model of patients' experience of living with the condition
  2. Definition/Conceptual framework for the MAH
  3. Evidence of importance to patients and/or caregivers
  4. Explanation of potential to show effect of intervention on MAH within the duration of the trial.

Section IV. Description and Justification for COA

Sections IV.A-B: Concept of Interest and Detailed COA Description

- Measured concept of interest (COI)
  - A description of the COI's relationship to the MAH
  - If the COI is different from the MAH, a definition/conceptual framework for COI
- Description of the COA
  - Exact version
  - Type of COA (e.g., PRO, ObsRO)
  - Details such as the recall period, mode of administration, scoring, etc.

Section IV.C Evidence-Based Rationale for Interpreting Scores from the COA as Reflecting the COI within the COU

Section IV.C: A. Concept of interest
Section IV.C: B. COA includes all the important parts of the COI

A. The concept of interest, [FILL IN], should be assessed by a [PRO/ObsRO/ClinRO/PerfO] because ...

This component may not always be needed, but it would be expected if multiple COA types are used.

B. [NAME OF MEASURE] includes all the important parts of [CONCEPT OF INTEREST]

The Sponsor should submit a summary of evidence that the COA captures all of the most clinically important concepts (e.g., literature review, expert input, qualitative research, summary of chronology from item generation to finalization)

Section IV.C: C. Measure administered as intended

C. [NAME OF MEASURE] is administered as intended by the measure developer.

The Sponsor should submit:

- User manuals
- Standardized training program completed by personnel at all sites
- Intermittent observation to ensure protocol adherence
- Rater training (if raters are used)

Section IV.C: D. Respondents Understand the Instructions and Items/Tasks of the Measure as Intended by the Measure Developer

D. [RESPONDENTS PROVIDING INFORMATION] understand the instructions and [ITEMS or TASKS] as intended by the measure developer.

The Sponsor should submit:

- For PROs, ObsROs, and ClinROs, evidence that instructions and responses are understood as intended (e.g., cognitive interviews)
- For PerfOs, evidence that respondents understand task instructions (e.g., cognitive interviews) and the task performance requested (e.g., pilot testing results)

Section IV.C: E. The method of scoring responses to the COA is appropriate for assessing the concept of interest

E1. The method of combining responses on items/tasks (i.e., scoring) is appropriate to assess [CONCEPT OF INTEREST]

The Sponsor should submit:

- The type of measurement model
- Measurement framework
- How responses to multiple items are combined
- Both quantitative evidence and qualitative evidence
- Justification any score transformations and normalizations, if applicable
- Scoring manual (if applicable)

Section IV.C: E. The method of scoring responses to the COA is appropriate for assessing the concept of interest

E1. The method of combining responses on items/tasks (i.e., scoring) is appropriate to assess [CONCEPT OF INTEREST]

E2. The assumption(s) of the psychometric model (e.g., partial-credit model, graded response model) used to generate COA scores are sufficiently met

Section IV.C: E. The method of scoring responses to the COA is appropriate for assessing the concept of interest

E1. The method of combining responses on items/tasks (i.e., scoring) is appropriate to assess [CONCEPT OF INTEREST]

E2. The assumption(s) of the psychometric model (e.g., partial-credit model, graded response model) used to generate COA scores are sufficiently met

E3. Empirical assessment of missing data rule shows the method for handling missing [ITEM or TASK] responses in scoring is appropriate for assessing [CONCEPT OF INTEREST]

- Missing data rules
- Justification for missing data rules
- Additional considerations for missing item or task responses
- Link to relevant section in scoring manual

E4. Scores from the COA are sufficiently sensitive to reflect changes in [CONCEPT OF INTEREST] within the COU

Section IV.C: F. Scores from the COA are not overly influenced by processes/concepts that are not part of the concept of interest

Sponsor should specify interfering influences on responses to items/tasks and assess the presence and strength of those influences that may influence score interpretation.

F1. Item or Task Interpretation or Relevance Does Not Differ Substantially According to Respondents’ Demographic Characteristics (Including Sex, Age, and Education Level) or Cultural/Linguistic Backgrounds

Section IV.C: F. Scores from the COA are not overly influenced by processes/concepts that are not part of the concept of interest

Sponsor should specify interfering influences on responses to items/tasks and assess the presence and strength of those influences that may influence score interpretation.

F1. Item or Task Interpretation or Relevance Does Not Differ Substantially According to Respondents' Demographic Characteristics (Including Sex, Age, and Education Level) or Cultural/Linguistic Backgrounds

F2. Recollection Errors Do Not Overly Influence Assessment of the Concept of Interest. (PROs, ObsROs, and ClinROs)

Section IV.C: F. Scores from the COA are not overly influenced by processes/concepts that are not part of the concept of interest

Sponsor should specify interfering influences on responses to items/tasks and assess the presence and strength of those influences that may influence score interpretation.

F1. Item or Task Interpretation or Relevance Does Not Differ Substantially According to Respondents' Demographic Characteristics (Including Sex, Age, and Education Level) or Cultural/Linguistic Backgrounds

F2. Recollection Errors Do Not Overly Influence Assessment of the Concept of Interest. (PROs, ObsROs, and ClinROs)

F3. Respondent fatigue or burden does not overly influence assessment of [CONCEPT OF INTEREST]

Section IV.C: F. Scores from the COA are not overly influenced by processes/concepts that are not part of the concept of interest

Sponsor should specify interfering influences on responses to items/tasks and assess the presence and strength of those influences that may influence score interpretation.

F1. Item or Task Interpretation or Relevance Does Not Differ Substantially According to Respondents' Demographic Characteristics (Including Sex, Age, and Education Level) or Cultural/Linguistic Backgrounds

F2. Recollection Errors Do Not Overly Influence Assessment of the Concept of Interest. (PROs, ObsROs, and ClinROs)

F3. Respondent fatigue or burden does not overly influence assessment of [CONCEPT OF INTEREST]

F4. Implementation and/or study design decisions do not overly influence assessment of [CONCEPT OF INTEREST]

Section IV.C: G. Scores from the COA are not overly influenced by measurement error

Sponsor should evaluate the most likely sources of variability in scores within the context of use and discuss quantitative study(ies) conducted to support that the COA scores are not overly influenced by measurement error

G1. Test-retest reliability coefficient indicates scores from the COA are not overly influenced by variation over time within clinically stable patients.

G2. Inter/Intra-rater reliability coefficient indicates scores from the COA are not overly influenced by variation over different COA raters (e.g., clinicians) or within the same COA rater over time.

- Summarize the study results in this section to include comparability of raters as well as patient characteristics between reliability study and pivotal trial(s)
- Submit rater qualifications/credentials and training materials in the Appendix
- If a standalone reliability study was conducted, include study protocol, which should include patient inclusion/exclusion criteria, and study results in the Appendix
- Describe methods used in detail

Section IV.C: H. Scores from the COA correspond to the meaningful aspect of health related to the concept of interest

H1. Correlation coefficients for the relationship between scores on the COA and other supportive COAs are as hypothesized a priori (e.g., convergent evidence)

- Specify which reference measures will be used to support convergent and divergent evidence
- Include the a priori hypothesized relationships among the concepts measured by any proposed reference measures and the concept(s) measured by the proposed COA
- Specify the criteria for describing the strength of a correlation

H2. Empirical comparisons of scores form the COA for patient groups known to differ with respect to the MAH show relationships hypothesized a priori (e.g., known-groups evidence)

- Specify appropriate reference measures and justify the corresponding cutoff values that represent clinically distinct levels of symptom severity and/or impairment
- Provide details of the proposed model and the hypothesis tests that will be performed

Section V. Description and Justification of COA-Based Endpoint(s)

Section V.A Computation of Proposed COA-Based Endpoints

1. Detailed algorithm, including the missing data rule(s) for the endpoints

   - Specify where the endpoint algorithm is described in the SAP and link the final SAP
   - Evidence to support the endpoint algorithm, including each missing data rule

Section V.B Rationale to Support Endpoint Selection

1. The MAH that the endpoint is thought to reflect

2. The clinical trial objective or hypothesis corresponding to the endpoint, ensuring the objective or hypothesis is specific
    – E.g., To assess the efficacy of product X by comparing the patient-reported physical functioning between arms at 24 weeks

3. The role of the endpoint

    - E.g., Primary, secondary, other

Section V.B Rationale to Support Endpoint Selection

4. How the COA-based endpoint will assess clinical benefit for the intended indication

5. Explain why the selected COA is fit-for-purpose

6. Evidence that the endpoint provides sufficient precision to estimate the effect of treatment when used with the proposed study design, sample size and analysis plan

Section V.B Rationale to Support Endpoint Selection

7. Support the importance of the endpoint to patients and/or caregivers from the literature and/or primary data collection

8. Strengths and limitations of the proposed endpoint

Section V.C Evaluating the Meaningfulness of COA-Based Endpoint Scores

1. Description and justification for the methodologies used to evaluate the meaningfulness of COA-based endpoint scores
    - Methodologies used to derive meaningful score differences (MSDs) or meaningful score regions (MSRs)
    - Methodologies used to apply derived MSDs and MSRs (e.g., average horizontal gap, vertical gap)
2. For the chosen methodology, provide descriptions of the analyses conducted and discussion of how the analyses informed the final MSDs or MSRs (e.g., **anchor-based analyses, qualitative interviews**, bookmarking, existing literature, other approaches)

Section V.C Evaluating the Meaningfulness of COA-Based Endpoint Scores

- When **anchor-based analyses** are used, include:
  - Evidence to support the use of the anchors (e.g., appropriateness of the response options of each anchor, that the concept assessed by each anchor variable is sufficiently related to the concept of interest of the COA)
  - Designation of each anchor measure as primary or secondary anchor that is supported by the hypothesized relationship of the anchor to the COI of the COA
  - Figures and tables to support the anchor-based analyses (e.g., empirical cumulative distribution function (eCDF) curves)

Section V.C Evaluating the Meaningfulness of COA-Based Endpoint Scores

- When **qualitative interviews and/or surveys** are used
  - Provide description of how qualitative data are used to inform selection of MSDs or MSRs
  - If interviews occurred during concept elicitation, if there were cognitive debriefing or exit interviews, or if exit surveys were conducted, include study documents (e.g., protocol, interviewer guide, transcripts)
  - Specify when interviews were conducted, e.g., as a stand-alone study, prior to start of the phase 3 trial, after completing the main portion of the study before unblinding

Section V.C Evaluating the Meaningfulness of COA-Based Endpoint Scores

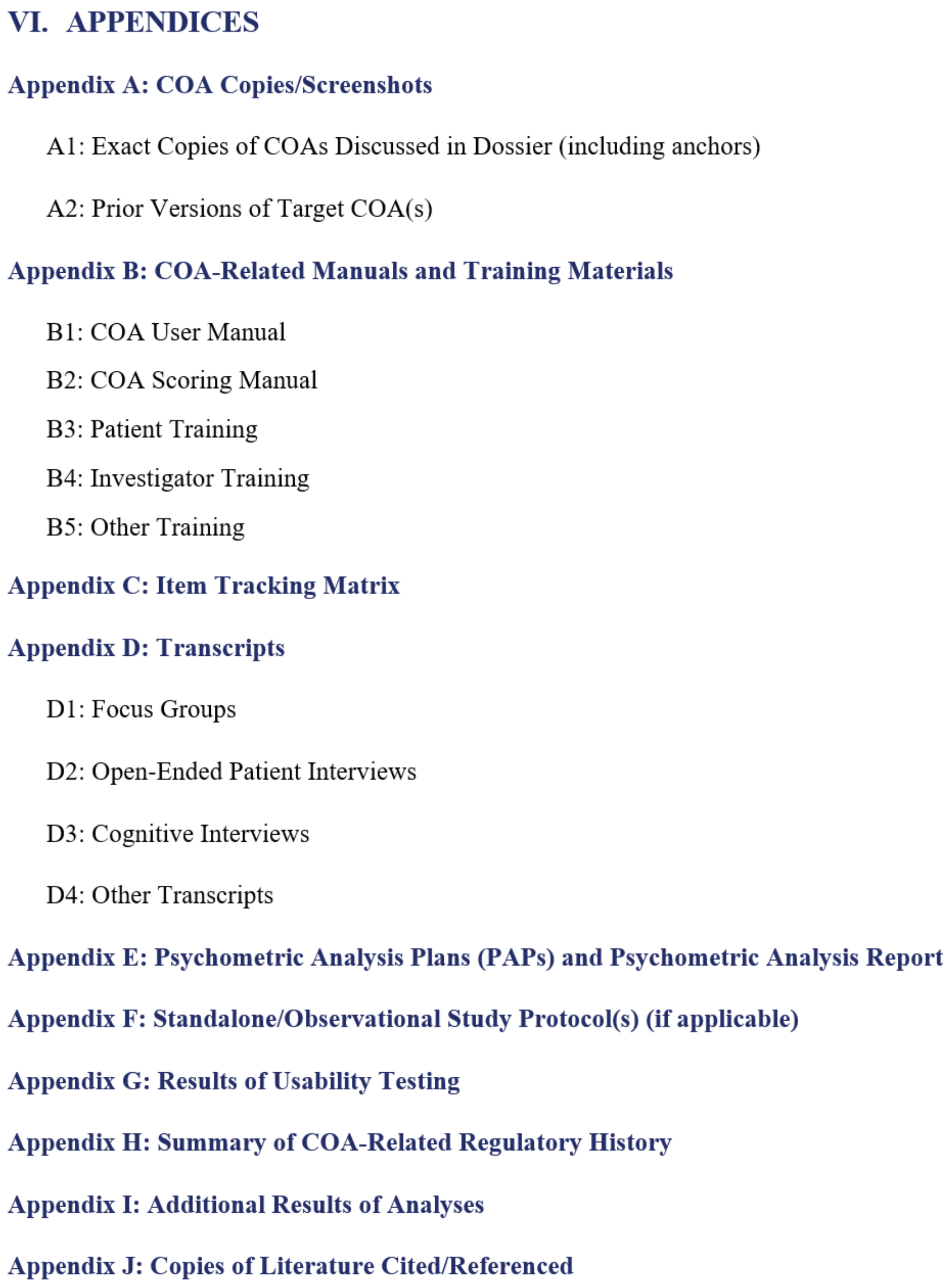
- Approach used to finalize MSD or MSR range
  - Final MSD range or MSRs
  - Discussion of rates of patient misclassification and impact of patient baseline symptom severity

Section V.C Evaluating the Meaningfulness of COA-Based Endpoint Scores

3. Details and discussion of how derived final range of MSDs or MSRs were applied to evaluate the meaningfulness of the COA endpoint scores in clinical trial(s)
    - Description of method(s) (e.g., average horizontal gap, vertical gap)
    - eCDF curves of COA endpoint scores by treatment arm annotated to aid interpretation (e.g., sample size, median, reference line(s) for proposed range of thresholds)
4. Overall conclusion regarding meaningfulness of COA-based endpoint scores in the context of the trial(s)

Should only be completed after trial results are available

Section VI



VI. APPENDICES

Appendix A: COA Copies/Screenshots

A1: Exact Copies of COAs Discussed in Dossier (including anchors)

A2: Prior Versions of Target COA(s)

Appendix B: COA-Related Manuals and Training Materials

B1: COA User Manual
B2: COA Scoring Manual
B3: Patient Training
B4: Investigator Training
B5: Other Training

Appendix C: Item Tracking Matrix

Appendix D: Transcripts

D1: Focus Groups

D2: Open-Ended Patient Interviews

D3: Cognitive Interviews

D4: Other Transcripts

Appendix E: Psychometric Analysis Plans (PAPs) and Psychometric Analysis Report

Appendix F: Standalone/Observational Study Protocol(s) (if applicable)

Appendix G: Results of Usability Testing

Appendix H: Summary of COA-Related Regulatory History

Appendix I: Additional Results of Analyses

Appendix J: Copies of Literature Cited/Referenced

Panel Discussion

followed by

Q+A

Break

Please return at 2:50 p.m. ET

Clinical Outcome Assessment Data: Interaction of Data Standardization and Statistical Programming

Overview of Session

Presentations (approx. 45 minutes)

Background

- Laura Lee Johnson - Division Director, Division of Biometrics III, Office of Biostatistics, Center for Drug Evaluation and Research, US Food and Drug Administration

NCI Case Study: Cancer Treatment Tolerability Consortium Resources and Tools

- Amylou Dueck - Professor of Biostatistics, Vice Chair, Department of Quantitative Health Sciences, Mayo Clinic

Learnings from Oncology: Submitting Patient-Reported Outcome Data from Cancer Clinical Trials

- Flora Mulkey - Lead Mathematical Statistician, Division of Biometrics V, Office of Biostatistics, Center for Drug Evaluation and Research

Use of Available Code Packages and Resources

- Ki Chung - Statistician, Division of Analytics and Informatics, Office of Biostatistics, Center for Drug Evaluation and Research

Panel Discussion (approx. 15 minutes)

- Pandu Kulkarni - Chief Executive Officer of Aparito, Chief Analytics Officer of Research and Development, Senior Vice President of Statistics, Data, and Analytics, Eli Lilly and Company
- Michael Otterstatter - Senior Principal Biostatistician, Emmes Group

Terms

- CDISC: Clinical Data Interchange Standards Consortium
- TCG: FDA Study Data Technical Conformance Guide
- FDA Study Data Standards Resources
- "Tech Spec": FDA Guidance for Industry Technical Specifications Document
- CTCAE: Common Terminology Criteria for Adverse Events
- PRO-CTCAE: is derived from the CTCAE version 4.0

Data Standards and Study Data Submission

- Data standards enable FDA to modernize and streamline the review process
- Standard way to exchange clinical and nonclinical research data between computer systems
- Consistent general framework for organizing study data, including templates for datasets, standard names for variables, and standard ways of doing calculations with common variables

➢ Not only for patient experience data

➢ CDER and CBER

https://www.fda.gov/industry/study-data-standards-resources/study-data-submission-cder-and-cber

EXAMPLE: TOLERABILITY OF ANTI-CANCER AGENTS

Tolerability Consortium:

Assessing Tolerability of Anti-Cancer Agents Using Clinician- and Patient-Reported Outcomes: Methods for Analyzing and Interpreting CTCAE and PRO-CTCAE Data

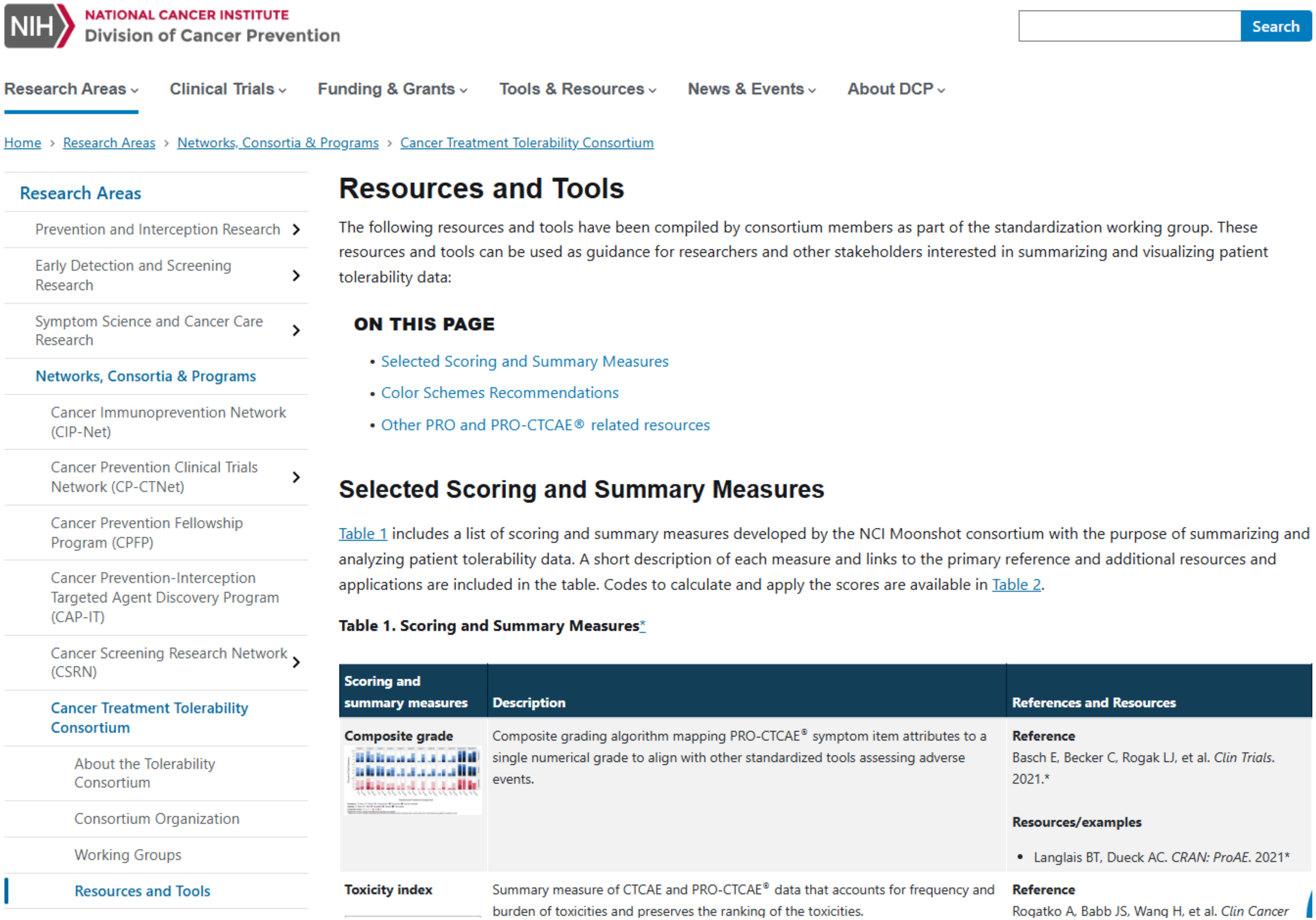
Slide from Lori Minasian, NCI

Funded Teams for the Tolerability Consortium

- UCLA-Cedar Sinai Team (with NRG investigators)
  - Integrate patient baseline risk factors into tolerability assessment
- Mayo Clinic-University of North Carolina Team (with Alliance investigators)
  - Analyzing CTCAE and PRO-CTCAE to characterize drug tolerability
- EVOLV Team (ECOG-ACRIN investigators)
  - Exploring the FACT-GP5 as a predictor for early discontinuation
- University Rochester Team
  - Understand relationships of aging and tolerability in older patients

Slide from Lori Minasian, NCI

https://prevention.cancer.gov/research-areas/networks-consortia-programs/cttc/resources-and-tools



Resources and Tools

The following resources and tools have been compiled by consortium members as part of the standardization working group. These resources and tools can be used as guidance for researchers and other stakeholders interested in summarizing and visualizing patient tolerability data:

ON THIS PAGE

- Selected Scoring and Summary Measures
- Color Schemes Recommendations
- Other PRO and PRO-CTCAE® related resources

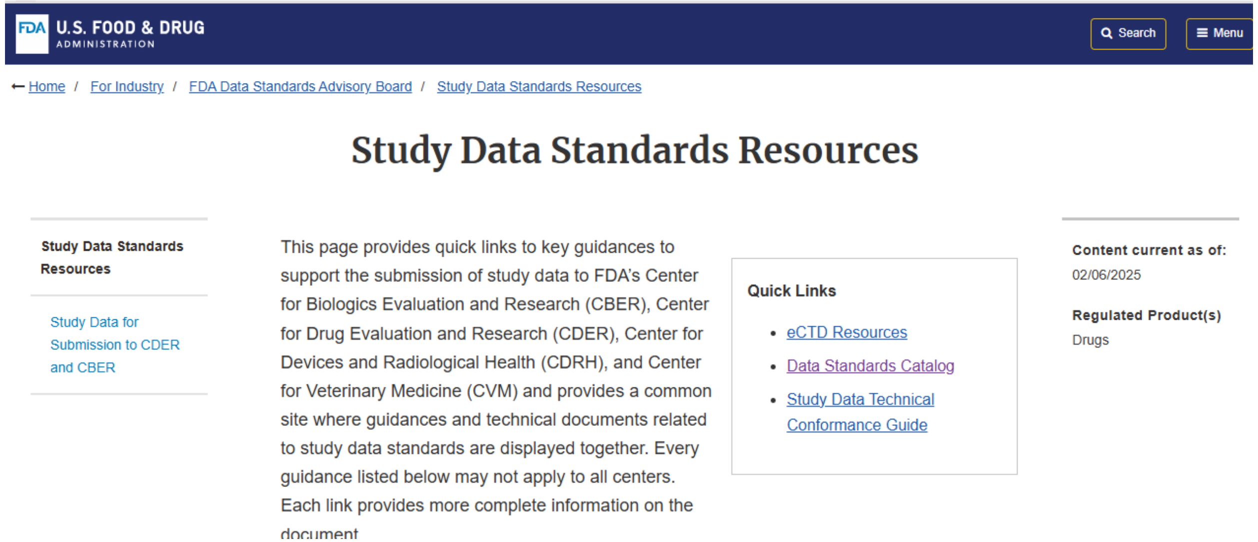
Selected Scoring and Summary Measures

Table 1 includes a list of scoring and summary measures developed by the NCI Moonshot consortium with the purpose of summarizing and analyzing patient tolerability data. A short description of each measure and links to the primary reference and additional resources and applications are included in the table. Codes to calculate and apply the scores are available in Table 2.

Table 1. Scoring and Summary Measures*

Scoring and summary measures	Description	References and Resources
Composite grade	Composite grading algorithm mapping PRO-CTCAE® symptom item attributes to a single numerical grade to align with other standardized tools assessing adverse events.	**Reference** Basch E, Becker C, Rogak LJ, et al. *Clin Trials*. 2021.* **Resources/examples** • Langlais BT, Dueck AC. *CRAN: ProAE*. 2021*
Toxicity index	Summary measure of CTCAE and PRO-CTCAE® data that accounts for frequency and burden of toxicities and preserves the ranking of the toxicities.	**Reference** Rogatko A, Babb JS, Wang H, et al. *Clin Cancer*

FDA: Oncology PRO Tech Spec



← Home / For Industry / FDA Data Standards Advisory Board / Study Data Standards Resources

Study Data Standards Resources

Study Data Standards Resources

Study Data for Submission to CDER and CBER

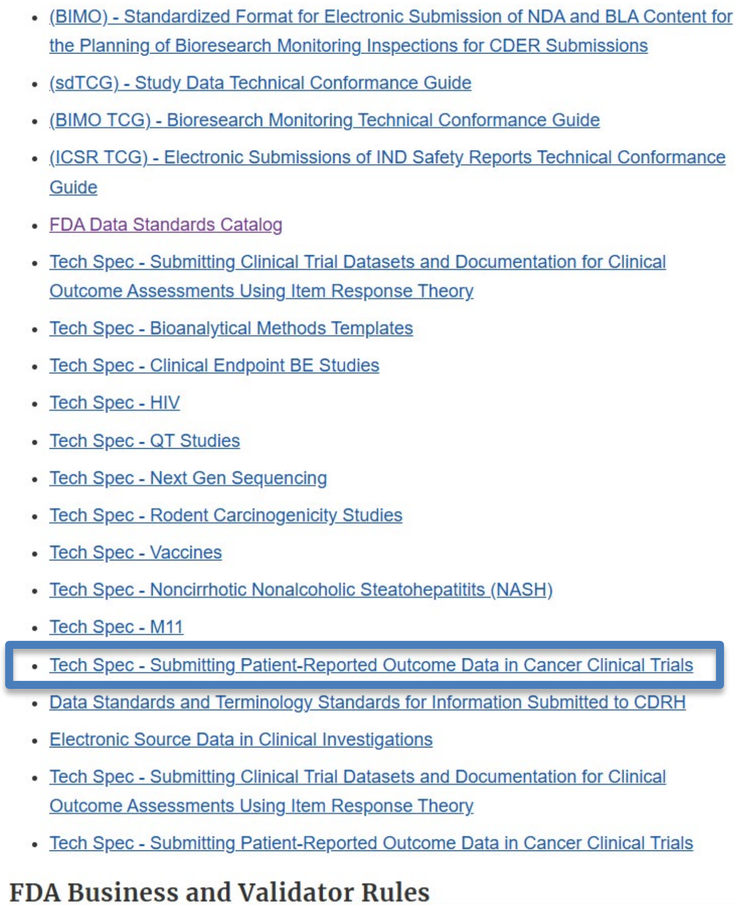
This page provides quick links to key guidances to support the submission of study data to FDA's Center for Biologics Evaluation and Research (CBER), Center for Drug Evaluation and Research (CDER), Center for Devices and Radiological Health (CDRH), and Center for Veterinary Medicine (CVM) and provides a common site where guidances and technical documents related to study data standards are displayed together. Every guidance listed below may not apply to all centers. Each link provides more complete information on the document

Quick Links

- eCTD Resources
- Data Standards Catalog
- Study Data Technical Conformance Guide

Content current as of:
02/06/2025

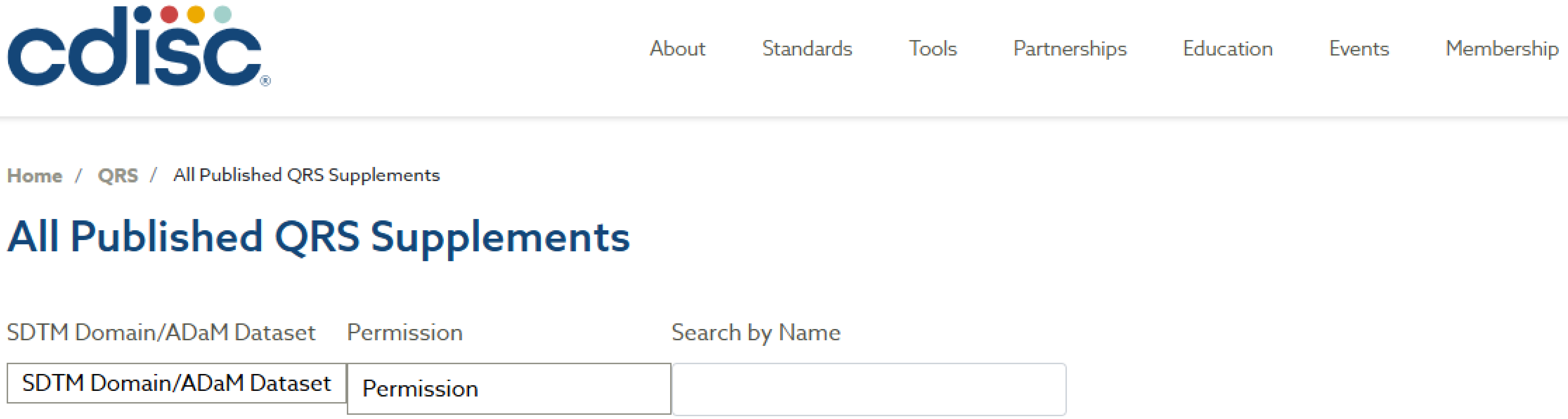
Regulated Product(s)
Drugs



- (BIMO) - Standardized Format for Electronic Submission of NDA and BLA Content for the Planning of Bioresearch Monitoring Inspections for CDER Submissions
- (sdTCG) - Study Data Technical Conformance Guide
- (BIMO TCG) - Bioresearch Monitoring Technical Conformance Guide
- (ICSR TCG) - Electronic Submissions of IND Safety Reports Technical Conformance Guide
- FDA Data Standards Catalog
- Tech Spec - Submitting Clinical Trial Datasets and Documentation for Clinical Outcome Assessments Using Item Response Theory
- Tech Spec - Bioanalytical Methods Templates
- Tech Spec - Clinical Endpoint BE Studies
- Tech Spec - HIV
- Tech Spec - QT Studies
- Tech Spec - Next Gen Sequencing
- Tech Spec - Rodent Carcinogenicity Studies
- Tech Spec - Vaccines
- Tech Spec - Noncirrhotic Nonalcoholic Steatohepatitis (NASH)
- Tech Spec - M11
- Tech Spec - Submitting Patient-Reported Outcome Data in Cancer Clinical Trials
- Data Standards and Terminology Standards for Information Submitted to CDRH
- Electronic Source Data in Clinical Investigations
- Tech Spec - Submitting Clinical Trial Datasets and Documentation for Clinical Outcome Assessments Using Item Response Theory
- Tech Spec - Submitting Patient-Reported Outcome Data in Cancer Clinical Trials

FDA Business and Validator Rules

CDISC QRS Supplements



All Published QRS Supplements

SDTM Domain/ADaM Dataset: SDTM Domain/ADaM Dataset

Permission: Permission

Search by Name:

Efficient Programming

- Data → Analysis
- Implementation
- Timing
- Differences and changes

NCI Case Study: Cancer Treatment Tolerability Consortium Resources and Tools

Amylou C. Dueck, PhD
Professor of Biostatistics
Mayo Clinic, Scottsdale/Phoenix, Arizona

Patient-Focused Drug Development:
Workshop #2 to Discuss Methodologic and Other Challenges
September 18th-19th, 2025

Disclosures

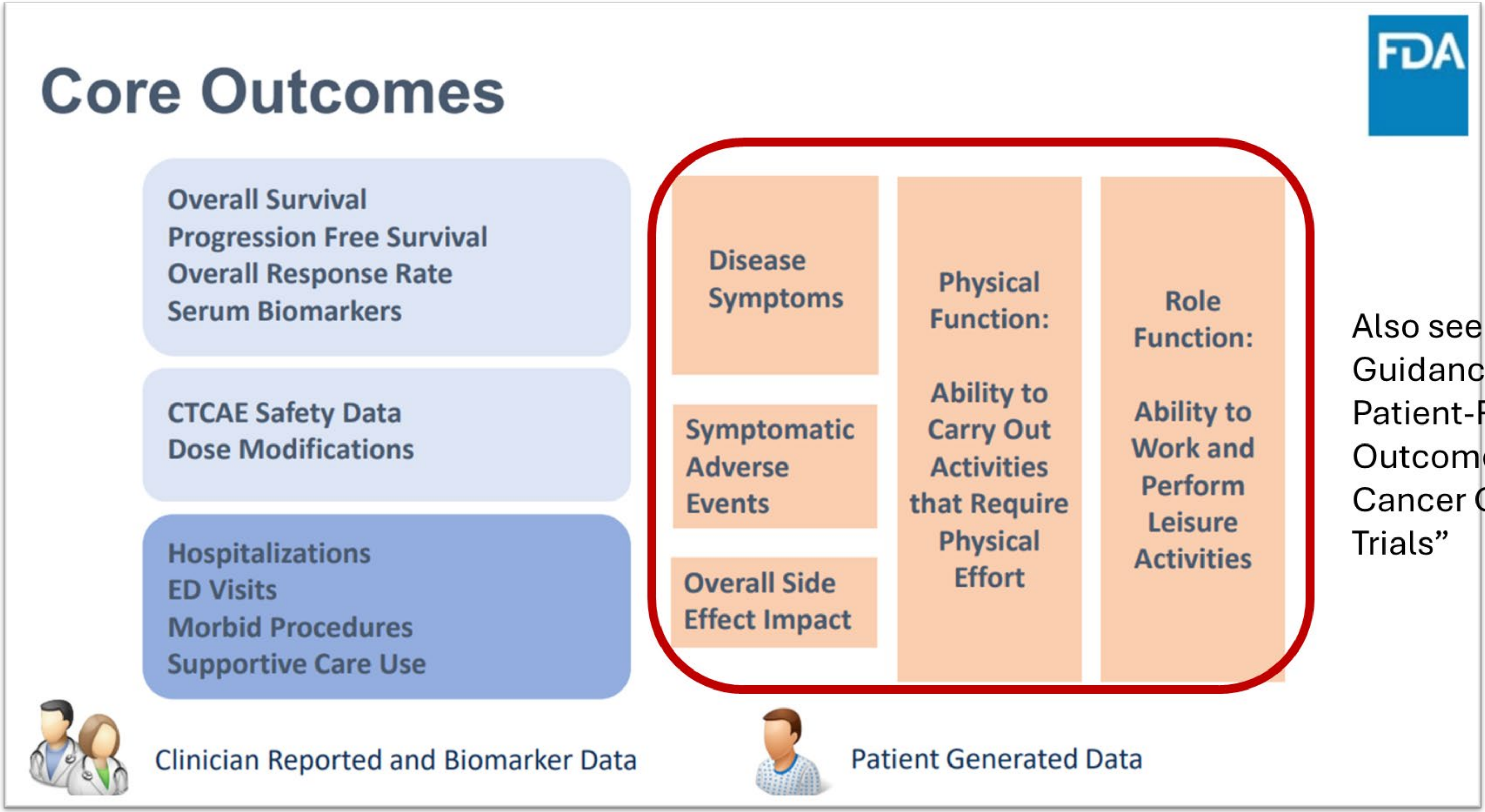
- Formal
  - Employment: Mayo Clinic
  - Advisory board: Novartis
  - IDMCs: NIH, Breast International Group, Merck

- Informal
  - I’m a US-based biostatistician who implements patient-reported outcomes (PROs) in cancer therapy clinical trials

Patient-Reported Outcomes (PROs) in cancer trials: Standardization across the protocol lifecycle

- **Protocol: SPIRIT-PRO Extension**
- **Selecting Measures: ISOQOL Recommendations & Others**
- **Statistical Analysis: SISAQOL & Graphical Recommendations**
- **Reporting: CONSORT-PRO Extension**
- **Education/Resources: PROTEUS - Trials**

SPIRIT-PRO: Calvert M, et al. JAMA. 2018; 319(5):483-494.
ISOQOL Recs & Others: Reeve BB, et al. Qual Life Res. 2013; 22(8):1889-905.
Crossnohere NL, et al. Qual Life Res. 2021; 30(1):21-40.
Piccinin C, et al. Lancet Oncol. 2023; 24(2):e86-e95.
SISAQOL & Graphics: Coens C, Pe M, et al. Lancet Oncol. 2020; 21(2):e83-e96.
Snyder C, et al. Qual Life Res. 2019; 28(2):345-356.
CONSORT-PRO: Calvert M, et al. JAMA. 2013; 309(8):814-22.
PROTEUS: https://theproteusconsortium.org/proteus-trials/



Core Outcomes

Overall Survival
Progression Free Survival
Overall Response Rate
Serum Biomarkers

CTCAE Safety Data
Dose Modifications

Hospitalizations
ED Visits
Morbid Procedures
Supportive Care Use

Disease Symptoms

Symptomatic Adverse Events

Overall Side Effect Impact

Physical Function:

Ability to Carry Out Activities that Require Physical Effort

Role Function:

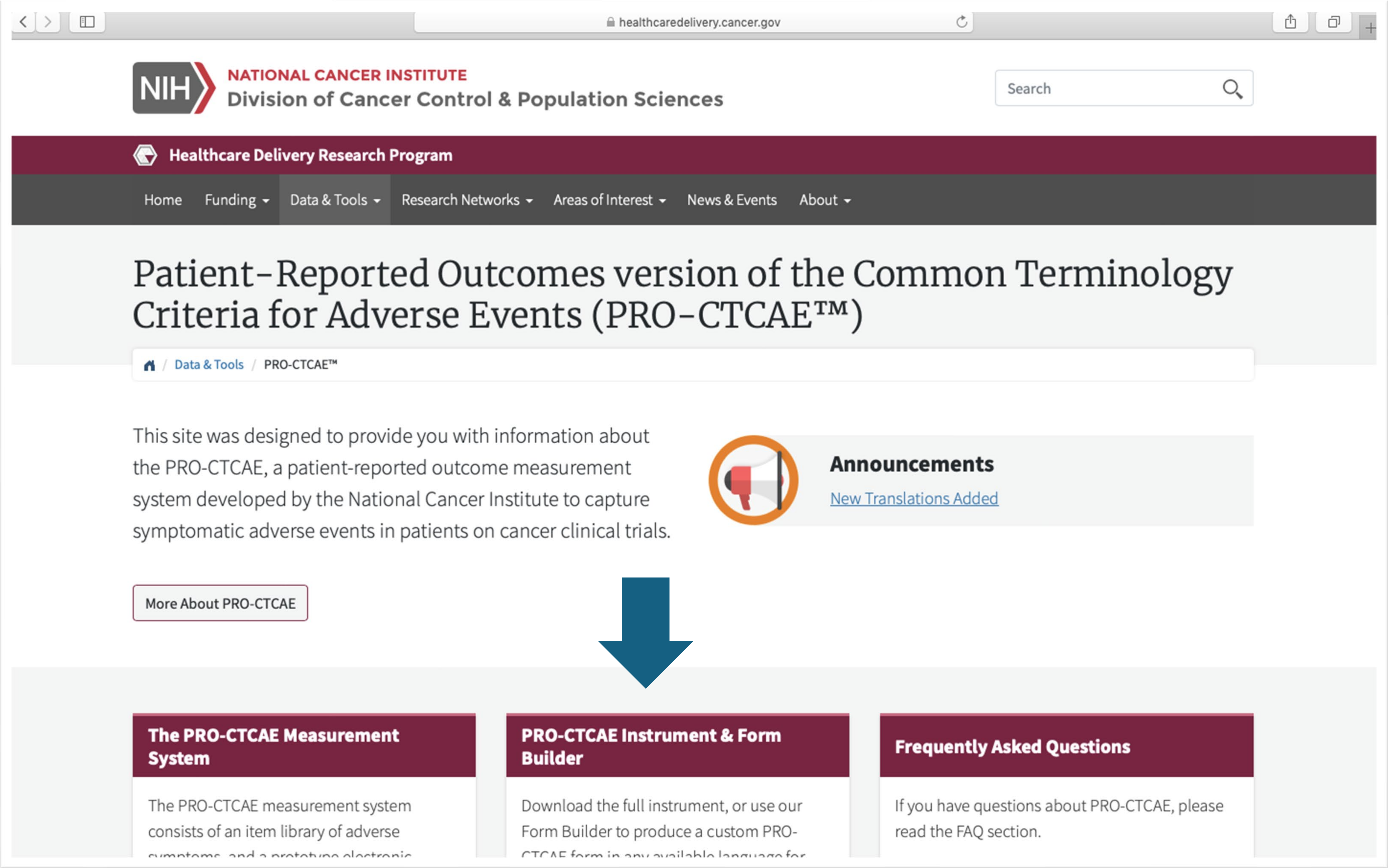
Ability to Work and Perform Leisure Activities

Clinician Reported and Biomarker Data

Patient Generated Data

Also see FDA Guidance on “Core Patient-Reported Outcomes in Cancer Clinical Trials”

Slide from: FDA WORKSHOP: 9th Annual Clinical Outcome Assessment in Cancer Clinical Trials Workshop https://www.fda.gov/media/180339/download



Patient-Reported Outcomes version of the Common Terminology Criteria for Adverse Events (PRO-CTCAE™)

This site was designed to provide you with information about the PRO-CTCAE, a patient-reported outcome measurement system developed by the National Cancer Institute to capture symptomatic adverse events in patients on cancer clinical trials.

The PRO-CTCAE Measurement System

The PRO-CTCAE measurement system consists of an item library of adverse

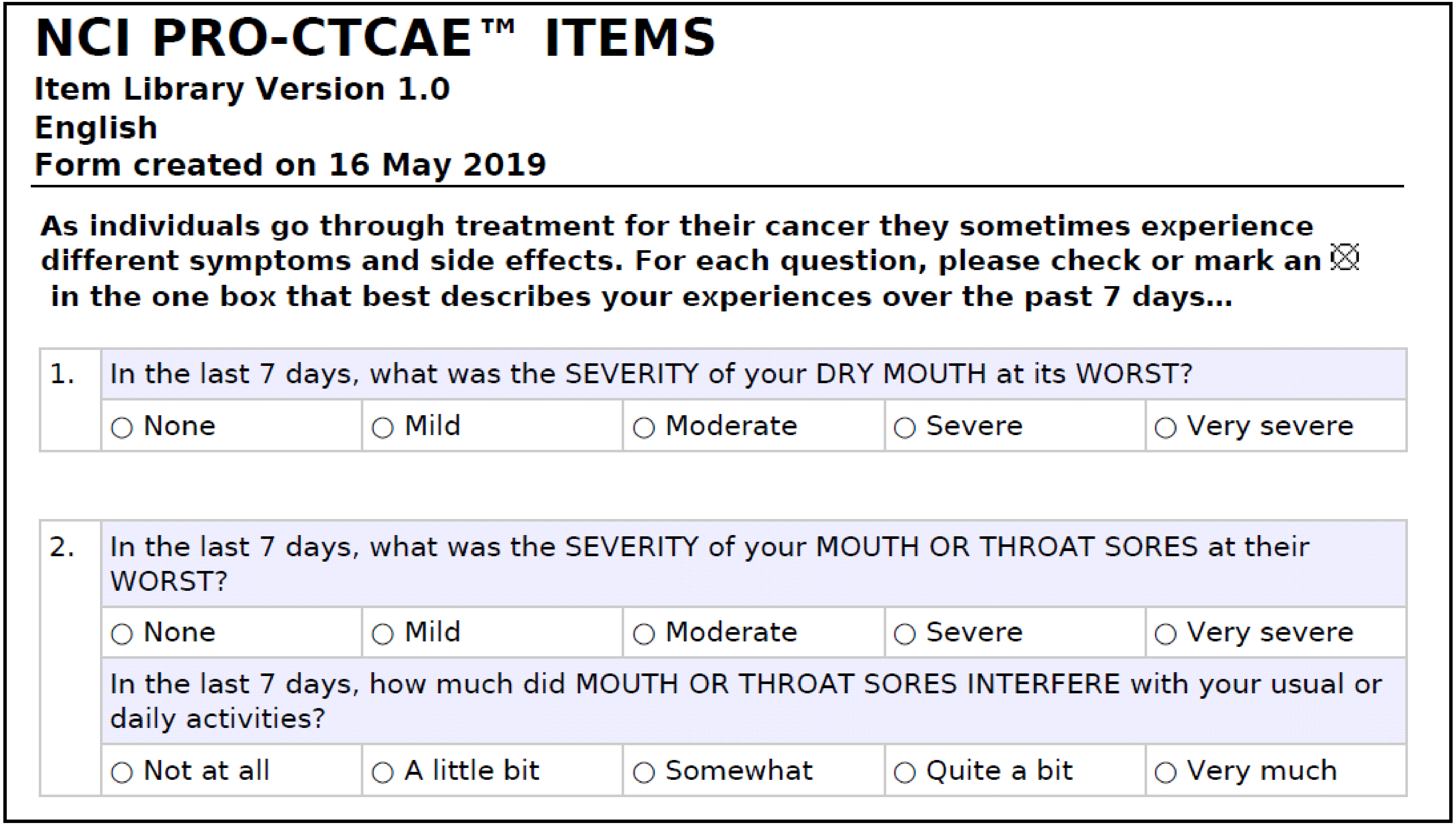
PRO-CTCAE Instrument & Form Builder

Download the full instrument, or use our Form Builder to produce a custom PRO-

Frequently Asked Questions

If you have questions about PRO-CTCAE, please read the FAQ section.

http://healthcaredelivery.cancer.gov/pro-ctcae/



NCI PRO-CTCAE™ ITEMS

Item Library Version 1.0
English
Form created on 16 May 2019

As individuals go through treatment for their cancer they sometimes experience different symptoms and side effects. For each question, please check or mark an ☒ in the one box that best describes your experiences over the past 7 days...

1.	In the last 7 days, what was the SEVERITY of your DRY MOUTH at its WORST?				
	○ None	○ Mild	○ Moderate	○ Severe	○ Very severe

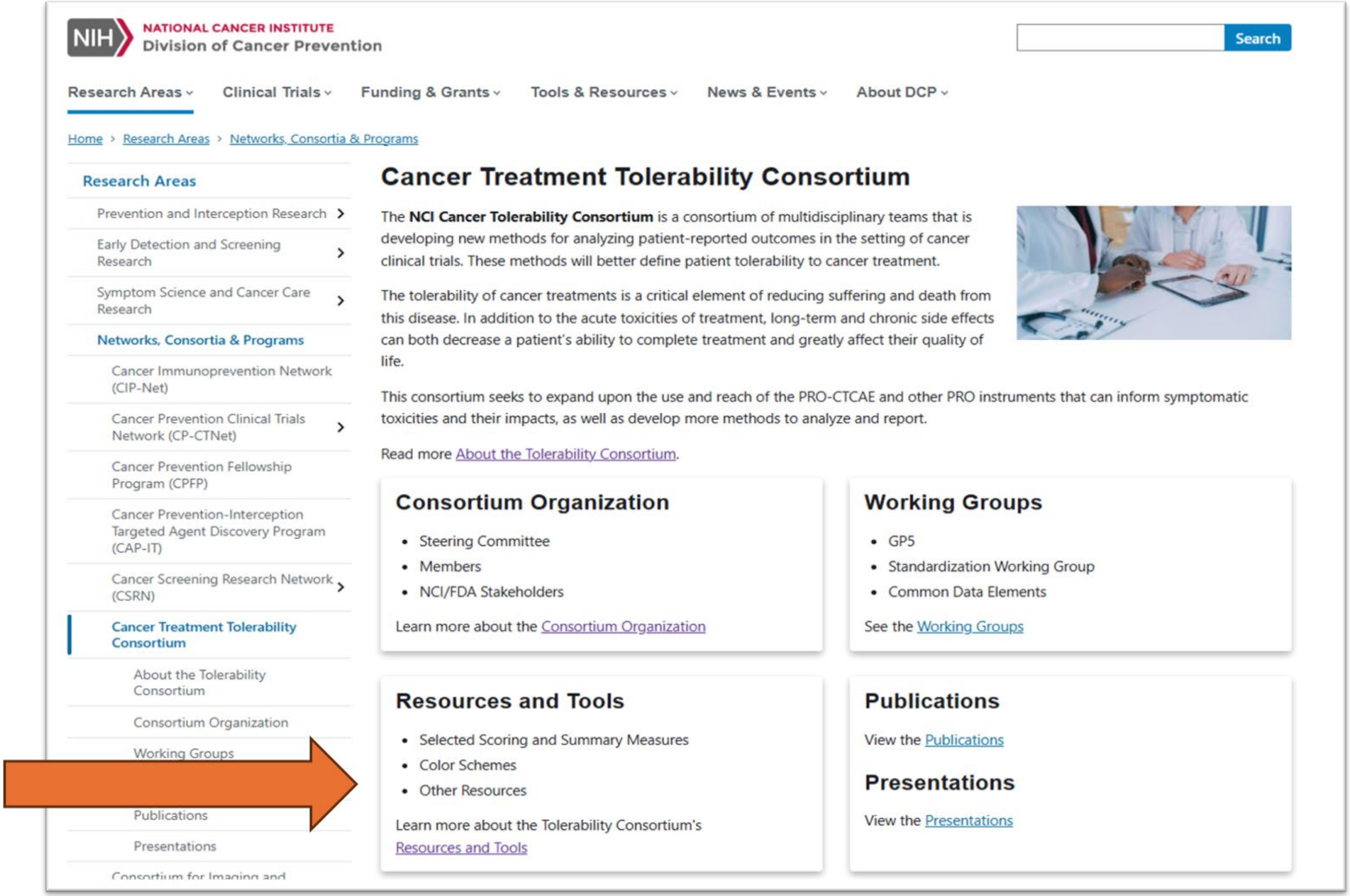
2.	In the last 7 days, what was the SEVERITY of your MOUTH OR THROAT SORES at their WORST?				
	○ None	○ Mild	○ Moderate	○ Severe	○ Very severe
	In the last 7 days, how much did MOUTH OR THROAT SORES INTERFERE with your usual or daily activities?				
	○ Not at all	○ A little bit	○ Somewhat	○ Quite a bit	○ Very much

NCI Cancer Treatment Tolerability Consortium (RFA-CA-17-052)

- Stimulate the development of methods for analyzing PRO-CTCAE, PRO, and CTCAE data in the context of definition(s) for tolerability
- Create a consortia of funded investigators to share analytic approaches
    - Basch / Dueck
    - Ganz / Rogatko
    - Gray / Wagner
    - Mohile

NIH NATIONAL CANCER INSTITUTE

https://prevention.cancer.gov/research-areas/networks-consortia-programs/cttc



NATIONAL CANCER INSTITUTE
Division of Cancer Prevention

Cancer Treatment Tolerability Consortium

The **NCI Cancer Tolerability Consortium** is a consortium of multidisciplinary teams that is developing new methods for analyzing patient-reported outcomes in the setting of cancer clinical trials. These methods will better define patient tolerability to cancer treatment.

The tolerability of cancer treatments is a critical element of reducing suffering and death from this disease. In addition to the acute toxicities of treatment, long-term and chronic side effects can both decrease a patient's ability to complete treatment and greatly affect their quality of life.

This consortium seeks to expand upon the use and reach of the PRO-CTCAE and other PRO instruments that can inform symptomatic toxicities and their impacts, as well as develop more methods to analyze and report.

Read more About the Tolerability Consortium.

Consortium Organization

- Steering Committee
- Members
- NCI/FDA Stakeholders

Learn more about the Consortium Organization

Working Groups

- GP5
- Standardization Working Group
- Common Data Elements

See the Working Groups

Resources and Tools

- Selected Scoring and Summary Measures
- Color Schemes
- Other Resources

Learn more about the Tolerability Consortium's Resources and Tools

Publications

View the Publications

Presentations

View the Presentations

https://prevention.cancer.gov/research-areas/networks-consortia-programs/cttc/resources-and-tools

Plus – forthcoming Tolerability Consortium Special Issue in ***Cancer***

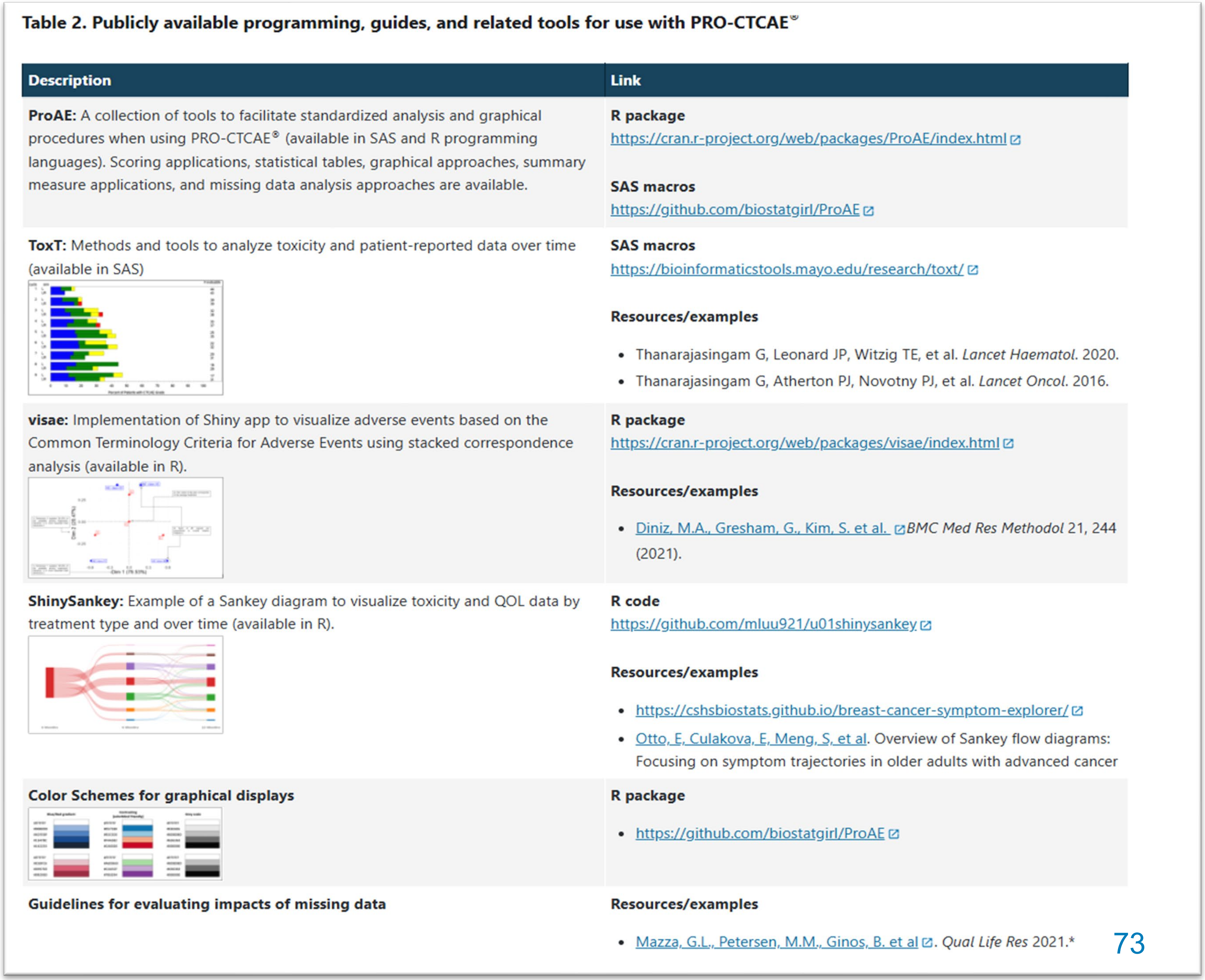
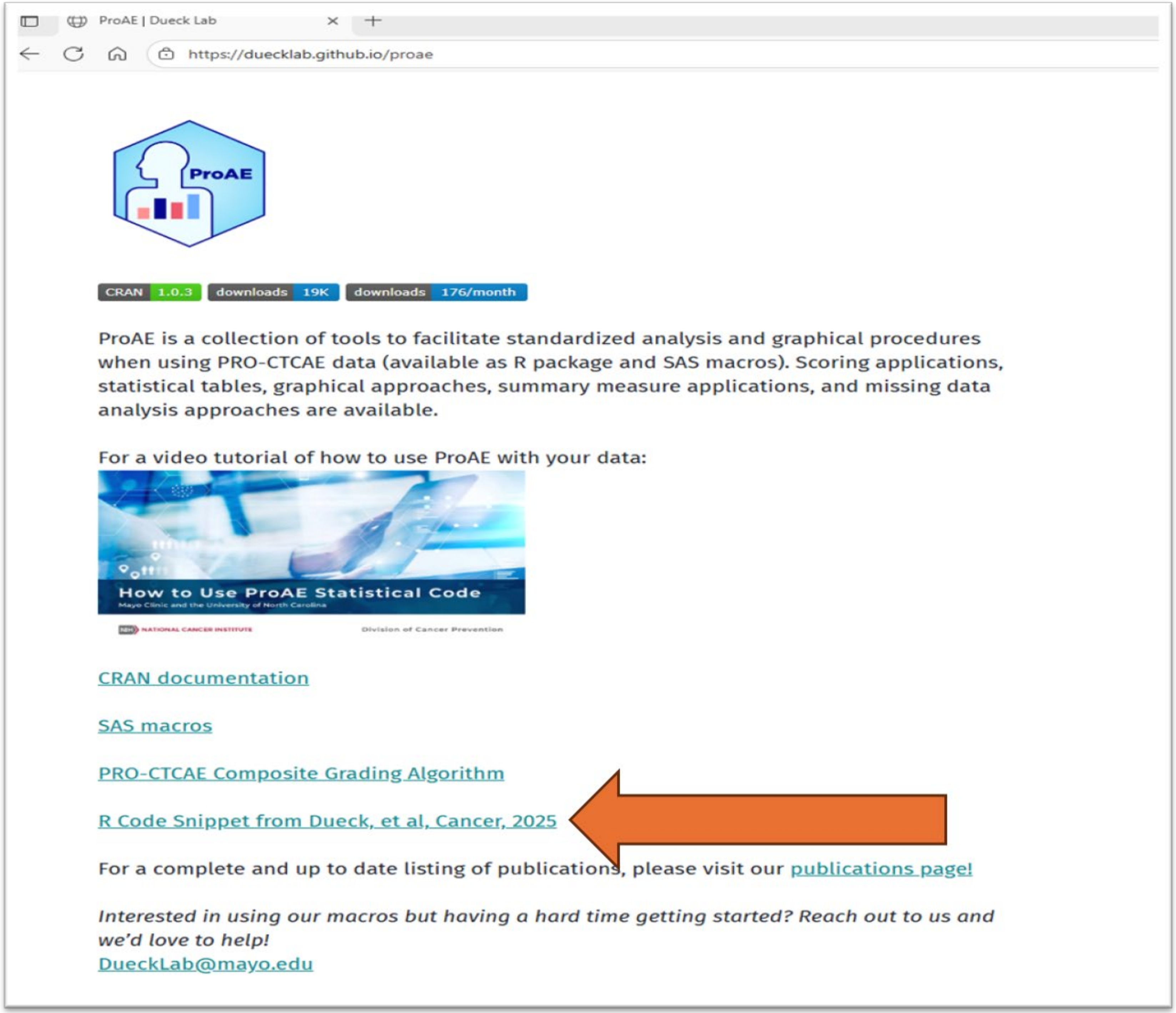


Table 2. Publicly available programming, guides, and related tools for use with PRO-CTCAE®

Description	Link
ProAE: A collection of tools to facilitate standardized analysis and graphical procedures when using PRO-CTCAE® (available in SAS and R programming languages). Scoring applications, statistical tables, graphical approaches, summary measure applications, and missing data analysis approaches are available.	**R package** https://cran.r-project.org/web/packages/ProAE/index.html **SAS macros** https://github.com/biostatgirl/ProAE
ToxT: Methods and tools to analyze toxicity and patient-reported data over time (available in SAS)	**SAS macros** https://bioinformaticstools.mayo.edu/research/toxt/ **Resources/examples** • Thanarajasingam G, Leonard JP, Witzig TE, et al. *Lancet Haematol*. 2020. • Thanarajasingam G, Atherton PJ, Novotny PJ, et al. *Lancet Oncol*. 2016.
visae: Implementation of Shiny app to visualize adverse events based on the Common Terminology Criteria for Adverse Events using stacked correspondence analysis (available in R).	**R package** https://cran.r-project.org/web/packages/visae/index.html **Resources/examples** • Diniz, M.A., Gresham, G., Kim, S. et al. *BMC Med Res Methodol* 21, 244 (2021).
ShinySankey: Example of a Sankey diagram to visualize toxicity and QOL data by treatment type and over time (available in R).	**R code** https://github.com/mluu921/u01shinysankey **Resources/examples** • https://cshsbiostats.github.io/breast-cancer-symptom-explorer/ • Otto, E, Culakova, E, Meng, S, et al. Overview of Sankey flow diagrams: Focusing on symptom trajectories in older adults with advanced cancer
Color Schemes for graphical displays	**R package** • https://github.com/biostatgirl/ProAE
Guidelines for evaluating impacts of missing data	**Resources/examples** • Mazza, G.L., Petersen, M.M., Ginos, B. et al. *Qual Life Res* 2021.*



ProAE is a collection of tools to facilitate standardized analysis and graphical procedures when using PRO-CTCAE data (available as R package and SAS macros). Scoring applications, statistical tables, graphical approaches, summary measure applications, and missing data analysis approaches are available.

For a video tutorial of how to use ProAE with your data:

[CRAN documentation]

[SAS macros]

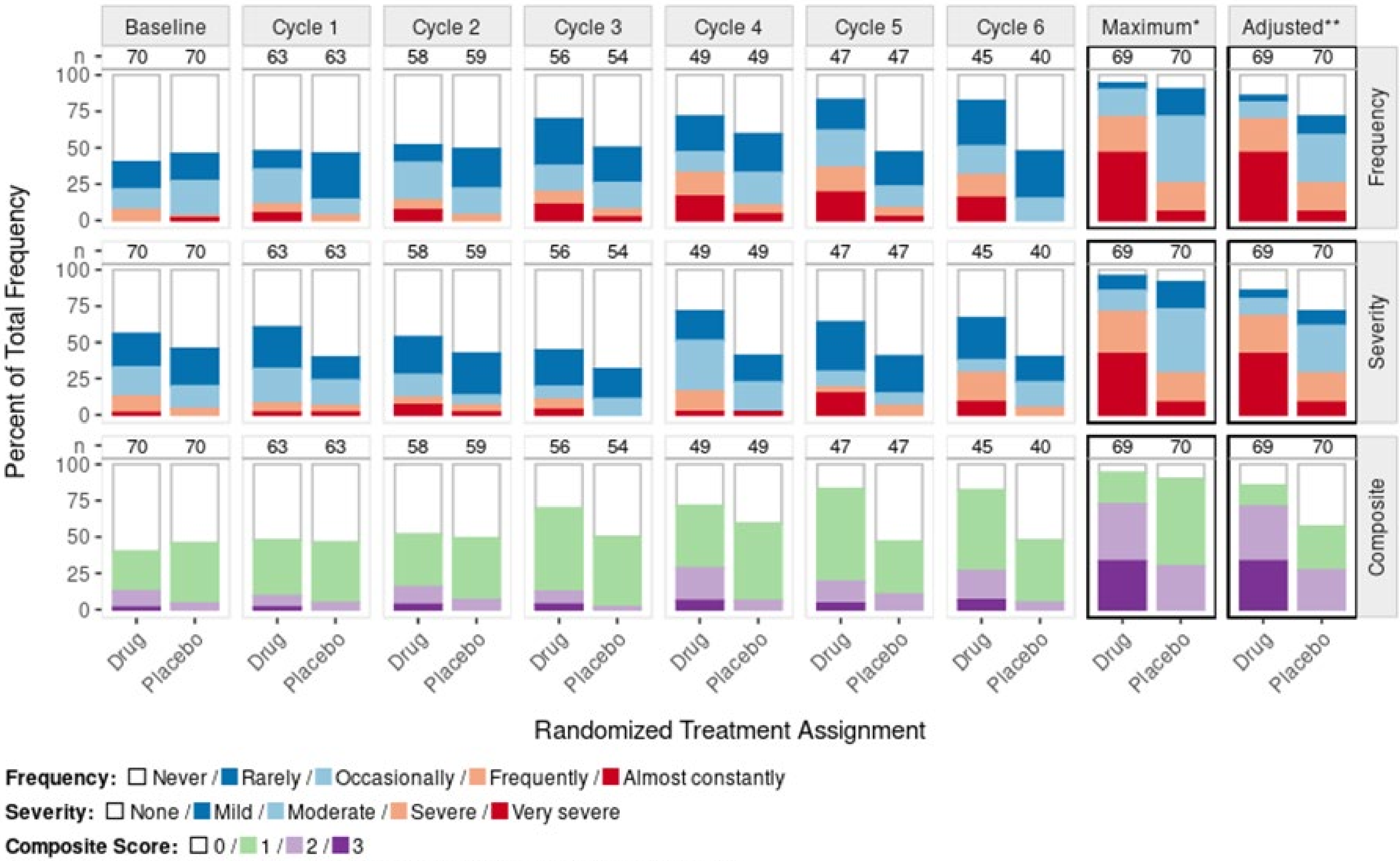
[PRO-CTCAE Composite Grading Algorithm]

[R Code Snippet from Dueck, et al, Cancer, 2025]

For a complete and up to date listing of publications, please visit our publications page!

Interested in using our macros but having a hard time getting started? Reach out to us and we'd love to help!
DueckLab@mayo.edu

https://duecklab.github.io



Column labels (n) show the number of subjects with an observed symptom score.
*Maximum score reported post-baseline per patient.
**Maximum score reported post-baseline per patient when including only scores that were worse than the patient's baseline score.

Inavolisib label (approved 10/10/2024 with palbociclib and fulvestrant for treatment of endocrine-resistant, PIK3CA-mutated, HR-positive, HER2-negative, advanced breast cancer)

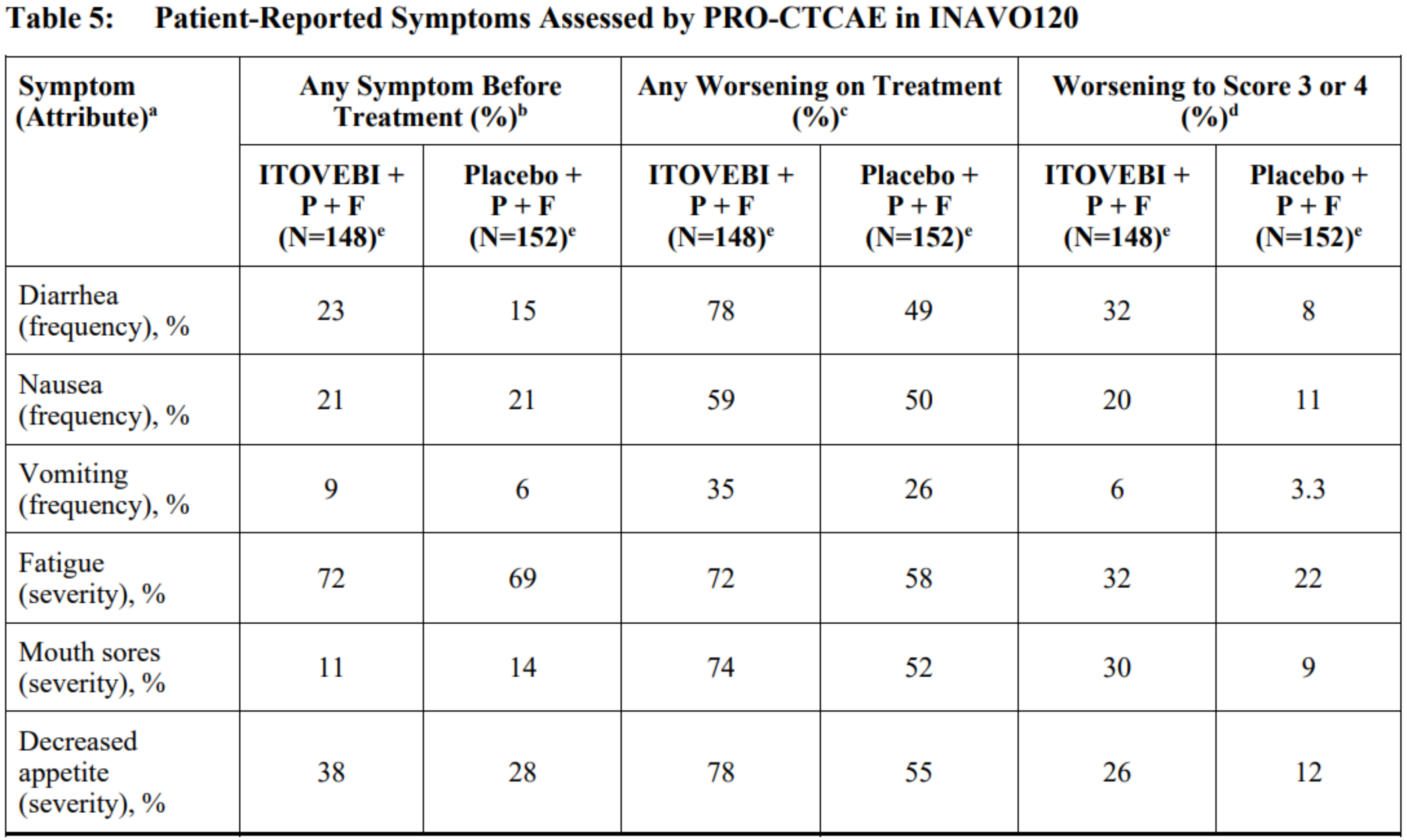


Table 5: Patient-Reported Symptoms Assessed by PRO-CTCAE in INAVO120

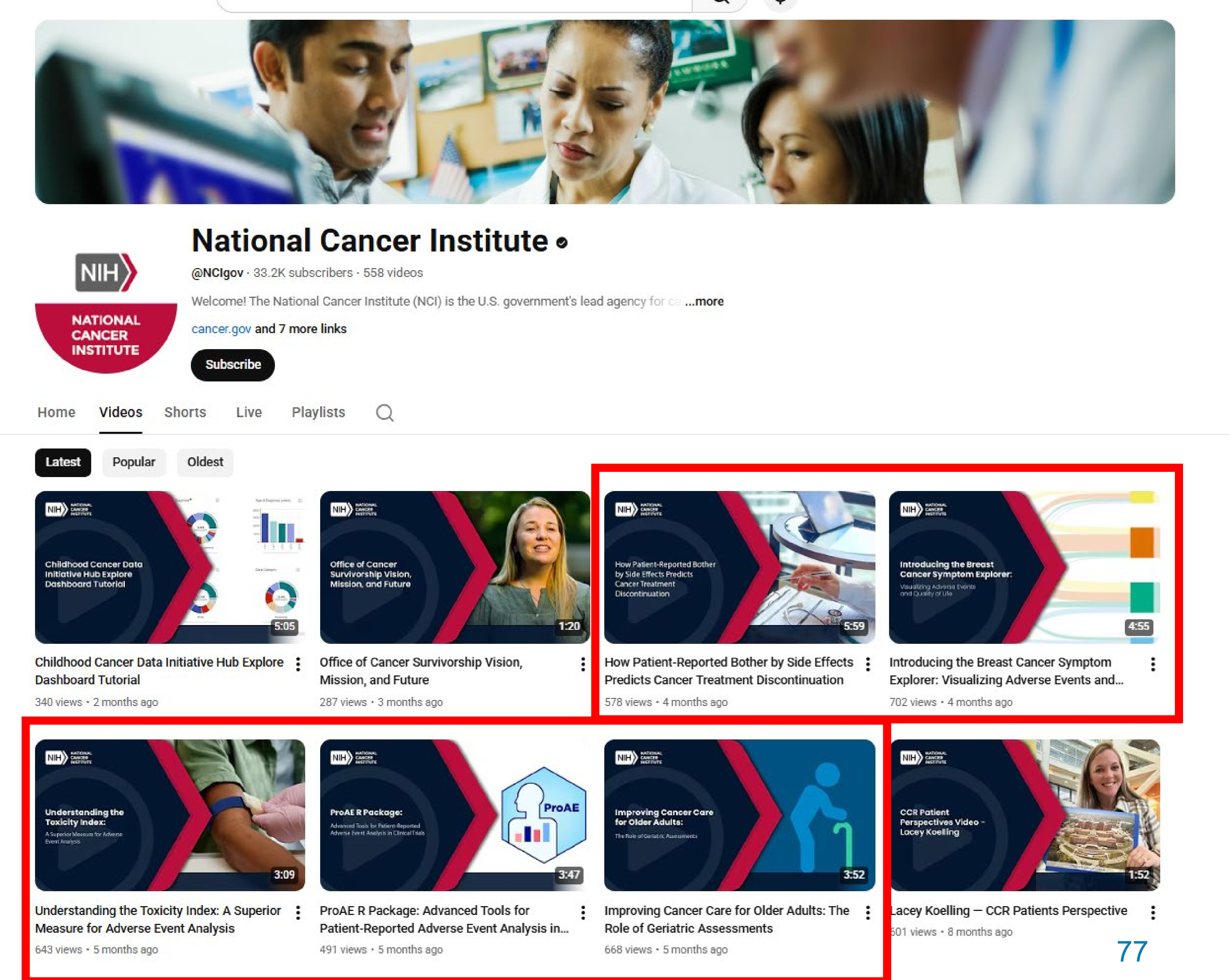
Symptom (Attribute)[a]	**Any Symptom Before Treatment (%)**[b]		**Any Worsening on Treatment (%)**[c]		**Worsening to Score 3 or 4 (%)**[d]	
	ITOVEBI + P + F (N=148)[e]	**Placebo + P + F (N=152)**[e]	**ITOVEBI + P + F (N=148)**[e]	**Placebo + P + F (N=152)**[e]	**ITOVEBI + P + F (N=148)**[e]	**Placebo + P + F (N=152)**[e]
Diarrhea (frequency), %	23	15	78	49	32	8
Nausea (frequency), %	21	21	59	50	20	11
Vomiting (frequency), %	9	6	35	26	6	3.3
Fatigue (severity), %	72	69	72	58	32	22
Mouth sores (severity), %	11	14	74	52	30	9
Decreased appetite (severity), %	38	28	78	55	26	12

At Cycle 2 Day 15, the proportion of patients with MBI responses of “not at all” were 25% in the ITOVEBI with palbociclib and fulvestrant arm and 53% in the placebo with palbociclib and fulvestrant arm. Through 31 cycles of treatment, patients in the ITOVEBI with palbociclib and fulvestrant arm reported more side effect bother compared to the placebo with palbociclib and fulvestrant arm.

https://www.accessdata.fda.gov/drugsatfda_docs/label/2024/219249s001lbl.pdf

NCI YouTube Channel

Tolerability Consortium videos highlighted in red boxes, including videos describing analytic approaches



National Cancer Institute

@NCIgov · 33.2K subscribers · 558 videos

Welcome! The National Cancer Institute (NCI) is the U.S. government's lead agency for ca ...more

cancer.gov and 7 more links

Childhood Cancer Data Initiative Hub Explore Dashboard Tutorial
340 views · 2 months ago

Office of Cancer Survivorship Vision, Mission, and Future
287 views · 3 months ago

How Patient-Reported Bother by Side Effects Predicts Cancer Treatment Discontinuation
578 views · 4 months ago

Introducing the Breast Cancer Symptom Explorer: Visualizing Adverse Events and...
702 views · 4 months ago

Understanding the Toxicity Index: A Superior Measure for Adverse Event Analysis
643 views · 5 months ago

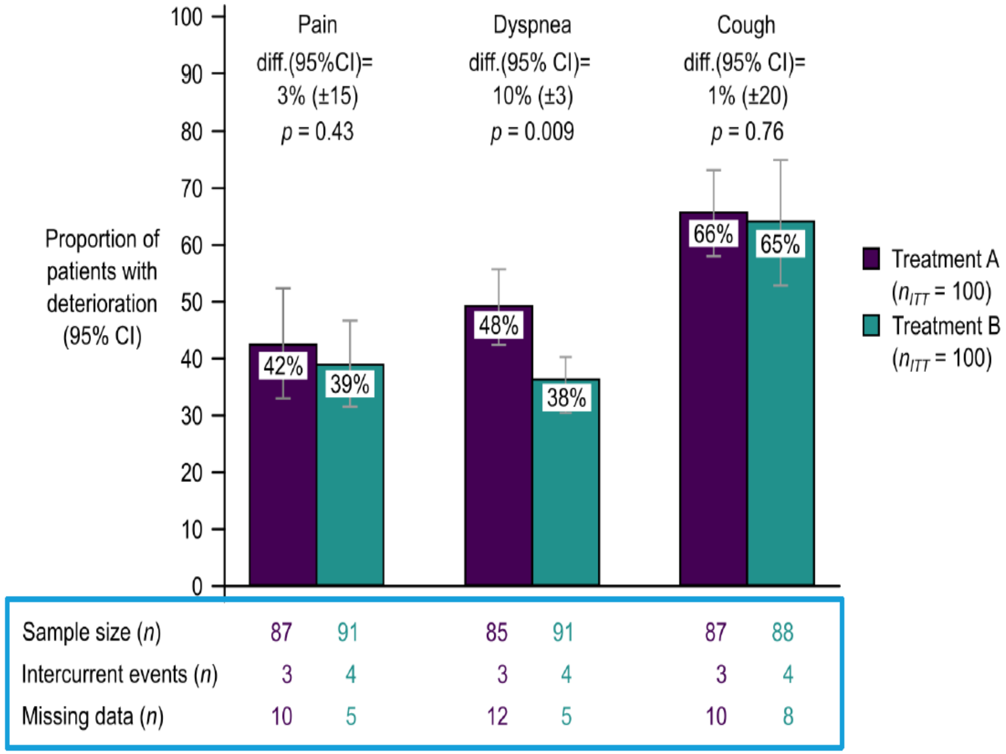
ProAE R Package: Advanced Tools for Patient-Reported Adverse Event Analysis in...
491 views · 5 months ago

Improving Cancer Care for Older Adults: The Role of Geriatric Assessments
668 views · 5 months ago

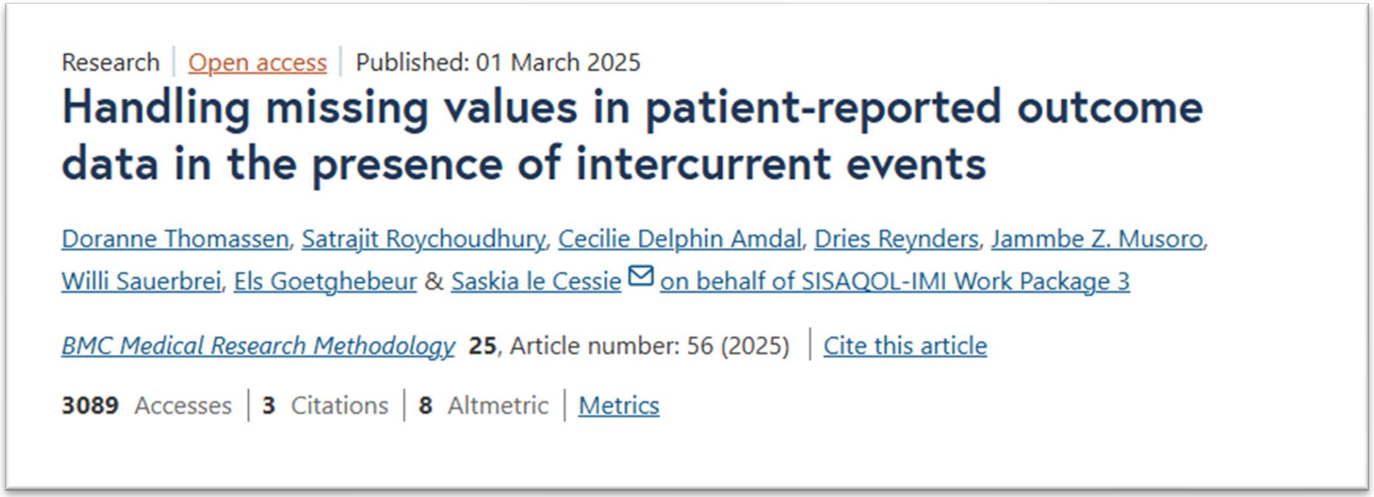
Lacey Koelling — CCR Patients Perspective
01 views · 8 months ago

Additional Code Resources Expected Soon

- [SISAQOL – Setting International Standards in Analyzing Patient-Reported Outcomes and Quality of Life endpoints](https://www.sisaqol-imi.org/) (https://www.sisaqol-imi.org/)
- Recommendations for the analysis of PROs in single-arm and randomized trials, including standardized visualizations
- Public launch of final recommendations in February 2025 (official publication pending)
- Anticipate code to be developed / made available



	Pain A	Pain B	Dyspnea A	Dyspnea B	Cough A	Cough B
Sample size (n)	87	91	85	91	87	88
Intercurrent events (n)	3	4	3	4	3	4
Missing data (n)	10	5	12	5	10	8



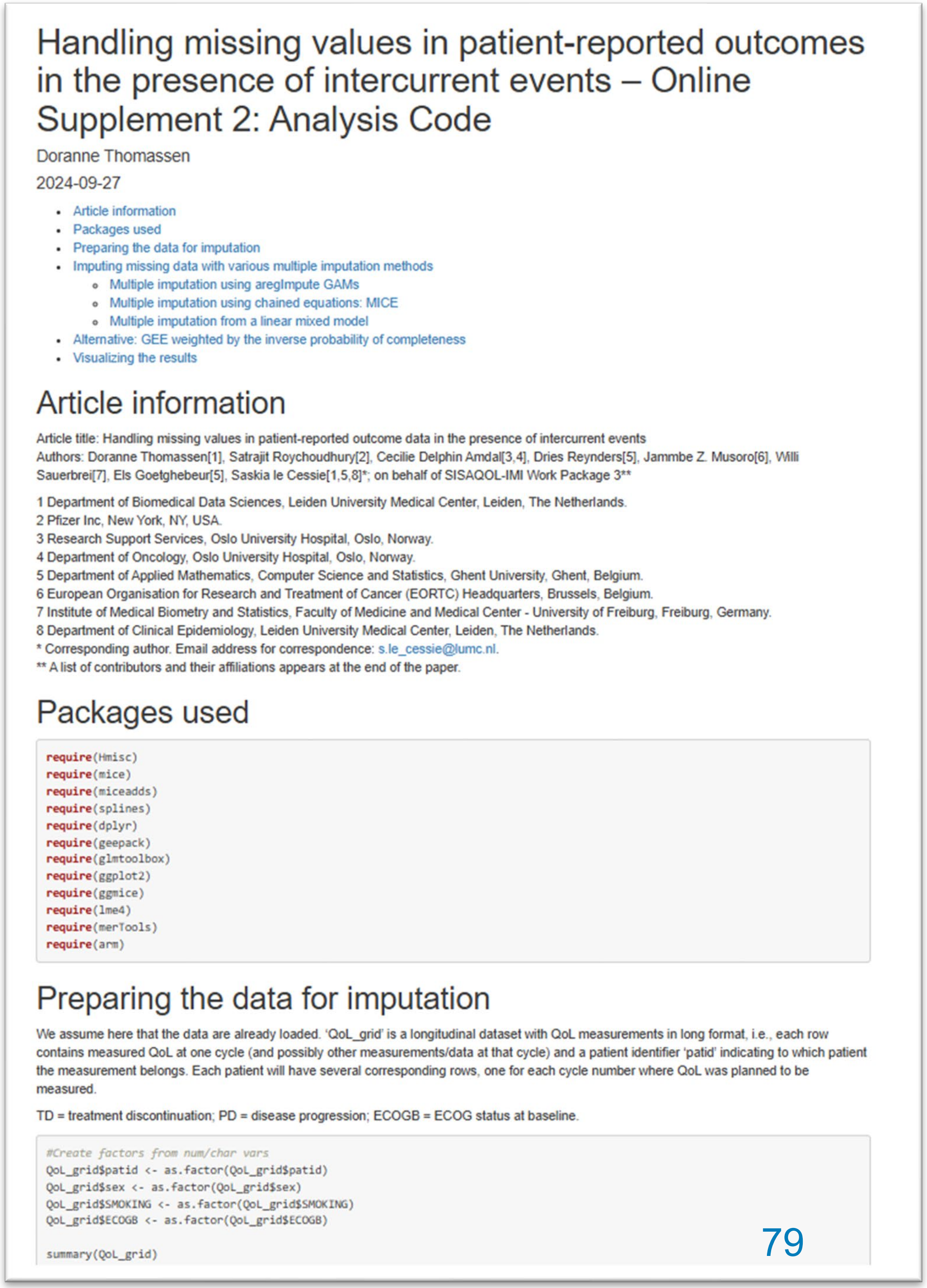
Research | Open access | Published: 01 March 2025

Handling missing values in patient-reported outcome data in the presence of intercurrent events

Doranne Thomassen, Satrajit Roychoudhury, Cecilie Delphin Amdal, Dries Reynders, Jammbe Z. Musoro, Willi Sauerbrei, Els Goetghebeur & Saskia le Cessie ✉ on behalf of SISAQOL-IMI Work Package 3

BMC Medical Research Methodology **25**, Article number: 56 (2025) | Cite this article

3089 Accesses | **3** Citations | **8** Altmetric | Metrics



Handling missing values in patient-reported outcomes in the presence of intercurrent events – Online Supplement 2: Analysis Code

Doranne Thomassen

2024-09-27

- Article information
- Packages used
- Preparing the data for imputation
- Imputing missing data with various multiple imputation methods
    - Multiple imputation using aregImpute GAMs
    - Multiple imputation using chained equations: MICE
    - Multiple imputation from a linear mixed model
- Alternative: GEE weighted by the inverse probability of completeness
- Visualizing the results

Article information

Article title: Handling missing values in patient-reported outcome data in the presence of intercurrent events
Authors: Doranne Thomassen[1], Satrajit Roychoudhury[2], Cecilie Delphin Amdal[3,4], Dries Reynders[5], Jammbe Z. Musoro[6], Willi Sauerbrei[7], Els Goetghebeur[5], Saskia le Cessie[1,5,8]*; on behalf of SISAQOL-IMI Work Package 3**

1 Department of Biomedical Data Sciences, Leiden University Medical Center, Leiden, The Netherlands.
2 Pfizer Inc, New York, NY, USA.
3 Research Support Services, Oslo University Hospital, Oslo, Norway.
4 Department of Oncology, Oslo University Hospital, Oslo, Norway.
5 Department of Applied Mathematics, Computer Science and Statistics, Ghent University, Ghent, Belgium.
6 European Organisation for Research and Treatment of Cancer (EORTC) Headquarters, Brussels, Belgium.
7 Institute of Medical Biometry and Statistics, Faculty of Medicine and Medical Center - University of Freiburg, Freiburg, Germany.
8 Department of Clinical Epidemiology, Leiden University Medical Center, Leiden, The Netherlands.
* Corresponding author. Email address for correspondence: s.le_cessie@lumc.nl.
** A list of contributors and their affiliations appears at the end of the paper.

Packages used

```
require(Hmisc)
require(mice)
require(miceadds)
require(splines)
require(dplyr)
require(geepack)
require(glmtoolbox)
require(ggplot2)
require(ggmice)
require(lme4)
require(merTools)
require(arm)
```

Preparing the data for imputation

We assume here that the data are already loaded. 'QoL_grid' is a longitudinal dataset with QoL measurements in long format, i.e., each row contains measured QoL at one cycle (and possibly other measurements/data at that cycle) and a patient identifier 'patid' indicating to which patient the measurement belongs. Each patient will have several corresponding rows, one for each cycle number where QoL was planned to be measured.

TD = treatment discontinuation; PD = disease progression; ECOGB = ECOG status at baseline.

```
#Create factors from num/char vars
QoL_grid$patid <- as.factor(QoL_grid$patid)
QoL_grid$sex <- as.factor(QoL_grid$sex)
QoL_grid$SMOKING <- as.factor(QoL_grid$SMOKING)
QoL_grid$ECOGB <- as.factor(QoL_grid$ECOGB)

summary(QoL_grid)
```

Concluding thoughts

- Challenges
  - Enormous range of study designs, systems, PRO measures, analytic approaches, and visualizations
  - Limited funding
  - No requirement to share analytic code

- Hope for future
  - Standardization efforts for PROs in (cancer) clinical trials have enabled investment in code development

Learnings from Oncology: Submitting Patient-Reported Outcome Data from Cancer Clinical Trials

Flora Mulkey
Division of Biometrics V
CDER/OTS/Office of Biostatistics

Standardizing Data Submissions

- FDA’s Guidance for Industry Technical Specification Document
- Why it is important and how it impacts regulatory review of PRO data

Patient-Reported Outcomes (PROs): Current State and Regulatory Context in Oncology

- Traditional Common Terminology Criteria for Adverse Event (CTCAE) grading may not adequately reflect patient experience or functional impact of treatment-related adverse events

- Accurate measurement of how patients feel and function provides valuable additional information to assess benefit-risk of cancer therapies (21st Century Cures Act)

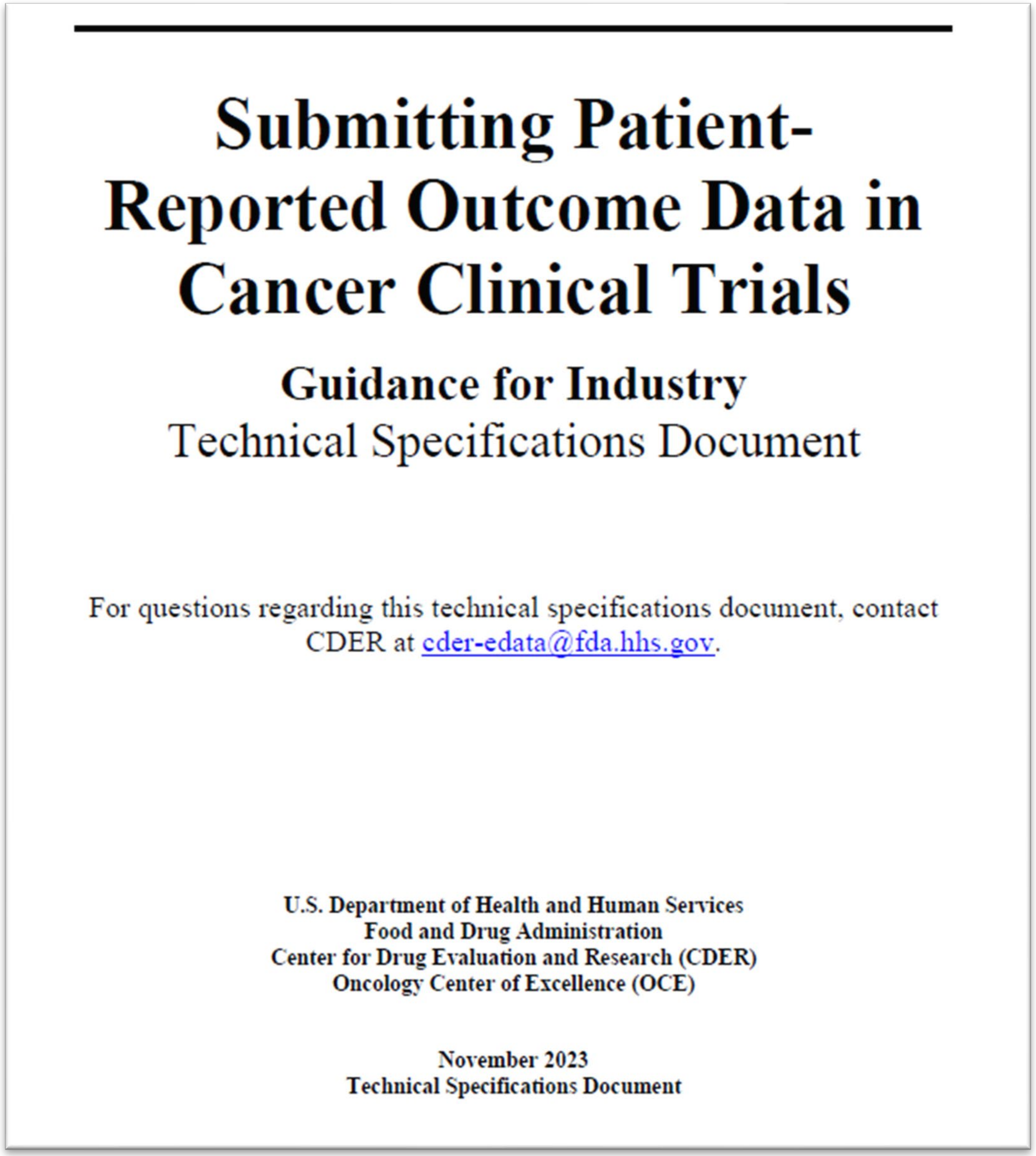
- PROs are complementary secondary or exploratory endpoints to traditional oncology efficacy endpoints

Patient-Reported Outcomes in Oncology

- Multi-faceted value of PROs in oncologic drug development
  - Inform tolerability in dose optimization
  - Contribute supportive evidence for regulatory decision-making
  - Provide valuable insights into the impact of treatment on patients' symptoms, functioning, and side effects

The collection and submission of PRO data has not always been consistent, nor has it always been conducive to analysis

Guidance Objectives:



Submitting Patient-Reported Outcome Data in Cancer Clinical Trials

Guidance for Industry

Technical Specifications Document

For questions regarding this technical specifications document, contact CDER at cder-edata@fda.hhs.gov.

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research (CDER)
Oncology Center of Excellence (OCE)

November 2023
Technical Specifications Document

- **Develop technical specifications** for the submission of standardized data and documentation for patient-reported outcome (PRO) data in cancer clinical trials
- **Describe how PRO objectives impact evaluation of**
    - Intercurrent events (Progression, death, AEs)
    - Analysis population
- **Harmonization of terminology** for
    - available data rate and completion rate
- **Required table/figure structure** for
    - Patient disposition
    - Available data rate and completion rate
    - PRO score distributions

https://www.fda.gov/regulatory-information/search-fda-guidance-documents/submitting-patient-reported-outcome-data-cancer-clinical-trials

Scope of this Technical Specifications Guidance

This guidance supplements:

- FDA's Core PRO in Cancer Clinical Trials Draft Guidance for Industry
- Patient Focused Drug Development Guidance series (PFDD G1-G4)

Not specific to a certain therapeutic area in oncology nor restricted to PROs:

- Specifications do not include all datasets, variables, or controlled terminologies
- May be relevant to other types of Clinical Outcome Assessments (clinician-reported, observer-reported, and performance outcome measures) in cancer clinical trials

CDISC: Controlled Data Interchange Standards Consortium

- FDA requires use of CDISC standards for data submitted in support of marketing applications for all trials initiated after December 17, 2016

- CDISC standards are required for the following types of data
  - **SDTM** (Study Data Tabulation Model: raw clinical trial data)
  - **ADaM** (Analysis Data Model: Analysis ready clinical data)
  - **SEND** (Standard for Exchange of Nonclinical Data: nonclinical study data)
  - **Define-XML** (Case Report Tabulation Data Definition Specification for the metadata that accompany SEND, SDTM, and ADaM datasets)

https://www.fda.gov/industry/fda-data-standards-advisory-board/study-data-standards-resources
https://www.fda.gov/regulatory-information/search-fda-guidance-documents/data-standards-catalog

SDTM and ADaM Dataset Specifications

SDTM (Study Data Tabulation Model) → ***Questionnaire (QS)*** dataset

- PRO-specific variables (e.g., **data collection mode**, **data collector**, and **language**)

  Includes recommendations for capturing reasons for and handling of missing data

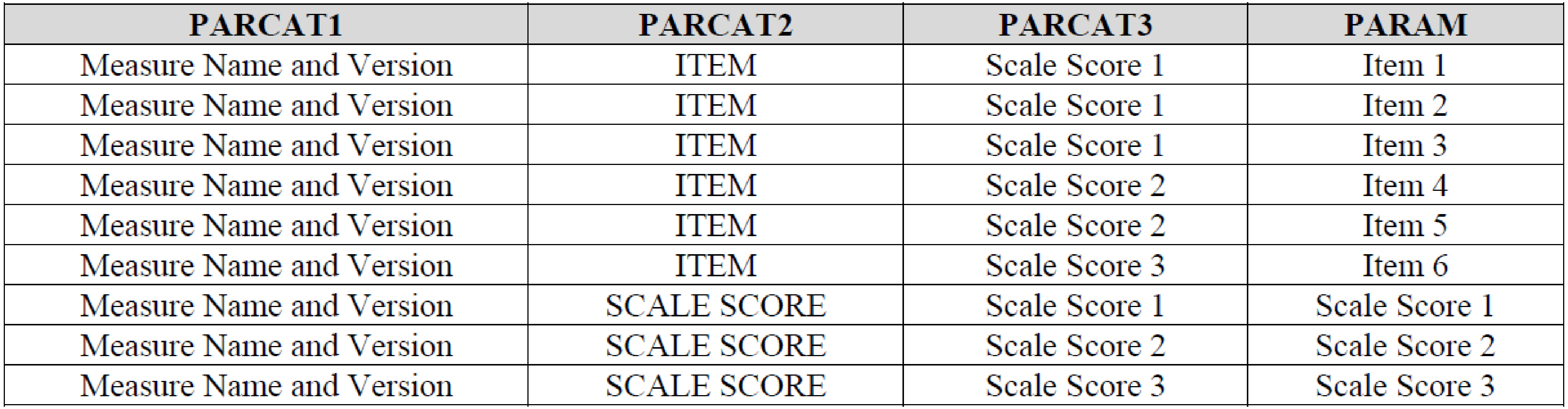
ADaM (Analysis Data Model) → ***ADQS*** dataset

- Newly-defined variables relevant to PROs:
  - **PROEXPFL**- indicator that the PRO parameter was expected to be assessed at that timepoint (use PROEX1FL and PROEX2FL if a single item will be used for both objectives)
  - **PROSCMFL**- PRO score completed indicator
  - **AREASND-** describes the reason for missing PRO data
  - **SCBLFL**- indicator that the baseline record was sourced from a screening assessment
  - **FPDDT-** first progressive disease date
    - If assessed by both investigator and independent review committee may differentiate using: **FPDINVDT** and **FPDIRCDT**

ADaM Dataset Specifications

Recommended dataset structure in ADaM using PARCATy variables is based on the type of PRO measure and the number of summary scores calculated (Guidance section 3.2.3)

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

Missing data and intercurrent events in ADaM (Guidance section 3.2.2)

- Each item score and summary score within the PRO measure should be presented for subjects in the randomized population **at each planned** (per protocol) **PRO assessment timepoint**
- Phantom records should be included in cases of treatment pause or failed administration

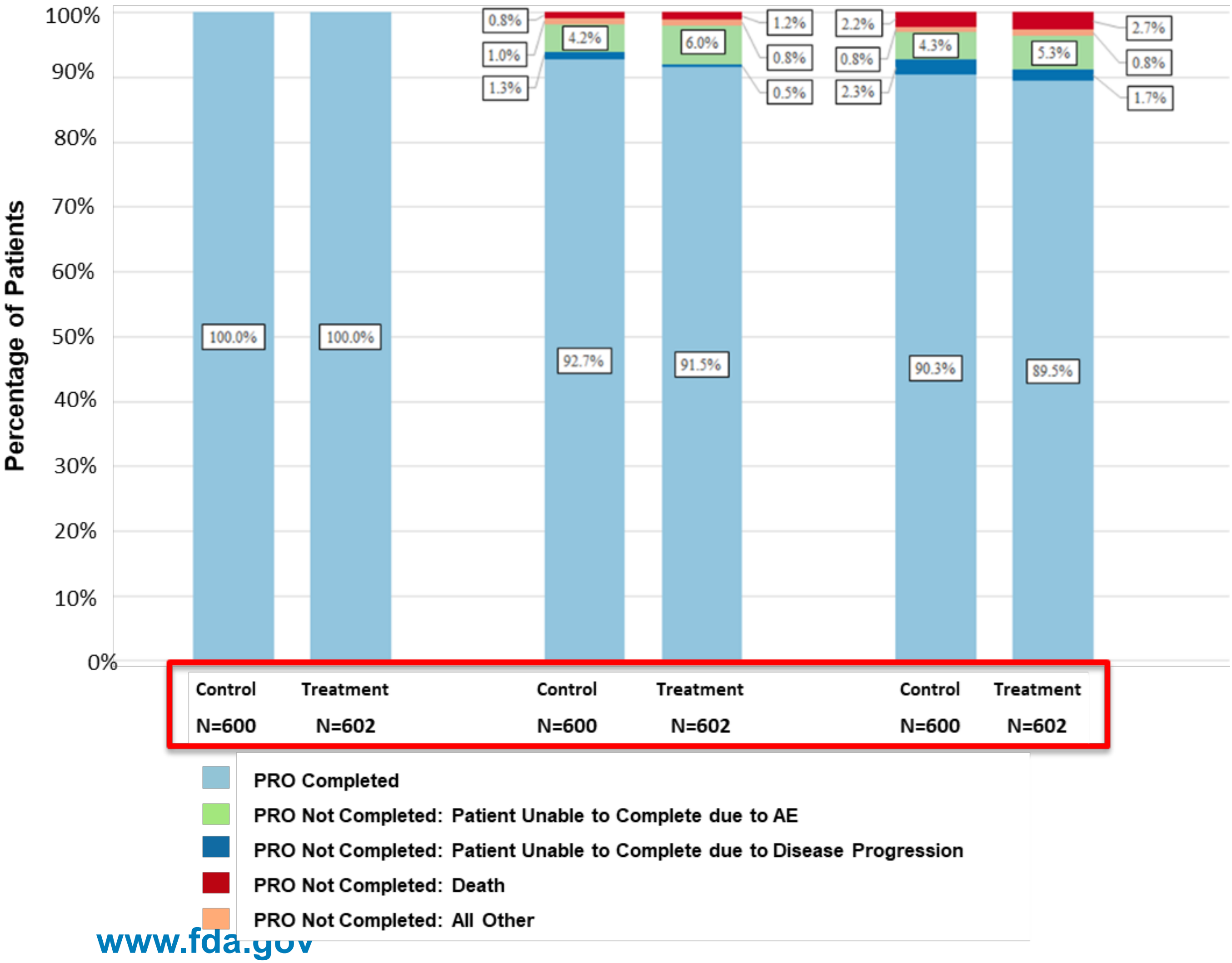
Specifications for Tables and Figures

Requested tables and figures include:

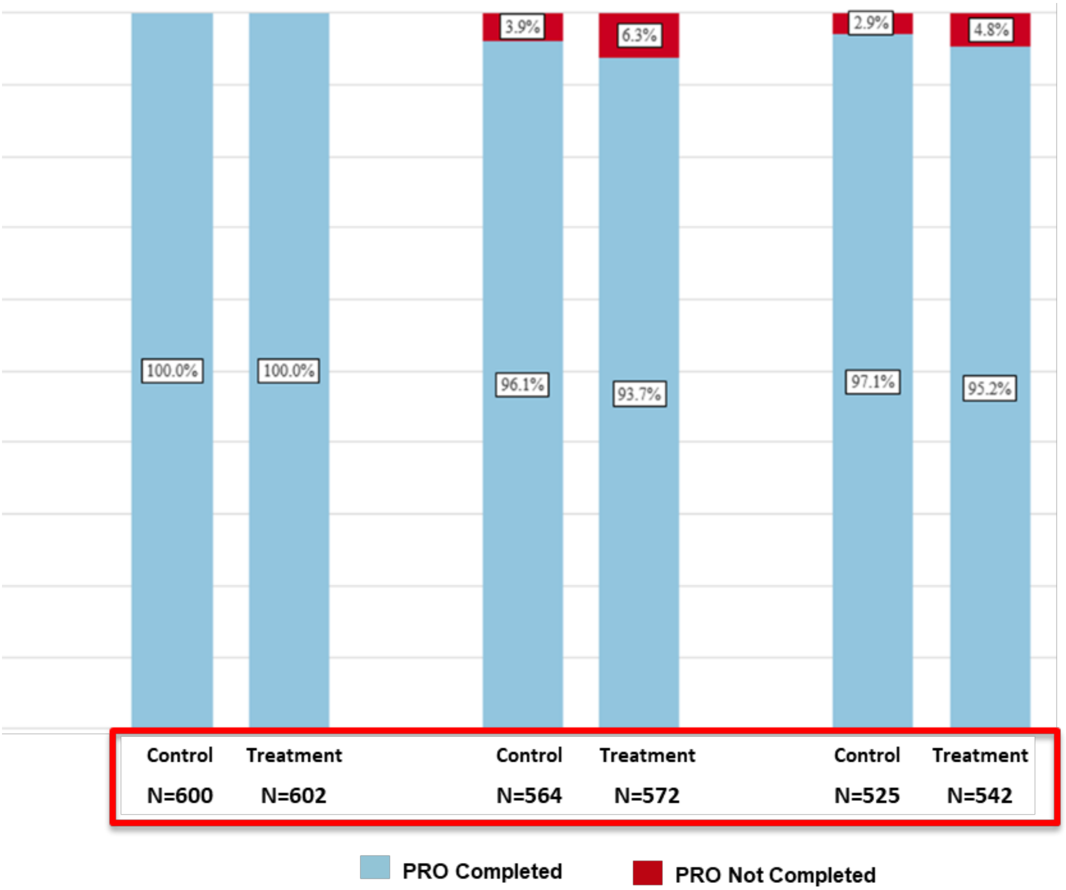
1. **Patient Disposition** (occurrence of intercurrent events)
2. **PRO Data Quality**
   - **Available Data Rate** (previously referred to as completion rate) for Clinical Benefit: based on randomized patients
   - **Completion Rate** (previously referred to as compliance rate) for Safety and Tolerability: based on patients where PRO is expected
3. **PRO Measure Distributions**
   - Distribution of Response (categorical & continuous)
   - Distribution of Change in Response from Baseline
4. **Incidence of Healthcare Utilization**

PRO Data Quality Figures

- **Available Data Rate:** Clinical Benefit



- **Completion Rate:** Safety and Tolerability

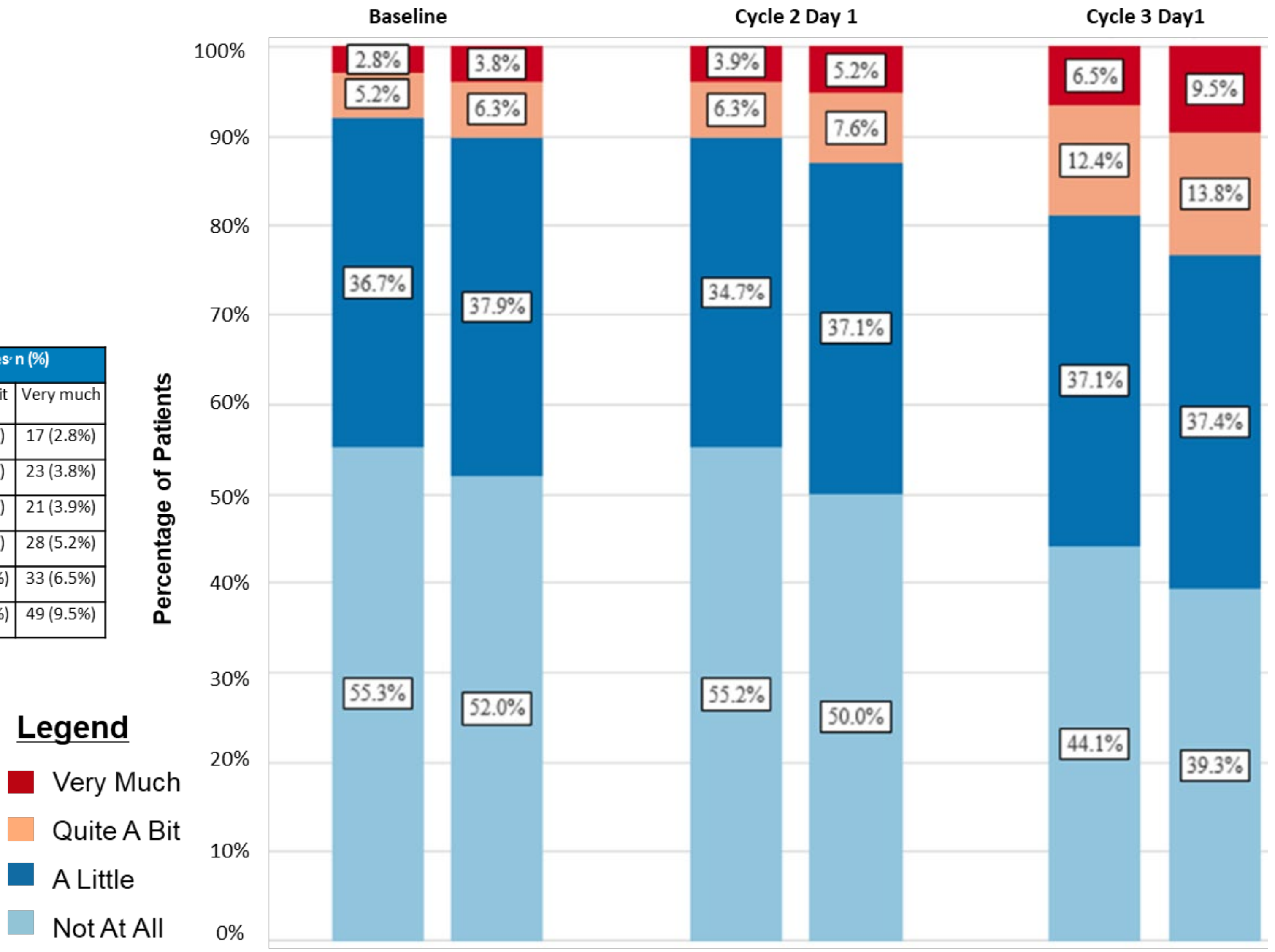


Distribution of Categorical Response

- Percentage in each category uses PRO completed as denominator

Analysis Visit	Treatment Arm	PRO Expected	PRO Completed, n (%)	PRO Not Completed, n (%)	Response Categories 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%)

Safety and Tolerability Example

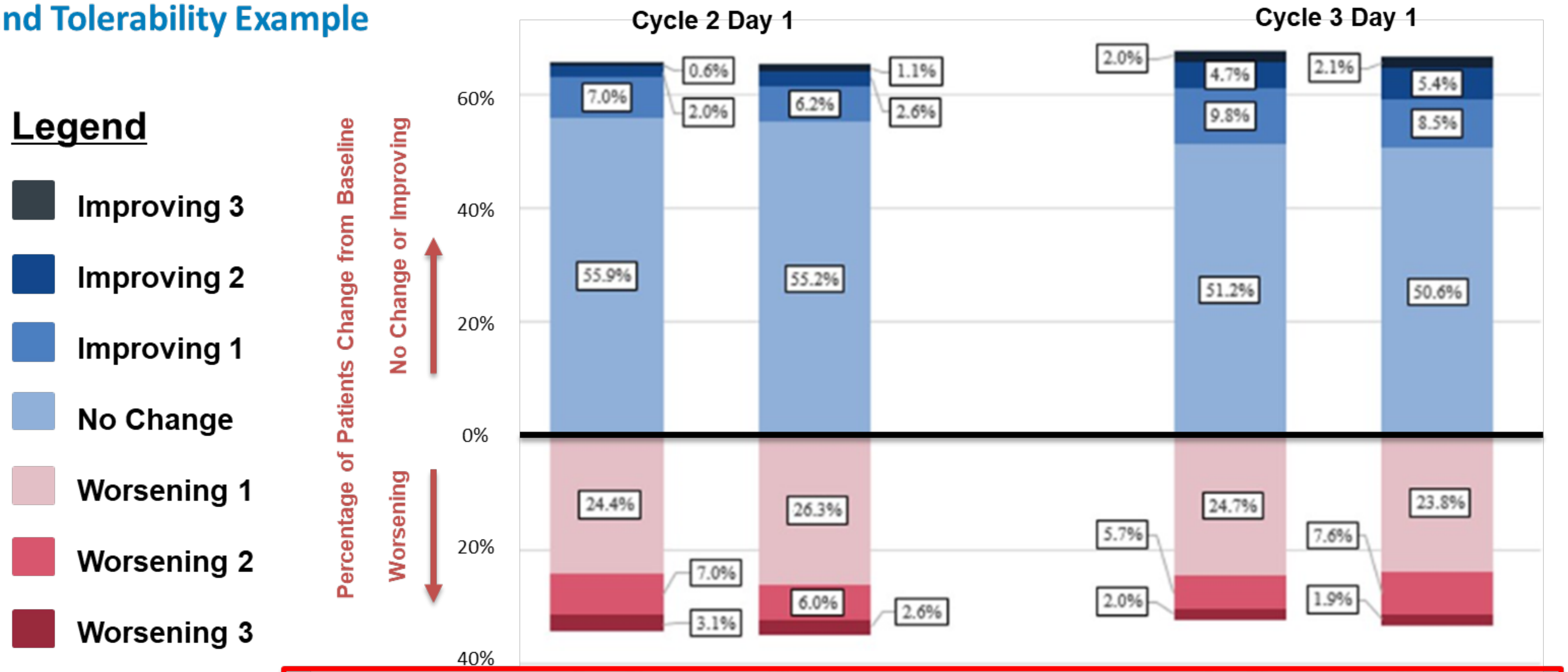


	Control	Treatment	Control	Treatment	Control	Treatment
PRO Completed (n)	600	602	542	536	510	516
PRO Not Completed (n)	0	0	22	36	15	26
PRO Expected (n)	600	602	564	572	525	542

Figures should include number of people in the sample (n)

Distribution of Change from Baseline: Categorical Response

Safety and Tolerability Example



	Control	Treatment	Control	Treatment
PRO Completed (n)	542	536	510	516
PRO Not Completed (n)	22	36	15	26
PRO Expected (n)	564	572	525	542

Conclusions

- PRO data collection and analysis population(s) depend on the PRO objective
  - Clinical benefit: All randomized patients (available data rate)
  - Safety and tolerability: PRO expected (completion rate)

- ADaM datasets should contain rows representing missing PRO data at each time point and include key indicator variables facilitating analysis and interpretation of the data

- Standardized tables and figures provide context and a broader picture of patient experience for FDA's benefit-risk assessment of anti-cancer therapies

Acknowledgements

I would like to extend thanks to the following for their contributions toward the development of the technical specifications document and this presentation

- Mallorie Fiero
- Ting-Yu (Jeff) Chen
- Vishal Bhatnagar
- Laura Lee Johnson

Use of Available Code Packages and Resources

Ki Heun Chung
Division of Analytics and Informatics
CDER/OTS/Office of Biostatistics

Overview

- **Why it is important (Planning)**
  - Statistical Analysis Plan
- **How it may affect review timeline (Execution)**
  - Statistical Programming Guidance

Statistical Software Clarification

- FDA does not require use of any specific software for statistical analyses
- Statistical software is not explicitly discussed in Title 21 of the Code of Federal Regulations [e.g., in 21CFR part 11]
- However, the software package(s) used for statistical analyses should be fully documented in the submission, including version and build identification

- Statistical Software Clarifying Statement

E9 Statistical Principle for Clinical Trials

- All important details should be clearly specified in a protocol written before the trial begins
  - Many of the more detailed aspects of presentation and tabulation should be finalized at or about the time of the blind review so that, by the time of the actual analysis, full plans exist for all its aspects
- The extent to which the procedures in the protocol are followed and is planned *a priori* will contribute to the degree of confidence in the final results and conclusions of the trial
- Documentation of appropriate software testing procedures should be available
- Any deviations from the pre-specified plan must be clearly identified and justified
- Formal records should be kept of when the statistical analysis plan (SAP) was finalized as well as when the blind was subsequently broken

- E9 Statistical Principles for Clinical Trials | FDA

Pre-specification

- Regulatory agencies require evidence that analyses were planned rather than exploratory
  - A proper rationale should be provided for estimation and hypothesis testing
- Pre-specification is fundamental to the credibility and interpretability of clinical trial results
  - An accurate execution should be provided for the planned analysis
  - A careful explanation should be provided for deviations from the planned analysis

Using Available Code Packages and Resources

- The agency focuses more on the appropriateness of statistical methods than the specific software used
- FDA statisticians often use multiple software concurrently for a review themselves
- Available resources may not correctly reflect the planned analysis
  - Leads to unintended/undetected deviation from the protocol, jeopardizing the analysis and delaying the review process

Disparity in Implementation

- Different approaches, both “correct” and under vaguely “pre-specified SAP"
- Different approaches, answering different or inappropriate question, but still under vaguely “pre-specified SAP"
- Wrongly done, not reflecting SAP properly

Code Validation

- Validation documentation
- Testing core statistical functions
- Documenting version control and maintaining change logs
- Appropriate publication references, i.e., theory foundation

- Consider the validation of a tool's dependencies as well
  - Particularly open-source resources

Consideration

- Confirm if an estimate is truly answering the right question in a right manner
  - Confirm if a numerical result is truly an estimate planned to be used
- Tools may be grouped under same name and purpose, but differ in their numeric result due to
  - Answering different hypothesis test under same model
  - Estimating in a different computational way
  - Implementation of different theoretical approach
- Be wary of “default” values/parameters and what is lurking underneath
  - It may not be appropriate
  - It may change in the next update
  - Complexity increases as more dependencies are introduced
  - Even author may not know
- Some may be trivial or negligible, whereas others may constitute a serious analytical deficiencies

Conclusion

- Reproducibility of an analysis is a foundation of good science
- Proper pre-specification is crucial for the conclusion of trial
  - Small ambiguity can lead to big disagreement in how analysis was to be conducted, possibly lengthening review process
- Rigorous execution of planned analysis is crucial for timely review
  - Available resources may not correctly reflect the planned analysis
  - Submissions should include source codes to generate analyses dataset and TLF that can be inspected and executed by reviewers.
    - macro/external dataset/etc.
    - importance of standardized data
  - Do not rely on “default ways” of a particular implementation to be appropriate for every submission
    - intention should be clear to all parties

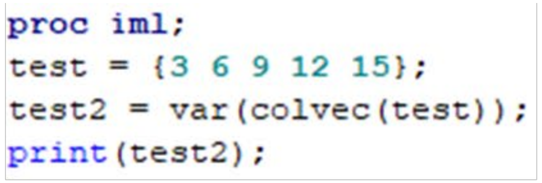
Statistical Analysis Plan (SAP)

- E9 Statistical Principles for Clinical Trials | FDA
  - *all important details (…) should be clearly specified in a protocol written before the trial begins.*
  - *extent to which (...) are followed and (…) is planned a priori will contribute to the (…) conclusions of the trial.*
  - *documentation of appropriate software testing procedures should be available.*
- *Pre-specification of statistical analysis approaches in published clinical trial protocols was inadequate* (Greenberg et al., 2018)
  - “No protocols adequately predefined all (…) aspects of the analysis.”
- *Evidence of unexplained discrepancies between planned and conducted statistical analyses: a review of randomised trials.* (Cro et al., 2020)
  - “… unexplained discrepancies in the statistical methods of randomised trials are common.”

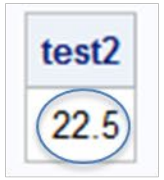
Statistical Software Discrepancies

- “Note that SAS […] uses the integral up to the last *event* time of each individual curve; we […] do not provide an option for that calculation.”
  - Dr. Therneau, on the summary of survival curve in `survival` package documentation
- “When SAS (or Stata, or Genstat/AS-REML or …) and R differ in their answers, R may not be wrong. Both SAS and R may be `right’ but proceeding in a different way/answering different questions/using a different philosophical approach (or both may be wrong …)”
  - Dr. Ben Bolker | lme4 FAQ

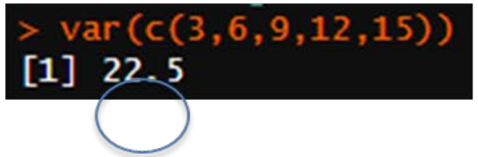
Statistical Software - Variance



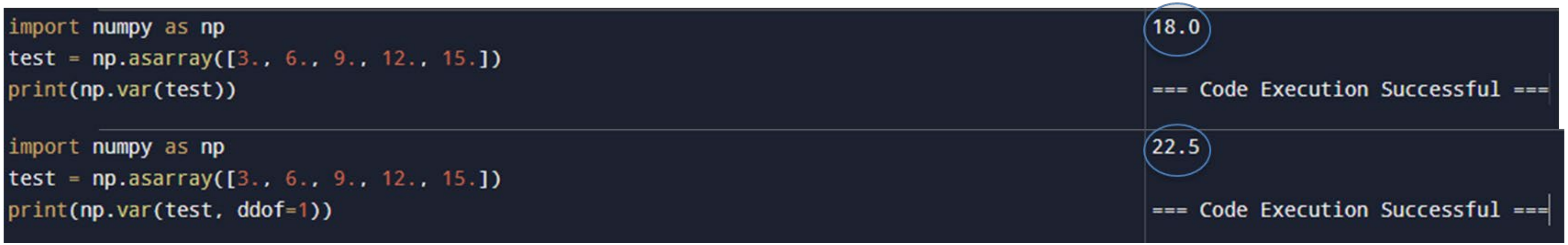
```
proc iml;
test = {3 6 9 12 15};
test2 = var(colvec(test));
print(test2);
```



```
test2
22.5
```



```
> var(c(3,6,9,12,15))
[1] 22.5
```



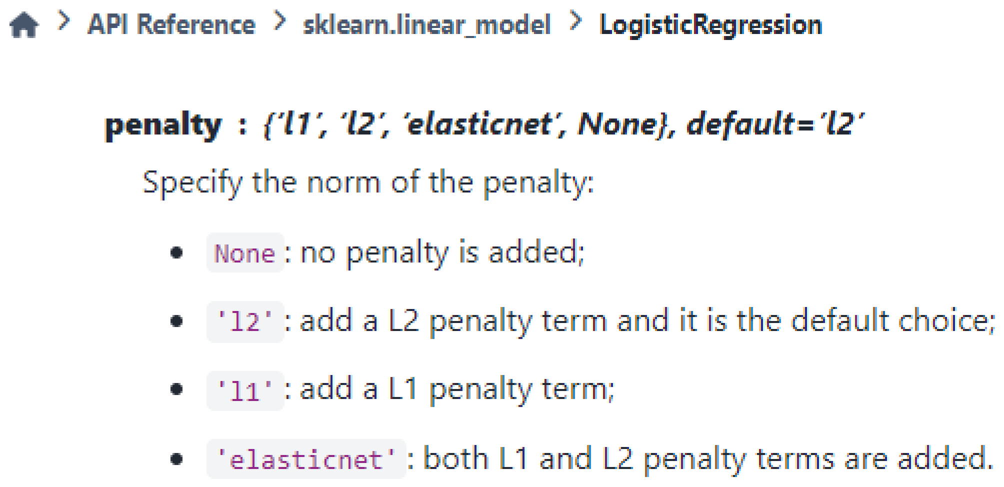
```
import numpy as np
test = np.asarray([3., 6., 9., 12., 15.])
print(np.var(test))
```

```
18.0

=== Code Execution Successful ===
```

```
import numpy as np
test = np.asarray([3., 6., 9., 12., 15.])
print(np.var(test, ddof=1))
```

```
22.5

=== Code Execution Successful ===
```

Statistical Software – Logistic Regression



API Reference > sklearn.linear_model > LogisticRegression

penalty : ***{'l1', 'l2', 'elasticnet', None}, default='l2'***

Specify the norm of the penalty:

- `None` : no penalty is added;
- `'l2'` : add a L2 penalty term and it is the default choice;
- `'l1'` : add a L1 penalty term;
- `'elasticnet'` : both L1 and L2 penalty terms are added.

How many millions of ML/stats/data-mining papers have been written by authors who didn't report (& honestly didn't think they were) using regularization?

- [Dr. Zachary Lipton, Carnegie Mellon University]

Potential Keywords of Ambiguity

- “robust”
- “bootstrap”
- “multiple imputation”
- “simulation”
- “wrapper function”
  - Bayesian (jags, etc.) / Large Data (spark, etc.) / ML (pytorch, etc.) / Visualization (Schoenfeld plot, etc.)
- “optimization”
- “adjusted”
- “contrast”
- “estimator”
- “degrees of freedom”
- “approach”

Panel Discussion

Closing Remarks

Robyn Bent, RN, MS

Director, Patient-Focused Drug Development Program
Office of Center Director
Center for Drug Evaluation and Research
U.S. Food and Drug Administration

Thank you!