Clinical Outcome Assessments in Cancer Trials

Paul G Kluetz, MD
Associate Director of Patient Outcomes
Oncology Center of Excellence, U.S. FDA

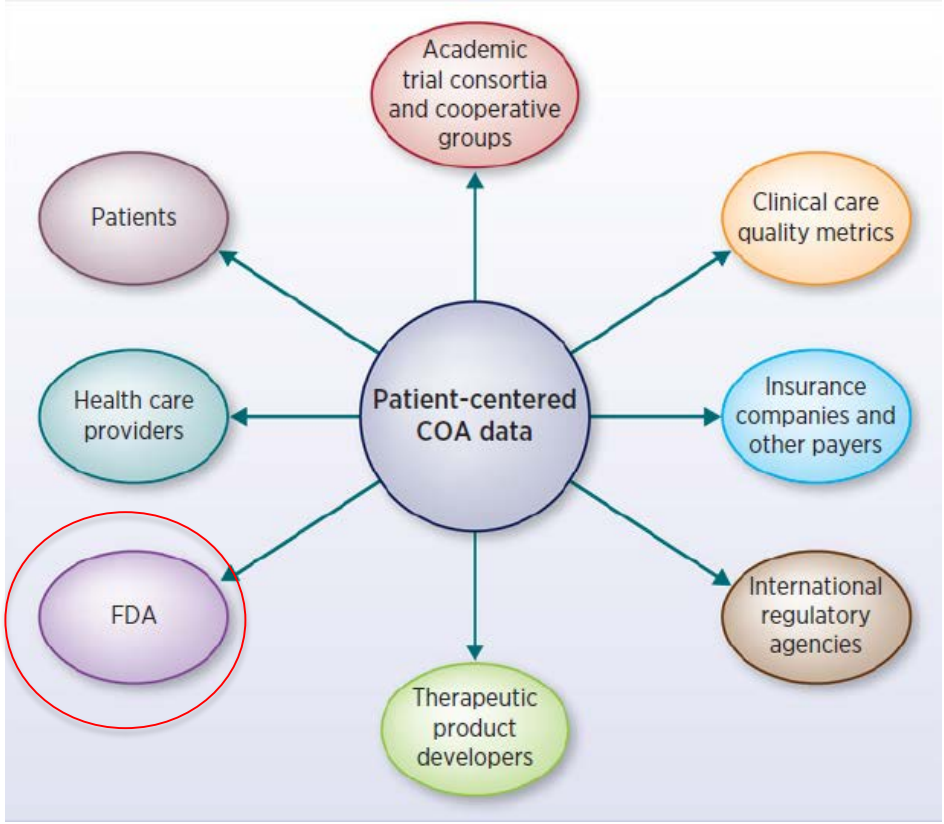
August 22, 2018 MEDCAC Meeting

Disclosures/Disclaimers

- I have no financial interests to disclose
- Specific PRO measurement tools are used as examples, and are not endorsements
- Consult with FDA for advice on PRO measurement in your specific trial

U.S. Regulatory Perspective:

- One of many stakeholders interested in assessing clinical outcomes



Outline

- Introduction- What is a Clinical Outcome?
- Core Concepts- What concepts to measure?
- Measurement Tools- How can we measure them?
- Endpoints- What questions can we ask?

Important Terminology

- **Concept-** an aspect of an individual's clinical, biological, physical or functional state or experience
  - Concepts can be clinical (pain) or non-clinical (pharmacodynamic biomarker)
- **Clinical outcome-** an outcome that describes or reflects how someone feels functions or survives
- **Clinical outcome assessments-** measures of clinical outcomes
  - A Patient-reported outcome is a type of clinical outcome assessment
- **Test, tool or instrument-** assessment system
  - For patient-reported outcomes this is a questionnaire and its scoring manual
- **Endpoint-** a precisely defined variable intended to reflect an outcome of interest that is statistical analyzed to address a particular research question.

Biomarkers, endpoints and other tools resource (BEST) https://www.ncbi.nlm.nih.gov/books/NBK338448/
Glossary 2009 FDA Guidance to Industry- PRO Measures: Use in Medical Product Development to Support Labeling Claims

Terminology Example – Pain as an Endpoint

Concept of Interest- pain

Pain is a ***Clinical Outcome***

Clinical Outcome Assessment is a ***Patient-Reported Outcome (PRO)***

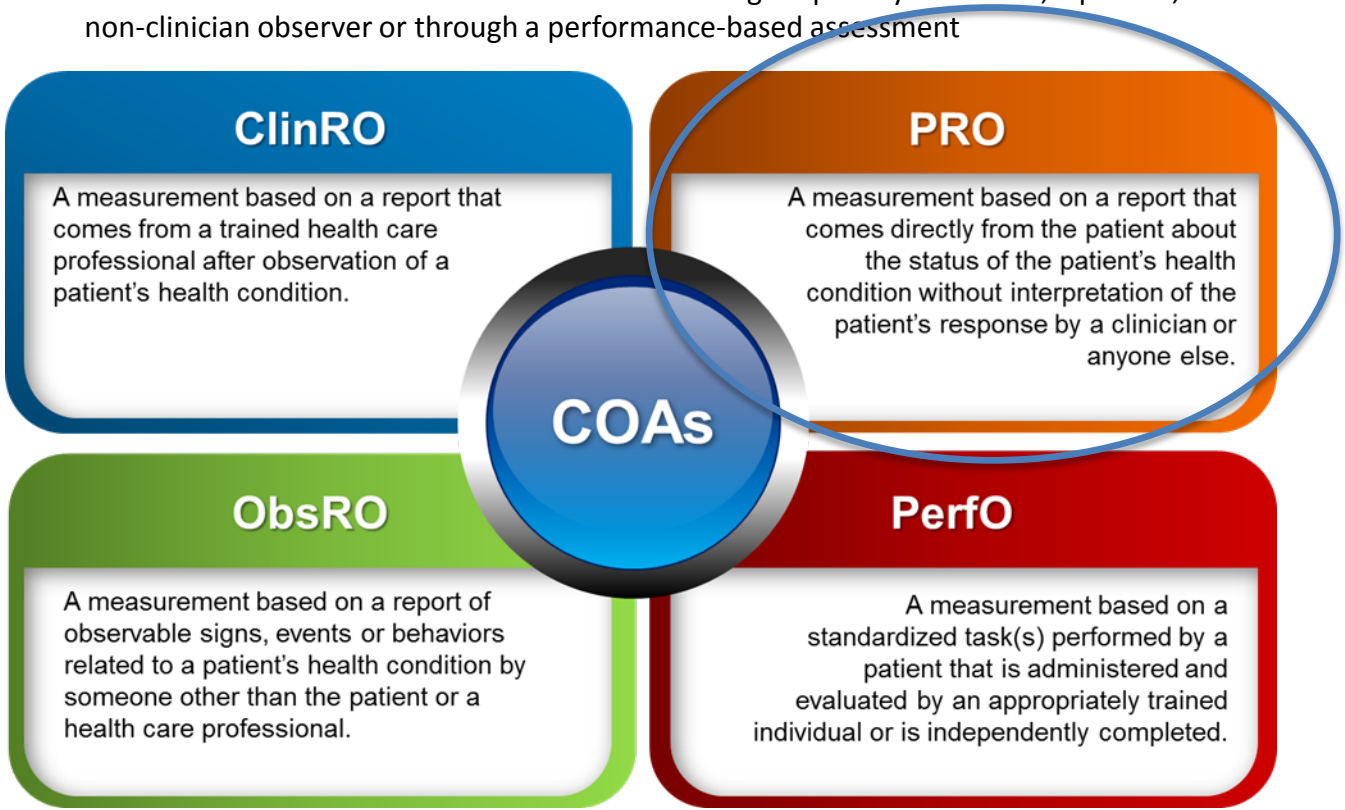
PRO Instrument may be Brief Pain Inventory Short Form (BPI-SF)

Endpoint may be 2 point decrease in item #3 of BPI-SF confirmed by second result at least 2 weeks later with no increase in analgesic use[1]

[1] For illustrative purposes, pain endpoint definitions may vary based on trial context

Clinical Outcome Assessments (COAs)

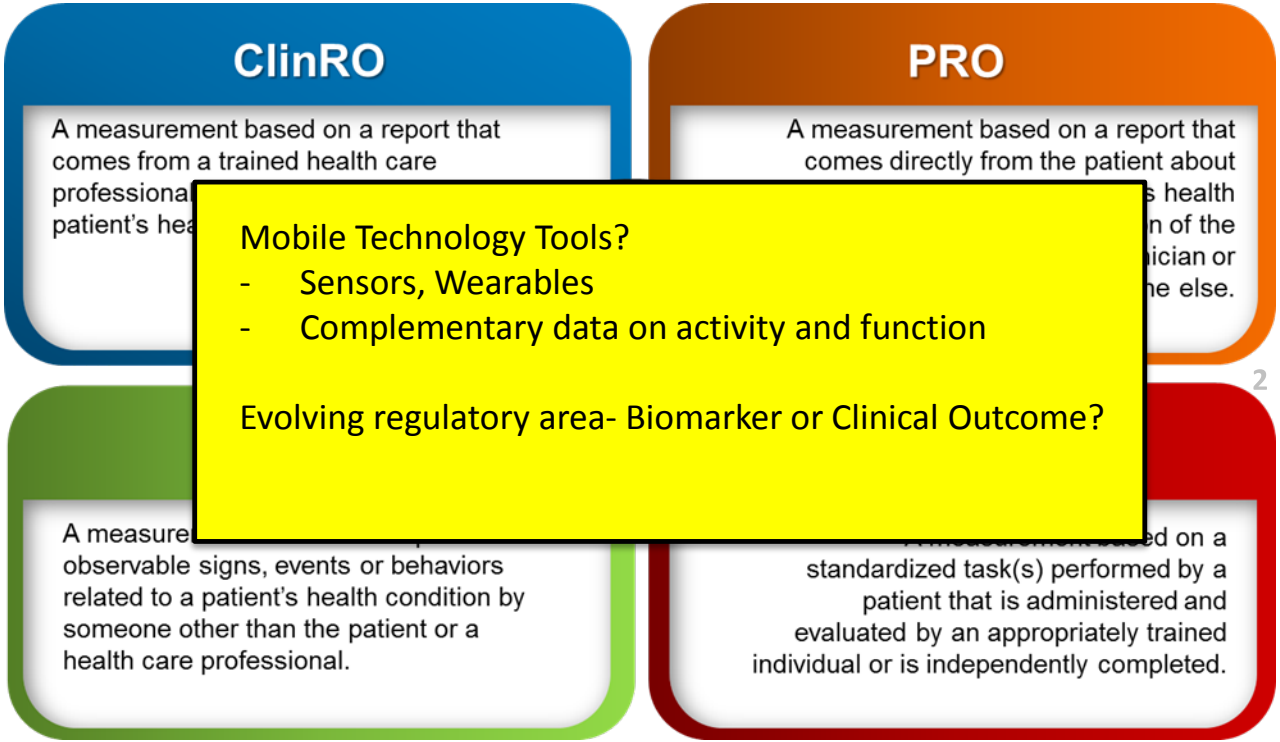
COA: Assessment of a clinical outcome made through report by a clinician, a patient, a non-clinician observer or through a performance-based assessment



Slide Courtesy of Selena Daniels, FDA Clinical Outcome Assessment Staff.

Clinical Outcome Assessments (COAs)

COA: Assessment of a clinical outcome made through report by a clinician, a patient, a non-clinician observer or through a performance-based assessment



Slide Courtesy of Selena Daniels, FDA Clinical Outcome Assessment Staff.

Enthusiasm for Clinical Outcomes: Why now?

- Recent Legislation- PDUFA and 21st Century Cures
- Informed choice in era of multiple therapies
- Technology is improving capabilities
  - ePRO->Structured clinical data = trial efficiencies, pipeline of structured RWD
  - Wearable devices and sensors
- Clinical care increasingly using PRO measures
- Item libraries and a focus on safety and tolerability

Outline

- Introduction- What is a Clinical Outcome?
- **Core Concepts- What concepts to measure?**
- Measurement Tools- How can we measure them?
- Endpoints- What questions can we ask?

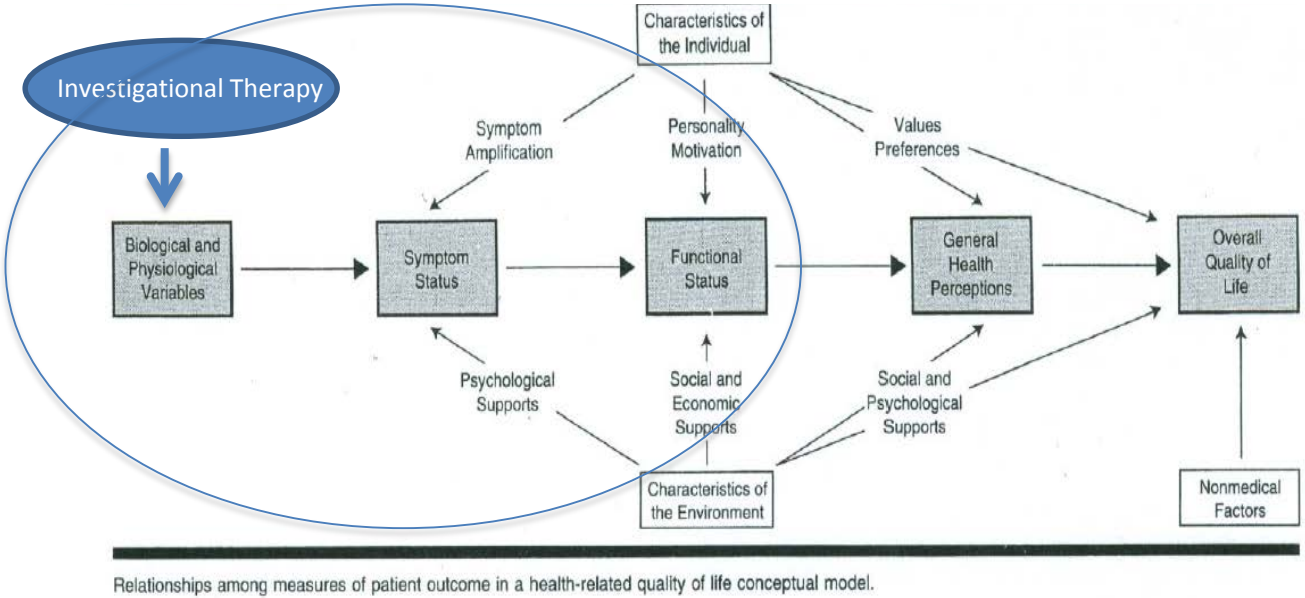
Core Concepts: What should we measure?

- Heterogeneity in PRO concepts assessed has been problematic
- A core set of PRO concepts can create consistency
- Focus should be on isolating the therapeutic effect

Distinguishing the Effect of the Drug: The Challenge of Overall Quality of Life

Proximal to the Drug Effect

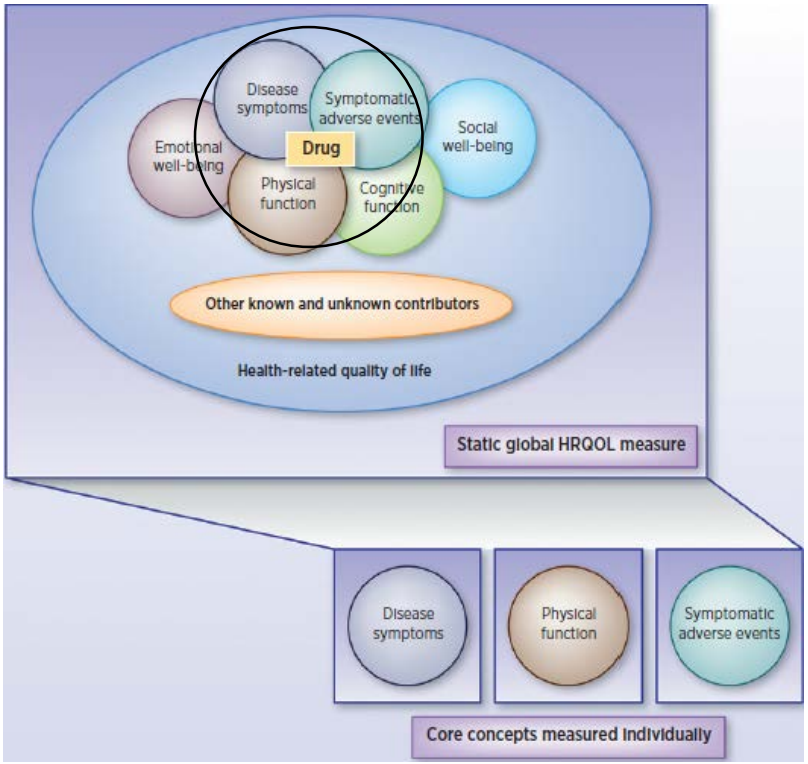
Distal to the Drug Effect



Relationships among measures of patient outcome in a health-related quality of life conceptual model.

Adapted from Wilson et al., Linking Clinical Variables with HRQOL. *JAMA*, 1995, (273): 59-65.

Core PRO symptom and functional concepts proximal to the therapeutic effect



Proximal symptom and functional assessments **favored for drug labeling**

Other aspects of HRQL can still be important to describe the patient experience (but may be more susceptible to bias and non-drug contributors)

Proximal concepts may not be the only PRO data to assess or measure

Focusing on core patient-reported outcomes in cancer clinical trials. *Clin Can Res* 22.7 (2016): 1553-1558.

Outline

- Introduction- What is a Clinical Outcome?
- Core Concepts- What concepts to measure?
- **Measurement Tools- How can we measure them?**
- Endpoints- What questions can we ask?

What is a “Fit For Purpose” PRO Instrument?

1. Appropriate for its intended use
   - Study design, Patient population, Therapy under study
2. Validly and reliably measures concepts that are:
   - Clinically relevant
   - Important to patients
3. Can be communicated in labeling in a way that is accurate, interpretable, and not misleading (i.e., **well-defined**)

What is meant by “well-defined”?

Example: Physical Function (PF) Scale

All questions within the PF domain score are measuring physical function

Two reasonable examples include:

PROMIS® PF

Are you able to do chores such as vacuuming or yard work?
Are you able to go up and down stairs at a normal pace?
Are you able to go for a walk of at least 15 minutes?
Are you able to run errands and shop?

Patient-reported outcomes measurement information system (PROMIS)

EORTC QLQC30 PF Domain

Do you have any trouble doing strenuous activities, like carrying a heavy shopping bag or a suitcase?

Do you have any trouble taking a long walk?

Do you have any trouble taking a short walk outside of the house?

Do you need to stay in bed or a chair during the day?

Do you need help with eating, dressing, washing yourself or using the toilet?

European Organisation for Research and Treatment of Cancer (EORTC)

*Examples for illustration- consult with FDA regarding the suitability of your PRO strategy

What is meant by “well-defined”?

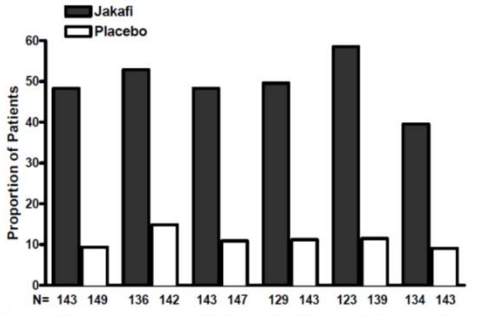
Disease Symptom Score

All questions within the symptom scale are measuring symptoms of the disease (to the extent feasible).

One reasonable example is the Myelofibrosis Symptom Assessment Form

- Abdominal Discomfort
- Pain under left ribs
- Early satiety
- Night sweats
- Itching
- Bone or Muscle Pain

Figure 3: Proportion of Patients With 50% or Greater Reduction in Individual Symptom Scores at Week 24



Individual score range = 0 to 10

Source: FDA product label, ruxolitinib- https://www.accessdata.fda.gov/drugsatfda_docs/label/2011/202192lbl.pdf

What is NOT well-defined?

- A score where one or more items are not measuring the concept of interest

- Hypothetical example: A fictitious **Fatigue Score**
  - How tired are you?
  - How much weakness do you have?
  - What is your energy level
  - What level of pain do you have?
  - How much numbness and tingling do you have?
  - How would you rate your quality of life?

What is NOT well-defined?

- Hypothetical example: A fictitious **Fatigue Score**
  - How tired are you?
  - How much weakness do you have?
  - What is your energy level
  - What level of pain do you have?
  - How much numbness and tingling do you have?
  - How would you rate your quality of life?

While pain and neuropathy may contribute to one's fatigue, they are not components of fatigue itself. While HRQL may be impacted by fatigue, it is not a component of fatigue.

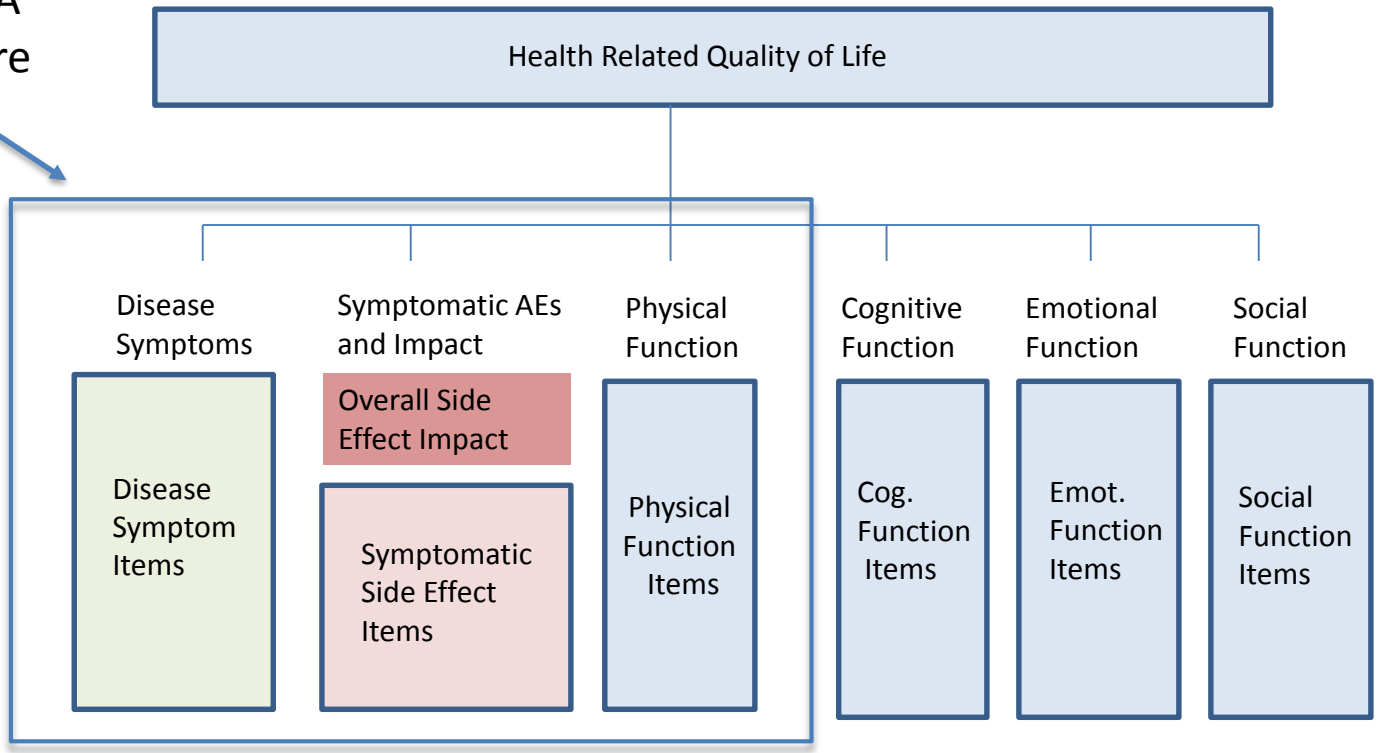
Reasonable PRO strategy for typical cancer trial?

"We favor a thoughtful combination of **static questionnaires** and **item banks** or **libraries** to create a balanced, flexible, and modular approach to PRO assessment to address targeted trial objectives and accommodate the needs of multiple stakeholders."

"Focusing on Core Patient-Reported Outcomes in Cancer Clinical Trials—Response." Clinical Cancer Research 22.22 (2016): 5618-5618.

Modular Approach- Data for all Stakeholders

Proposed FDA Oncology Core Concepts



Example of Modular Instrument:

- EORTC QLQC30
- One example of a “Modular” health-related quality of life instrument:
- Each scale/item is well-defined and can form their own endpoint

	Scale	Number of items	Item range*	**Version 3.0** Item numbers	Function scales
Global health status / QoL					
Global health status/QoL (revised)†	QL2	2	6	29, 30	
Functional scales					
Physical functioning (revised)†	PF2	5	3	1 to 5	F
Role functioning (revised)†	RF2	2	3	6, 7	F
Emotional functioning	EF	4	3	21 to 24	F
Cognitive functioning	CF	2	3	20, 25	F
Social functioning	SF	2	3	26, 27	F
Symptom scales / items					
Fatigue	FA	3	3	10, 12, 18	
Nausea and vomiting	NV	2	3	14, 15	
Pain	PA	2	3	9, 19	
Dyspnoea	DY	1	3	8	
Insomnia	SL	1	3	11	
Appetite loss	AP	1	3	13	
Constipation	CO	1	3	16	
Diarrhoea	DI	1	3	17	
Financial difficulties	FI	1	3	28	

* *Item range* is the difference between the possible maximum and the minimum response to individual items; most items take values from 1 to 4, giving *range* = 3.

† (revised) scales are those that have been changed since version 1.0, and their short names are indicated in this manual by a suffix “2” – for example, PF2.

For illustrative purposes, consult with FDA on appropriateness of your PRO strategy in your trial context

Figure used with permission- EORTC User Manual https://www.eortc.be/qol/files/SCManualQLQ-C30.pdf

Example of a modular measurement system

PRO Measurement Information System

- National Institutes of Health (NIH)
- Multiple short form scales
- Each scale/item should be well-defined and can form an endpoint
- Scales are flexible and can be adapted to trial context

http://www.healthmeasures.net/explore-measurement-systems/promis/intro-to-promis/list-of-adult-measures

*For illustrative purposes, consult with FDA on appropriateness of your PRO strategy in your trial context

Example of a modular approach: Item Library

- National Cancer Institute's PROCTCAE
- Library of 78 Symptomatic Adverse Events
- Can select relevant symptoms for expected therapeutic toxicity profile

PATIENT-REPORTED OUTCOMES VERSION OF THE COMMON TERMINOLOGY CRITERIA FOR ADVERSE EVENTS (PRO-CTCAE™) ITEM LIBRARY (Version 1.0)

Oral	
Dry mouth	S
Difficulty swallowing	S
Mouth/throat sores	SI
Cracking at the corners of the mouth (cheilosis/cheilitis)	S
Voice quality changes	P
Hoarseness	S

Gastrointestinal	
Taste changes	S
Decreased appetite	SI
Nausea	FS
Vomiting	FS
Heartburn	FS
Gas	P
Bloating	FS
Hiccups	FS
Constipation	S
Diarrhea	F
Abdominal pain	FSI
Fecal incontinence	FI

Respiratory	
Shortness of breath	SI
Cough	SI
Wheezing	S

Cardio/Circulatory	
Swelling	FSI
Heart palpitations	FS

Cutaneous	
Rash	P
Skin dryness	S
Acne	S
Hair loss	A
Itching	S
Hives	P
Hand-foot syndrome	S
Nail loss	P
Nail ridging	P
Nail discoloration	P
Sensitivity to sunlight	P
Bed/pressure sores	P
Radiation skin reaction	S
Skin darkening	P
Stretch marks	P

Neurological	
Numbness & tingling	SI
Dizziness	SI

Visual/Perceptual	
Blurred vision	SI
Flashing lights	P
Visual floaters	P
Watery eyes	SI
Ringing in ears	S

Attention/Memory	
Concentration	SI
Memory	SI

Pain	
General pain	FSI
Headache	FSI
Muscle pain	FSI
Joint pain	FSI

Sleep/Wake	
Insomnia	SI
Fatigue	SI

Mood	
Anxious	FSI
Discouraged	FSI
Sad	FSI

Gynecologic/Urinary	
Irregular periods/vaginal bleeding	P
Missed expected menstrual period	P
Vaginal discharge	A
Vaginal dryness	S
Painful urination	S
Urinary urgency	FI
Urinary frequency	FI
Change in usual urine color	P
Urinary incontinence	FI

Sexual	
Achieve and maintain erection	S
Ejaculation	F
Decreased libido	S
Delayed orgasm	P
Unable to have orgasm	P
Pain w/sexual intercourse	S

Miscellaneous	
Breast swelling and tenderness	S
Bruising	P
Chills	FS
Increased sweating	FS
Decreased sweating	P
Hot flashes	FS
Nosebleed	FS
Pain and swelling at injection site	P
Body odor	S

Attributes	
F: Frequency	I: Interference
S: Severity	P: Presence/Absence
A: Amount	

https://healthcaredelivery.cancer.gov/pro-ctcae/measurement.html

Version date: 6/22/2018

*For illustrative purposes, consult with FDA on appropriateness of your PRO strategy in your trial context

Outline

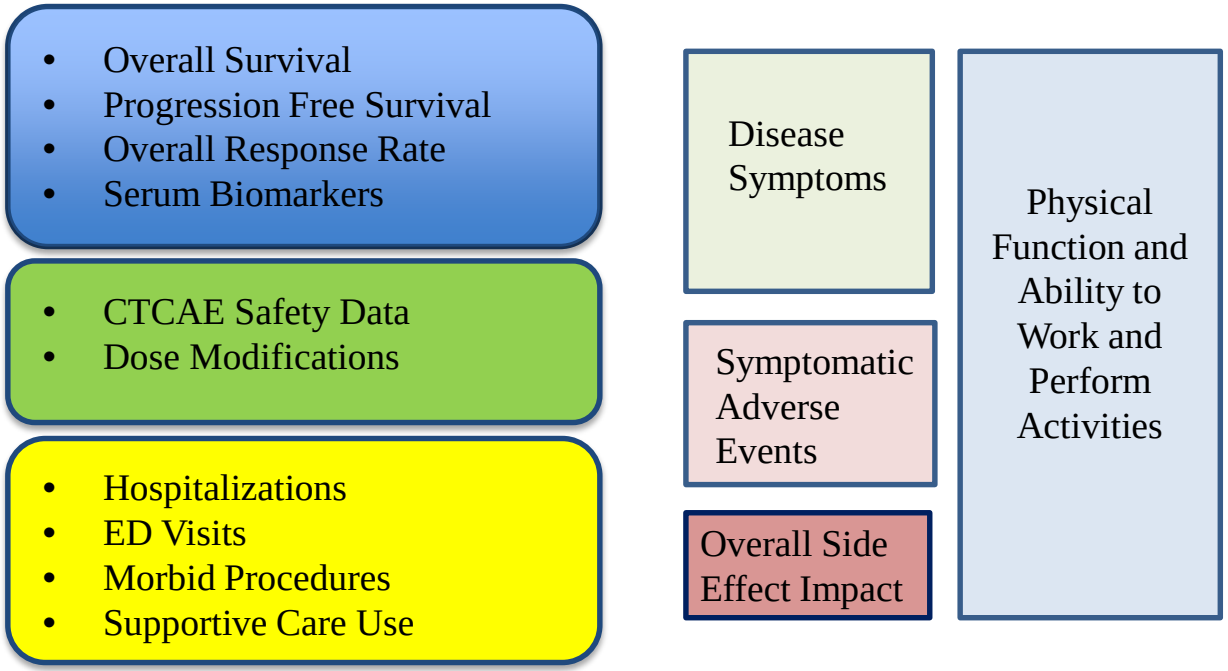
- Introduction- What is a Clinical Outcome?
- Core Outcomes- What concepts to measure?
- Measurement Tools- How can we measure them?
- **Concepts to Instruments to Endpoints**

Endpoint- a precisely defined variable intended to reflect an outcome of interest that is statistical analyzed **to address a particular research question**.

What is the research question?

- Safety and Tolerability
  - Complement standard safety assessment to inform tolerability
  - Longitudinal symptomatic adverse event data
  - Overall side effect impact
- Efficacy
  - Complement standard tumor and survival efficacy findings
  - Improvement in core disease symptoms or function

Important Data Elements in Cancer Trials



Standard Measurement and Analysis Needed

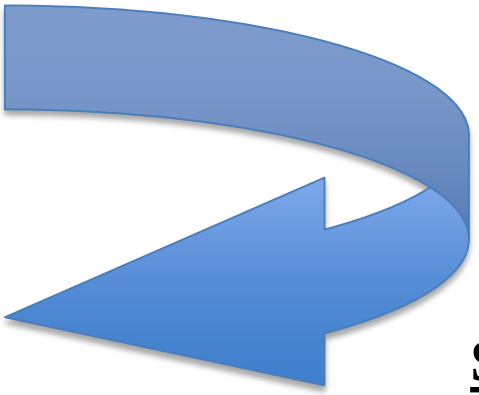
Outcomes	Measure	Standardization
Overall Survival	OS	High
Tumor Progression	TTP, PFS, DFS	High (RECIST)
Tumor Shrinkage	ORR	High (RECIST)
Adverse Events	Clin-RO	High (CTCAE)
Symptomatic Adverse Events	PRO	Low*
Overall Side Effect Impact	PRO	Low*
Physical Function	PRO	Low*
Disease Symptoms	PRO	Low*

TTP- Time to Progression, PFS- Progression-free survival, DFS- Disease-free survival, ORR- Objective response rate, Clin-RO- Clinician-reported outcome, PRO- Patient-reported outcome, HRQL- Health-related quality of life

* Standardization with respect to measurement tool, endpoint definition and analysis methods

Future FDA Oncology Work Will Focus on PRO Objectives/Endpoints/Analytics

Identify and Expect Core PRO Concepts



Use Well Defined Scales and Item Libraries

Standard Objectives / Endpoints / Analytics

Example: Concept to Instrument to Endpoint

Concept of Interest- Disease symptoms

Instrument- Myelofibrosis symptom assessment form

Endpoint- 50% or greater reduction in symptom score by week 24

Table 8: Improvement in Total Symptom Score

	Jakafi (N=148)	Placebo (N=152)
Number (%) of Patients with 50% or Greater Reduction in Total Symptom Score by Week 24	68 (45.9)	8 (5.3)
P-value	< 0.0001	

FDA Label for Ruxolitinib https://www.accessdata.fda.gov/

Example: Concept to Instrument to Endpoint

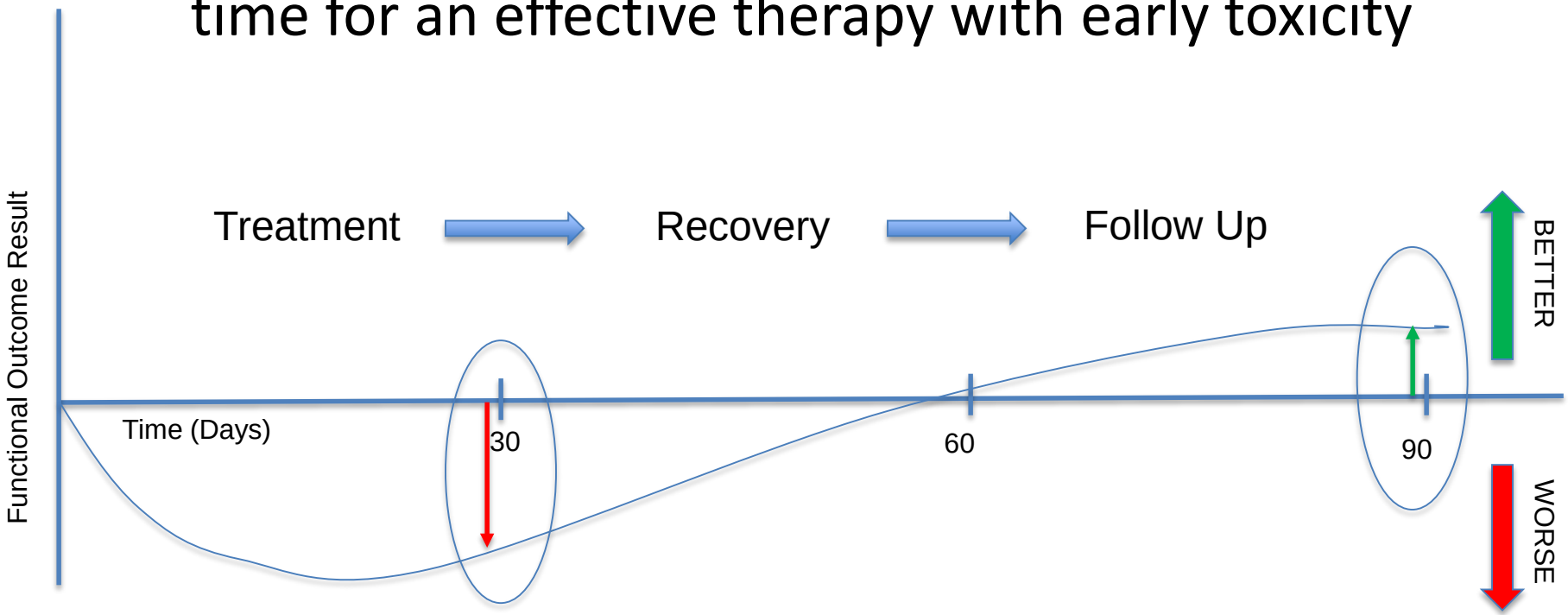
Concept of Interest- Physical Function

Instrument- PROMIS 10a Physical Function Short Form

Endpoint- Change from baseline in physical function at 12 months in the ITT population

* For illustrative purposes, consult with FDA on appropriateness of your PRO strategy in your trial context

Fictitious example of physical function over time for an effective therapy with early toxicity



Opposite conclusions are reached if endpoint is 30 day change from baseline versus 90 day change from baseline...

PRO information must be considered within the totality of data and the treatment context

- Adjuvant / Curative
  - Short term significant toxicity BUT
  - Defined treatment duration
  - Potential for cure
- Palliative / Non-Curative
  - Treatment until progression
  - Cumulative toxicity
  - Benefit typically limited to time on therapy

Conclusion

- Clinical outcomes can complement (not replace) survival, tumor and safety measures in cancer trials; PRO is one type of clinical outcome
- PRO Functional scales and symptom scales should be well-defined to allow for clear communication to patients and providers in labeling
- Several PRO measurement systems exist, consult with FDA to assure they are fit for purpose for your trial
- It is CRITICAL to understand your treatment context and carefully consider the research objective to inform endpoint definition