WORKSHOP ON
CLINICAL OUTCOME ASSESSMENTS (COAs)
IN CANCER CLINICAL TRIALS

April 26, 2016  ■  Silver Spring, MD

Co-sponsored by

[Logos for FDA and Critical Path Institute]
Session 2
Using Multiple Instruments to Create a Comprehensive PRO Assessment Strategy in Cancer Trials

WORKSHOP ON CLINICAL OUTCOME ASSESSMENTS (COAs) IN CANCER CLINICAL TRIALS

April 26, 2016 ▪ Silver Spring, MD

Co-sponsored by

FDA
CRITICAL PATH INSTITUTE
Disclaimer

• The views and opinions expressed in the following slides are those of the individual presenters and should not be attributed to their respective organizations/companies, the U.S. Food and Drug Administration or the Critical Path Institute.

• These slides are the intellectual property of the individual presenters and are protected under the copyright laws of the United States of America and other countries. Used by permission. All rights reserved. All trademarks are the property of their respective owners.
Session Participants

Chair
– Stephen Joel Coons, PhD – C-Path

Presenters
– Paul G. Kluetz, MD - FDA
– Patty Spears – Cancer Information and Support Network
– Alicyn Campbell, MPH – Genentech
– Charles S. Cleeland, PhD – M.D. Anderson Cancer Center
– Sandra A. Mitchell, PhD – NCI, NIH
– Andrew Bottomley, PhD – EORTC
– David Cella, PhD – Northwestern University Feinberg School of Medicine

Panelists
– Elektra Papadopoulos, MD, MPH – FDA
– Jeff A. Sloan, PhD – Mayo Clinic
Incorporation of Multiple PRO Instruments in Clinical Trials

Paul G. Kluetz
Office of Hematology and Oncology Products
U.S. Food and Drug Administration
• This talk represents current thinking in an evolving area of scientific and health care policy

• The views expressed do not necessarily reflect the official position of the U.S. Food and Drug Administration
Session 1 Review: FDA Use of PRO Data

• FDA **reviews** all submitted PRO data as important supportive data in the overall benefit:risk assessment
  – All PRO data should be assessed rigorously with this in mind
  – Proximal symptom/function is the current focus for product labeling
  – Safety/tolerability relevant objective for PRO measures in many contexts

• To be considered for product labeling:
  – PRO assessments must be well-defined and reliable
  – Claim of treatment benefit should be tested in statistical hierarchy
  – **Goal: Provide informative, interpretable patient-centered data to further inform the safety and efficacy of an anti-cancer therapy**
Opportunities and Challenges for PRO Measurement in Contemporary Cancer Trials

• Sustained commitment to research in biology, immunology and genetics = **unprecedented treatment effects**
  – FDA programs used to expedite deliver of cancer therapies to patients
  – More single arm trial submissions
  – Randomized trials commonly open-label

• Multiple different mechanistic classes of agents = more diverse symptomatic side effect profiles

• **Opportunity: Expanding options for PRO measurement**
Expanding PRO Toolbox

• **FACIT Measurement System:**
  – [http://www.facit.org/FACITOrg/Questionnaires](http://www.facit.org/FACITOrg/Questionnaires)
  – FACT-G, Cancer Specific Measures, Cancer Specific Symptom Indexes, Treatment Specific Measures

• **EORTC Quality of Life Questionnaires:**
  – EORTC-QLQ-C30 and Disease Specific Modules

• **MD Anderson Symptom Inventory (MDASI)**
  – [https://www.mdanderson.org/education-and-research](https://www.mdanderson.org/education-and-research)
  – MDASI Core Symptom Inventory, Disease Modules, BPI-SF, BFI-SF

• **PROMIS® Tools**
  – [http://www.nihpromis.org](http://www.nihpromis.org)
  – Measures of physical, mental and social well-being and global health

• **Patient-Reported Outcome version of CTCAE (PRO-CTCAE)**
  – Item library of symptomatic adverse events

• Many other disease specific and symptom specific measures that are completed or in development
Assessment of Symptomatic Adverse Events

• The assessment of safety and tolerability is important across clinical trial contexts, including single arm trials

• There is interest in exploring the NCI PRO-CTCAE symptomatic adverse event item library
  – Systematically incorporate the patient voice into safety assessments
  – Complement clinician reported safety findings
Patient-Reported Outcomes version of the Common Terminology Criteria for Adverse Events (PRO-CTCAE)

• **Strengths**
  – Patient-reported (Standard oncology safety assessment is Clinician Reported)
  – Systematically and rigorously developed
  – Standardized: provides a standard item library and platform
  – Flexible: Select a subset of the 78 items to create a study-specific survey
  – Complements existing safety evaluation (CTCAE)

• **Early in Implementation: Much Work to be Done**
  – Translations and Cultural Adaptations
  – Scoring
  – Unbiased Item Selection
  – Most informative method to analyze and present data

For more information and permission to use PRO-CTCAE, visit: [http://healthcaredelivery.cancer.gov/pro-ctcae/](http://healthcaredelivery.cancer.gov/pro-ctcae/)
What is the Standard PRO Assessment Strategy?

- For registration trials, we typically receive a 27-30 item Health-Related Quality of Life (HRQoL) instrument

- A Disease Module is typically added

- Current Item Burden: In general, approximately 40-50 questions are asked of patients at various assessment frequencies.

- EQ-5D commonly submitted (5 items and a visual analogue scale)
Respondent Burden

• Assessment frequencies must be frequent enough (especially early in treatment) for the clinical context
  – In the advanced/metastatic cancer setting, patients may be off-study for progression at 2-3 months

• There is a concern for increased respondent burden and duplication with the integration of additional instruments to standard PRO assessment approaches
Adding Tools to the Standard PRO Assessment Strategy- PRO-CTCAE

• Minor duplication occurs with common strategy of combining HRQoL assessments, disease module, and EQ-5D

• Addition of PRO-CTCAE symptomatic adverse event assessment can add 20-30+ additional questions
  – May be at higher assessment frequency than the remainder of PRO assessments in a trial
  – Likely additional duplication with treatment symptom items from disease modules and HRQoL instruments
Adding Tools to the Standard PRO Assessment Strategy- Other Measures of Interest

- Burden when adding new tools is not unique to PRO-CTCAE. Trial objectives may call for a more detailed assessment of other patient-reported symptom and function measures
  - Pain
  - Fatigue
  - Physical Function
  - Other symptom or functional outcome assessments unique to the disease and therapeutic context (e.g. swallowing in esophageal cancer, cognitive function in neuro-oncology)
Session 2 Overview:

• PRO objectives are different depending on the type of trial. We aren’t looking to identify a specific maximum number of questions per trial…

• Session 2 is about opportunities and challenges associated with integrating multiple PRO assessments in cancer trials

• **Goal of a PRO assessment strategy:** Informative, relevant and rigorous patient-centered data while managing burden
Incorporation of Multiple PRO Tools in a Single Trial:

- Rigorous PRO data describing the patient experience is an expectation for multiple stakeholders who make treatment, regulatory and health policy decisions

- Discuss the issue of **respondent burden**
  - What are considerations for addressing respondent burden?

- Discuss the issue of **duplication**
  - What are considerations for addressing duplication?
Why PROs?

PROs Consider:
Not only **WHAT IS THE MATTER** with the patient
But also **WHAT MATTERS** to the patient

— Sandra Finestone

“A PRO is any report of the **status of a patient’s health condition** that comes directly from the patient, without interpretation of the patient’s response by a clinician or anyone else.”

Assessing harms and benefits are very subjective and needs to be done consistently and reliably to actually predict benefits and harms that **ARE IMPORTANT** to patients.
Aid in Decision Making: Benefits vs. Harms

Once a drug is approved, patients are faced with treatment decisions.

Decision making on the part of the physician and patient is important and very complex.

What is necessary to make that decision?

**BENEFITS**
What benefits are being measured and are they important to patients?

**HARMS**
What harms are being measured and are they important to patients?

The best decision is made when everyone is fully informed of the actual benefits and harms that are important to patients!
The Clinical Trial Landscape is Changing

- **Patients** are more involved in research, in clinical trials and in their own healthcare.

- **Patients** want to have a voice.

- **Patients** need the information from other patients on trials to make **Patient Informed Decisions** about their treatment.

- **The process to approve drugs is changing** – accelerated approval, breakthrough designations.

- **Precision Medicine Initiative** – it’s changing the way we do trials.
So, what’s the problem?

Historically:

- QOL questionnaires have been around since the ‘70s
- HR-QOL questionnaires have been around since the ‘90s
- Short forms have been developed for diseases (FACT-B, FACT-P, etc)
- Specific forms for certain side effects have been developed (Anxiety, Depression, etc)

Why is it still so challenging?
Challenges

◆ One size fits all approach (chemotherapy vs targeted therapy)

◆ Although health-related focused, the questionnaire is still very broad

◆ New short forms are added on top of each other

◆ Measuring items not associated DIRECTLY with the patient experience as it relates to disease and treatment

◆ Not a primary objective and not often a secondary objective

◆ Analysis and reporting is not done at the same time as the efficacy clinical trial results (later and different journals)

◆ Information of HR-QOL is not getting back to the patient and is not aiding patients decision-making.
What do patients think are barriers?

- **Redundancy**
  - Being asked the same question over and over
  - Redundant questions irritate patients

- **Complexity**
  - Vague or unclear questions
  - The answer isn’t there and no write in options
  - Not fully understanding the question

- **Time**
  - The time it takes to fill them out - too long
  - The pressure they feel to do it quickly - rushed
  - Too many questions
What do patients think will help?

- **What is being asked is important**
  - Asking about what matters to them
  - Relevant questions to what they are experiencing

- **Knowing why the questions are being asked**
  - Knowing what the results are being used for
  - Relevance to them and other patients

- **Feedback is important**
  - That their responses are being used in their personal care
  - Being informed about the results and how they were used

- **Length/Time matters!**
  - The number of questions and/or the time required
  - Less questions can be asked more often, more questions less often
What changes are needed?

**Now**
- Used in phase 3 trials
- Analyzed separate from efficacy and published at different times
- Information is not shared with patients
- Instrument validation
- Combination of multiple instruments
- Global HR-QOL

**Future**
- Use in early phase trials (1, 2 and 3)
- Analyzed along with efficacy and published together
- Information is shared with public and patients
- Item (question) validation
- Combination of specific items
- Targeted measurements
How do you effectively collect PRO data in trials?

- **Start Early** - collecting PRO data in early trials (Phase I and II) to better inform the collection of PROs in larger later trials (Phase III)

- **Develop targeted (precision) PRO assessments for ALL trials**
  - Asks patients to report what matters to them
  - Make PRO measures acceptable to patients to complete
  - Make PRO measures an important part of the trial with well-defined purpose and use
  - Make the PRO endpoints meaningful to patients
  - Make the PRO information available to patients

**Patient Informed PRO Measures** - **ASKING:**
- The Right Question ~ at The Right Time ~ in The Right Way
How does this happen?

- **Rigor** in development and incorporation of PRO measures in trials

- **Include patients** and other stakeholders in discussions about relevance, timing and construction of PRO measures

- **Process of PRO inclusion during trial development**
  - What is the Treatment (chemo, immuno, biological)?
  - What is the Disease (Stage, type)?
  - What is the Endpoint?
    - What measures are needed?
    - How often will they be collected?
    - When will they be analyzed?

**Patient Informed PRO Measures - ASKING:**

The Right Question ~ at The Right Time ~ in The Right Way
Potential Benefits - Gains

◆ **Less missing data** (potentially)
  ◆ Patients more likely to fill out the questionnaires that are relevant, take less time and they understand their importance.

◆ **Relevant information for informed decision-making**
  ◆ Physicians and patients will have access to the results to make a more informed decision.

◆ **Value will be added**
  ◆ New agents can be approved based upon improving how a patient feels and functions in addition to efficacy and safety.

**Patient Informed PRO Measures** - ASKING:
*The Right Question ~ at The Right Time ~ in The Right Way*
An Industry Example of Instrument Modification

Alicyn Campbell, MPH
Global Head, Patient-Centered Outcomes Research for Oncology
Genentech, a Member of the Roche Group

WORKSHOP ON CLINICAL OUTCOME ASSESSMENTS (COAS) IN CANCER CLINICAL TRIALS

April 26, 2016  ■  Silver Spring, MD
Co-sponsored by

FDA
CRITICAL PATH INSTITUTE

collaborate  ·  innovate  ·  accelerate
Disclosures

- I am currently an employee of Genentech, a Member of the Roche Group
- The opinions and thoughts expressed in this presentation are my own and do not reflect nor represent those of F. Hoffmann-La Roche AG, nor of Genentech, a Member of the Roche Group
Problem Statement

• Tool modification to **increase precision** on concepts assessed (e.g. removing items that are not relevant for the target population or endpoint) and **minimize responder burden** is becoming increasingly common.

• How can modification minimize burden and duplication when using multiple instruments?

• What is an example of an instrument modification that has been done?
Rationale for Modifying Instruments

• Sponsors need to create endpoints to assess outcomes of interest that demonstrate clinical benefit

• Once patient-relevant concepts of interest are identified, sponsors are often limited by:
  – Trial timelines (frequent lack of Ph2 in Oncology)
  – Tools available
  – Study design
  – Differing evidence requirements of global stakeholders
Modification Example from Early Breast Cancer (EBC)
Treatment Burden and Information Gaps in EBC

- Treatment-related symptoms define the patient experience: 90% of EBC patients are asymptomatic at diagnosis\(^1\)
- Patients report the information they receive regarding experience and duration of treatment-related symptoms is inadequate\(^2\)
- Currently available therapies may have negative impact on patients’ ability to function and conduct activities of daily living\(^3,4\), despite reducing the risk of recurrence and death

Collaboration with EBC Patients to Identify Measurement Concepts

• Endpoint workshop conducted with eight breast cancer patient advocates\(^1\)
• Reviewed the study design and participated in open-ended concept elicitation to identify:
  – Key concepts and symptoms for patients undergoing neo(adjuvant) breast cancer therapy
  – Functional impairments
  – Other distressing aspects of disease for EBC patients
• Follow-up qualitative research conducted\(^2\)

**Impact:** Patient-relevant concepts identified to inform trial endpoints and future instrument development

• Reviewed currently available tools for coverage of key patient identified measurement concepts: treatment-related symptoms and impacts (i.e., physical function)

• Recommendation: Adapt the EORTC QLQ-C30 and BR23
  – Modified the QLQ-BR23 in consultation with the EORTC for use in 4 neo-adjuvant and adjuvant studies
  – Removed items 47–53 assessing symptoms and side effects experienced only in Metastatic Breast Cancer
  – Inclusion of peripheral neuropathy and skin problems (2) items from the EORTC Item Bank Approved by EORTC Executive Committee

Steps to Ensure Patient-Relevant Trial Endpoints$^{1,2}$


- Advocate focus groups (former breast cancer patients) to identify concepts
- Review of the literature / measures
- Qualitative interviews with EBC patients / concept saturation
- Analyze patient feedback and consult with EORTC on adaptation for EBC
- Inclusion in Ph3 EBC studies

2012 | 2014
During the past week:

<table>
<thead>
<tr>
<th>Question</th>
<th>Not at All</th>
<th>A Little</th>
<th>Quite a Bit</th>
<th>Very Much</th>
</tr>
</thead>
<tbody>
<tr>
<td>31. Did you have a dry mouth?</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>32. Did food and drink taste different than usual?</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>33. Were your eyes painful, irritated or watery?</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>34. Have you lost any hair?</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>35. Answer this question only if you had any hair loss: Were you upset by the loss of your hair?</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>36. Did you feel ill or unwell?</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>37. Did you have hot flushes?</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>38. Did you have headaches?</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>39. Have you felt physically less attractive as a result of your disease or treatment?</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>40. Have you been feeling less feminine as a result of your disease or treatment?</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>41. Did you find it difficult to look at yourself naked?</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>42. Have you been dissatisfied with your body?</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>43. Were you worried about your health in the future?</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>44. Have you had skin problems (e.g. itchy, dry)?</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>45. Did itching of your skin bother you?</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>46. Did you have tingling hands or feet?</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>
The EORTC QLQ-C30, BR23, and additional single item symptoms scales from the item bank scored according to their published guidelines\(^1\)
- Single items to be assessed by frequency of response option

Analysis to be performed on the modified BR23 in parallel with the final data analysis to confirm the psychometric properties of the instrument
- Pre-specified in the protocol and statistical analysis plan

By using validated single items, translations were available for the composite measure, enabling use in global Ph3 trials

---

1. Fayers et al. EORTC 2001
Conclusions

• Yes you can!
  – Seek patient input
  – Construct a custom, patient-relevant endpoint
  – Remove items not relevant to the patient population under study
  – Meet tight (~12 weeks) timelines in a global trial
  – Differentiate clinical benefit
Considerations in Addressing Respondent Burden

Charles S. Cleeland, PhD
Sandra A. Mitchell, PhD
Andrew Bottomley, PhD
David Cella, PhD

WORKSHOP ON CLINICAL OUTCOME ASSESSMENTS (COAS) IN CANCER CLINICAL TRIALS

April 26, 2016 ■ Silver Spring, MD
Co-sponsored by

FDA
CRITICAL PATH INSTITUTE
In the Beginning...

Deployment of PRO measures in Early Drug Development: Modification of the MD Anderson Symptom Inventory

Charles Cleeland, Lori Williams, Tito Mendoza, Qiuling Shi, Xin Shelley Wang

Department of Symptom Research, MD Anderson Cancer Center
Approach – Developing “Fit for Purpose”

- What could be expected of the investigational agent?
  - Improvement or stabilization of disease-related symptoms or interference with functioning
  - Lengthening of time to symptom deterioration
  - Improvement in tolerability

- How do we begin establishing reasonable PRO endpoints for the clinical trial?
  - Literature review
  - Patient interviews
  - Feedback from standard of care and Phase I dosing trials

- Maximize patient and clinician input
- Use what we know – “incremental validity”
Target populations:

- Patients with the trial’s disease and stage undergoing standard of care
- Patients with the target disease or other cancers receiving the agent (or combination) in dosing and extension trials
Modification Process

- Qualitative interviews for symptom identification and naming – establish content validity
- Comparison of symptoms between groups for item attribution
- Expert panel (patients, caregivers, clinicians) review for symptom relevance, item reduction
- Cognitively debrief final items – item modification
- Administer to validation sample
- Elimination of low-yield items
- Establish psychometric properties and sensitivity

Example of Development: MDASI-OC

MDASI Core Items

Symptom Items
- Fatigue/tiredness
- Nausea
- Pain
- Vomiting
- Sadness
- Distress (emotional)
- Difficulty remembering
- Numbness/tingling
- Lack of appetite
- Shortness of breath
- Disturbed sleep
- Drowsiness
- Dry mouth

Interference items
- General activity
- Mood
- Work
- Relations with others
- Walking
- Enjoyment of life
Candidate Items for MDASI-OC

- Abdominal Pain
- Fullness/bloating
- Constipation
- Problem paying attention
- Back pain
- Leg cramps or leg muscle pain
- Urinary frequency/urgency *
- Pain or burning with urination*

* Clinician suggested

Perioperative Symptom Burden from Laparotomy for GYN Cancer: MDASI-OC

Top Symptom Items

Average score of systemic symptoms:
- fatigue
- dry mouth
- drowsiness

Days (from surgery)

Group
- Standard peri-operative care (n=74)
- Enhanced recovery program (n=154)
“Your appointment’s been cancelled. You took too long filling out those forms.”
Considerations In Addressing Respondent Burden and PRO Item/Construct Overlap in Cancer Clinical Trials

Sandra A. Mitchell, PhD, CRNP
Research Scientist and Program Director
NCI Scientific Director, PRO-CTCAE
Outcomes Research Branch
Division of Cancer Control and Population Sciences
National Cancer Institute
Rockville, MD

mitchells@mail.nih.gov
Is there a burden to respondents and study staff associated with inclusion of PROs in cancer clinical trials?

- Little empiric knowledge of respondent or administrative burden
- Patient perceptions of burden are individual-, context- and measure-specific:
  - Individual: Health status, reading speed, visual acuity, dexterity, understanding of why providing the information is important
  - Context: Frequency, convenience, timing
  - Measure: Time to complete, ease of cognitive task, emotional distress/novelty-engagement
- Many researchers and decision-makers value this information
  - Can add to trial complexity if not accommodated thoughtfully into clinical workflow
  - Handled as any other protocol-related evaluation
In the Contemporary PRO Context, Burden Is Increasingly Addressed Through Technology

- Electronic administration reduces the number of items that must be asked of patients
  - Conditional branching
  - Computer adaptive testing
- Technology has enabled data collection to be more efficient, customized, mobile, and responsive
  - Smart phone/hand-held devices/Interactive Voice Response (IVRS) for data collection
  - Customize time of day for assessment, text size on screen, and mode of administration
  - Reminders to patients and staff for missed surveys
  - Eliminate need for data entry (and the associated errors)
  - Mobile devices may improve engagement in data collection and convenience
What do we really mean when we say that there is duplication or overlap?

- Prima facie evaluation may suggest overlap
- BUT important to consider:
  - Construct (anxiety vs. worry)
  - Attribute/dimension being assessed (e.g. symptom severity vs. interference vs. bother)
  - Recall period (past day, past week, past month)
  - Response scale (Likert, LASA? 4, 5 or 10 point?)
- Duplicative items may be an essential component of a scale
  - Bodily pain (2 items) and vigor (1 item) are part of SF-36 Physical Component Summary (PCS) Score, and are required for a summary score to be calculated
  - There is minor overlap from a conceptual perspective with a measure of pain or fatigue
  - Such redundancy is minor from a practical perspective and must be preserved for a valid interpretation of SF-36 PCS score
Why not just address any redundancy by omitting a ‘duplicative’ item from a scale or index? Isn’t that fundamentally what an item library or item bank does?

- Alter measurement properties
  - Reliability, validity, responsiveness
  - Change established factor structure
- Lose the ability to compare sample values with other population values/historical controls
- Shift scoring, particularly if it is a weighted scale or a summated index
- For example with an index, a high score on one indicator may compensate for a low score on another; removing one or more items (particularly if that item reflects a rare or common symptom) can significantly shift score interpretation
- Item library is a collection of items, some of which may be nested or grouped
- Item bank is a repository of items along with information about their alignment or calibration against a standard, allowing for scores to be expressed on a common psychometric scale
Why not just create a ‘customized’ PRO strategy by selecting relevant single items from validated measures?

- Single items are valuable in measuring narrow constructs or providing global assessments
- Single items may be useful for screening or as part of a branching strategy
- Lower reliability and restricted range for responsiveness to change
- Ceiling and floor effects (e.g., single item to measure fatigue severity may provide no range for worsening in respondents rating fatigue as severe or very severe at baseline)
- Collection of single items also invites problems of multiplicity
- Summation of single items can be less than informative and may obscure change over time
  - High score on one indicator may compensate for a low score on another
How much symptom domain overlap exists among PRO measures commonly used in oncology trials?

<table>
<thead>
<tr>
<th>SF-36</th>
<th>PROMIS-29</th>
<th>FACT-G</th>
<th>EORTC-QLQ C-30</th>
<th>MD Anderson Symptom Inventory</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fatigue</td>
<td>Fatigue</td>
<td>Fatigue</td>
<td>Fatigue</td>
<td>Fatigue</td>
</tr>
<tr>
<td>Pain</td>
<td>Pain</td>
<td>Nausea</td>
<td>Pain</td>
<td>Pain</td>
</tr>
<tr>
<td>Mood</td>
<td>Depression</td>
<td>Depression</td>
<td>Depression</td>
<td>Depression</td>
</tr>
<tr>
<td>Anxiety</td>
<td>Anxiety</td>
<td>Insomnia</td>
<td>Anxiety</td>
<td>Nausea</td>
</tr>
<tr>
<td>Insomnia</td>
<td>Insomnia</td>
<td>Insomnia</td>
<td>Insomnia</td>
<td>Nausea</td>
</tr>
<tr>
<td>Shortness of breath</td>
<td>Shortness of breath</td>
<td>Shortness of breath</td>
<td>Shortness of breath</td>
<td>Numbness</td>
</tr>
<tr>
<td>Anorexia</td>
<td>Anorexia</td>
<td>Anorexia</td>
<td>Anorexia</td>
<td>Dry mouth</td>
</tr>
<tr>
<td>Nausea</td>
<td>Nausea</td>
<td>Nausea</td>
<td>Nausea</td>
<td>Distress</td>
</tr>
<tr>
<td>Vomiting</td>
<td>Vomiting</td>
<td>Vomiting</td>
<td>Vomiting</td>
<td>Drowsiness</td>
</tr>
<tr>
<td>Constipation</td>
<td>Constipation</td>
<td>Constipation</td>
<td>Constipation</td>
<td>Global Symptom</td>
</tr>
<tr>
<td>Diarrhea</td>
<td>Diarrhea</td>
<td>Diarrhea</td>
<td>Diarrhea</td>
<td>Interference</td>
</tr>
<tr>
<td>Concentration problems</td>
<td>Concentration problems</td>
<td>Concentration problems</td>
<td>Concentration problems</td>
<td></td>
</tr>
<tr>
<td>Memory problems</td>
<td>Memory problems</td>
<td>Memory problems</td>
<td>Memory problems</td>
<td></td>
</tr>
</tbody>
</table>

Estimated duplication (<11 items) created by use of PRO-CTCAE weekly along with any of these measures.

<11 items <11 items <11 items <11 items
Key Points

• Estimated additional time burden introduced by overlap is estimated to be 2 minutes or less per HRQL or disease-specific symptom measure
  – Given that many trials include 4-6 post-baseline assessments for an HRQL endpoint, use of a measure of disease-related symptoms or HRQL together with weekly PRO-CTCAE reporting, adds an estimated 15 minutes or less of additional burden attributable to ‘overlap’, across the course of a trial

• Strive for a parsimonious measurement strategy
  – Hypothesis driven and focused on a specific set of constructs (e.g. symptomatic toxicity, physical function, disease-related symptoms, )
  – Timing of assessments based on clinical knowledge of the anticipated treatment effects and tumor response

“Everything should be made as simple as possible. But not simpler.”
Albert Einstein
Key Points

- CAT and IRT-based measurement allows for parsimony and precision and reduces non-informative overlap
- Pulling apart established measures by selecting a subset of items has the potential to set our field back:
  - Lose the established measurement properties of the scale/subscale
  - Chance ceiling or floor effects
  - Abandon interpretive norms and minimally important difference
  - Invite challenges of dealing with multiplicity
  - Attract the interpretive difficulties associated with aggregated (or summed scoring)
- Single items have an important role in screening, in well-validated global impression items, and as part of a conditional branching strategy
- Measures can be tailored over time through rigorous empiric methods

“Everything should be made as simple as possible. But not simpler.”

Albert Einstein
Reducing patient burden issues in cancer clinical trials: An EORTC perspective

Andrew Bottomley, PhD.
Assistant Director,
Head Quality of Life Department
What is the EORTC experience in tool development?

1. **Questionnaire development**
   - In older trials, prior to 1980, burden caused by filling in QOL tools was a major problem in EORTC trials, which sometimes failed for this reason (e.g. Diary!).
   - Nowadays this is a very rare problem.
     - EORTC AML 13 trial had a compliance of 40% at baseline
   - Since 1990, EORTC instrument guidelines\(^1\) and clinical trials manuals\(^2\) stress the fact that we respect patients views. We do our best to never burden them and always make efforts to reduce burden.
   - Developing new tools takes over 6 years and involves 1000s of patients. Asking for redundant items over the years is critical and a major part

2. **Trial development and designing**
   - EORTC trials with QOL (over 160) usually have 2 tools plus EQ5D. Rarely more.
   - EORTC found no difference when patients complete 1, 2 or 3 (EQ5D) tools

   Are we sure patient burden is as much of a challenge as we think?

---

Challenges of multiple static tools

EORTC is aware that sometimes not all items are valid in the whole trial population, as static tools are developed for a broad ranges of patients within a disease (e.g. Lung covers SCLC and NSCL).

- Not all patients are the same, all are different!

*Examples where items in EORTC tools may be less important to some patients:*-

**Patient specific**
- Does your hair loss bother you? - How is this relevant to a bald man?
- Unable to walk up a flight of stairs? - How is this relevant for people using elevators?

**Population/ design specific**
- Nausea and vomiting - How is this relevant to a cancer survivor off treatment one year post chemo?
- Diarrhea - How is this relevant to an early breast cancer patient 12 months post therapy, assessed every year for 5 years?
- Fear of recurrence when patients in a trial of recurrent disease
How does EORTC cope with this in trials, as all trials are different with new therapies or novel combinations?

- We accept that there are limitations with static tools: newer IT based tools can be a future solution.
- However, we are not currently running with new IT tools, or will be for the next 5 years in any standardized way. So what do we do?

- Accept the limitations of static tools (as we have for decades), but also recognize advantages (e.g. can be done piece meal: when wanted, without IT such as PC, internet, WIFI, electrify or batter power, telephone (IVR))
- Design trials with attention to over burden (only critical tools) and key assessments time points
- Understand that patients (at least in the EU) rarely complain about too many tools.
  - One patient per trial, X 12 assessment points, x 15 mins = 3 hours effort (often in a clinic waiting).
  - In the perspective of a patient is 3 hours over a trial of 2-3 years so much?
  - Would they rate hospital parking, waiting for HCP, waiting for treatment maybe all more of an issue than spending 15 minutes in a waiting rooms filling in a measure?

- Evidence that patients like to complete relevant QOL measures

- Solutions are not always to cut down tools to the endpoints of scales that you think are critical, as you may miss points. Eg. We hope we select the best 5-6 scales, but even if we rely on data from phase 1-2 trials and on clinical judgments, we cannot always predict correctly.
What can help us address patient burden concerns? An EORTC approach

• **A good design**
  - Don’t include too many tools.
  - Try not to over assess patients every month.
    - If you have too many questions ethics committees, clinicians and patients will object and this will lead to problems.
  - Be clear than some good quality data is important.
  - Push for making the investigators aware: kick off meetings, training, newsletter, awards for cakes, manning and shaming, education, funding compliance. Trial motivator.

• **Explain** in the informed consent that some questions may seem less relevant to you but could be important for patients who may have these issues, and explain that some items may overlap, but this is so we can have a detailed picture of the issues which we are not sure about.

• We have several trials in EORTC where we did not have a scale or full tools, so we used the EORTC *item bank* to select items, to be added as a check list (e.g. Muscatus in a bladder trial?).
  - Our item bank has 1500 items and all translation, so it is an easy tool, free for all academics!
The Future

• Over the next 10 years, newer IT based tools will help refine measurement practices in RCTs.

• However, patients may want all options including pencil and paper, and while many patients like ePRO, some studies show that in less economically advanced rural areas in EU patients prefer pen and paper.

• Keep the challenge of patient burden in perspective.
  • Patient burden is a challenge, and may grow as demand from users increases, but remember patients are helpful, will be happy to help: after all, in EORTC experiences, patient burden rarely emerges as a problem.
Considerations in Addressing Respondent Burden

David Cella, PhD
Chair, Department of Medical Social Science, Northwestern Feinberg School of Medicine
Rules of thumb

For the typical oncology clinical trial:

- 4-6 assessments over 1-2 years
- Fewer than 50 items per assessment
- Less than 15 minutes per assessment
- For every overburdened patient there are 10 grateful ones
Guideline for item burden based on assessment frequency

Assessment frequency

- Daily (x30)
- Weekly (x12)
- Monthly (x3-6)
- Quarterly (x4-8)
- Once

# of items

- Daily (x30)
- Weekly (x12)
- Monthly (x3-6)
- Quarterly (x4-8)
- Once
Leveraging Existing and Emerging Tools to Optimize PRO Assessment Strategies in Cancer Trials
Session Participants

Chair
- Stephen Joel Coons, PhD – C-Path

Presenters
- Paul G. Kluetz, MD - FDA
- Patty Spears – Cancer Information and Support Network
- Alicyn Campbell, MPH – Genentech
- Charles S. Cleeland, PhD – M.D. Anderson Cancer Center
- Sandra A. Mitchell, PhD – NCI, NIH
- Andrew Bottomley, PhD – EORTC
- David Cella, PhD – Northwestern University Feinberg School of Medicine

Panelists
- Elektra Papadopoulos, MD, MPH – FDA
- Jeff A. Sloan, PhD – Mayo Clinic
WORKSHOP ON
CLINICAL OUTCOME ASSESSMENTS (COAs)
in Cancer Clinical Trials

April 26, 2016  Silver Spring, MD

Co-sponsored by