CDDF 10TH ALPINE CONFERENCE
CURRENT AND FUTURE CHALLENGES OF INNOVATIVE ONCOLOGY DRUG DEVELOPMENT
26 - 28 February 2018
Innsbruck, Austria

PATIENTS' REPORTED OUTCOMES IN ONCOLOGY PATIENT- CENTERED DRUG DEVELOPMENT:

OPPORTUNITIES & CHALLENGES

10th Alpine Conference
Innsbruck, Austria, 27 February 2018

Disclaimer

The views and opinions expressed in this session are those of the individual presenters and panelists and should not be attributed to or considered binding on their respective organizations or to CDDF

Co-Chairs:

- Eva Skovlund, Norwegian University of Science and Technology
- Sarah Bobiak, Merck

Panelists:

- Andrew Bottomley, EORTC
- Daniel O'Connor, MHRA
- Amylou Dueck, Mayo Clinic
- Dan Danielson, Premera
- Paul Kluetz, FDA

Organizers:

- Irmela Radtke, Regulatory, Roche
- Jan Gross, Regulatory, EMD Serono
- Elisabeth Piault-Louis, PCOR, Genentech

Objective:

Discuss the use of Patient Reported Outcomes (symptoms, function, quality of life, etc) to document clinical benefit in oncology drug development.

Format:

One panelist to introduce each of the three topics for discussion followed by panel discussion and audience questions

- Study design
- Endpoints
- Instrument

Patient Reported Outcomes in Oncology

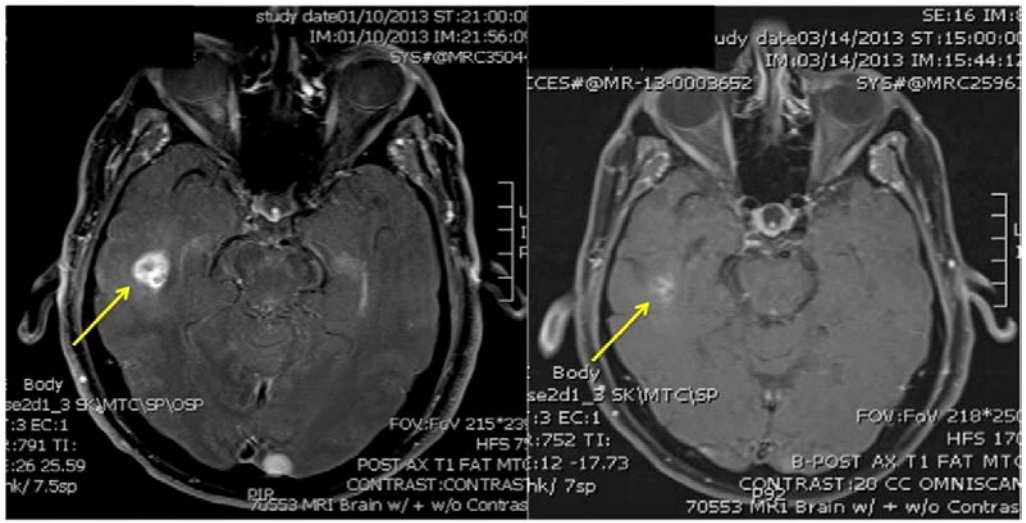
Sarah Bobiak, PhD
Scientific Director, Immuno-oncology
Global Evidence & Value Development | R&D
Merck

Clinical benefit defined

FDA	Impact of a treatment on how an individual feels, functions, or survives
EMA	Convincingly demonstrate favourable effects on survival are. Acceptable primary endpoints [to document clinical benefit] include cure rate, OS and PFS/DFS.
HTA (e.g. GBA)	The benefit of a new drug is the patient relevant therapeutic effect with regard to an increased health status, reduction of duration of disease, prolongation of survival, reduction of adverse events or increased QoL.

How is clinical benefit of a treatment measured in oncology clinical trials?

10603 (started February 1, 2013)
CH5424802 760 mg bid



January 10, 2013 March 14, 2013

What does this MRI tell us about how the disease and treatment impact the patient' day to day life?

Patient Reported Outcomes

Please rate your pain by tapping the one number that best describes your pain at its **worst** in the last week.

0	1	2	3	4	5	6	7	8	9	10
No Pain									Pain as bad as you can imagine	

Tap the one number that describes how, during the past week, pain has interfered with your:

A. General Activity

0	1	2	3	4	5	6	7	8	9	10
Does not interfere									Completely interferes	

B. Mood

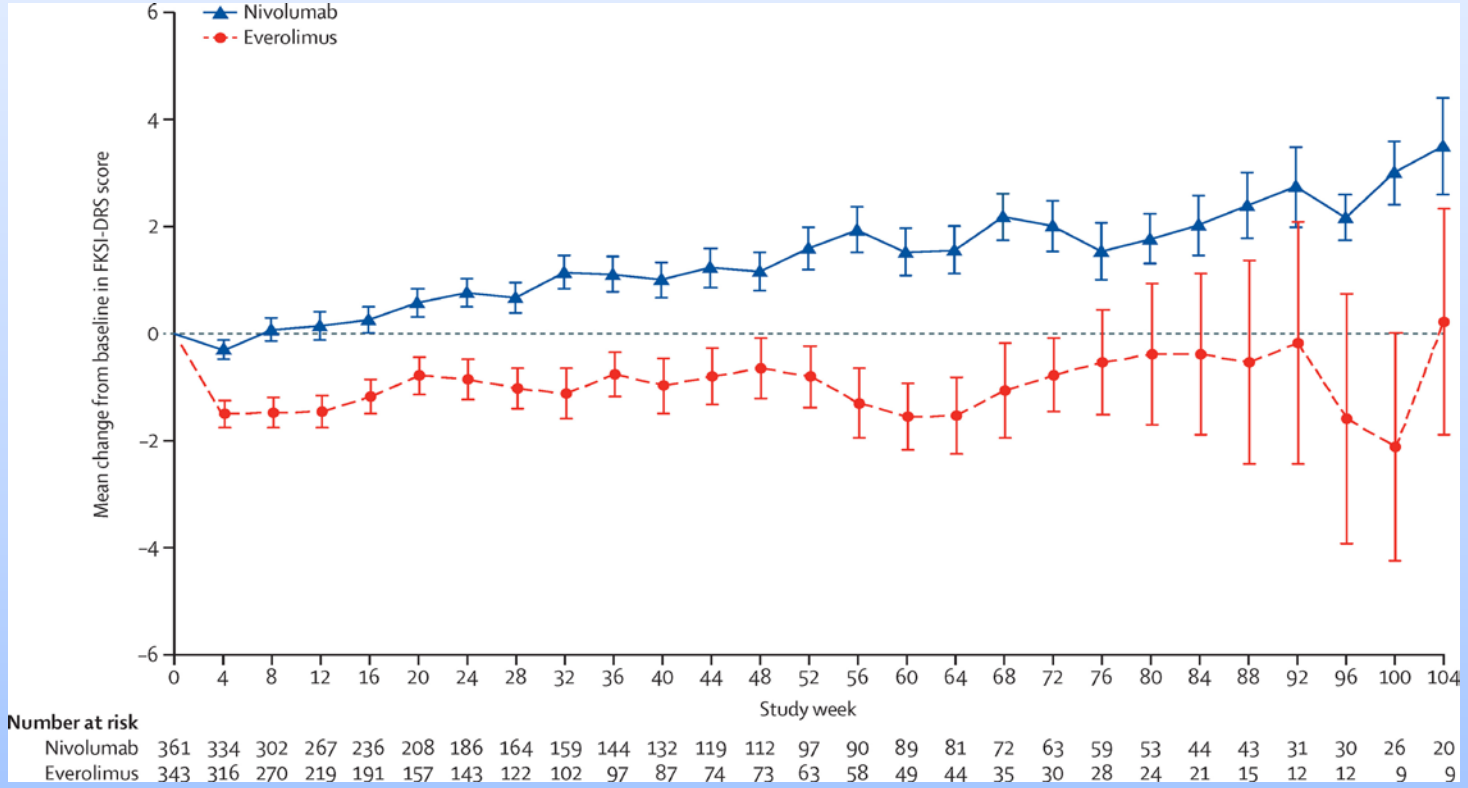
0	1	2	3	4	5	6	7	8	9	10
Does not interfere									Completely interferes	

C. Walking Ability

0	1	2	3	4	5	6	7	8	9	10
Does not interfere									Completely interferes	

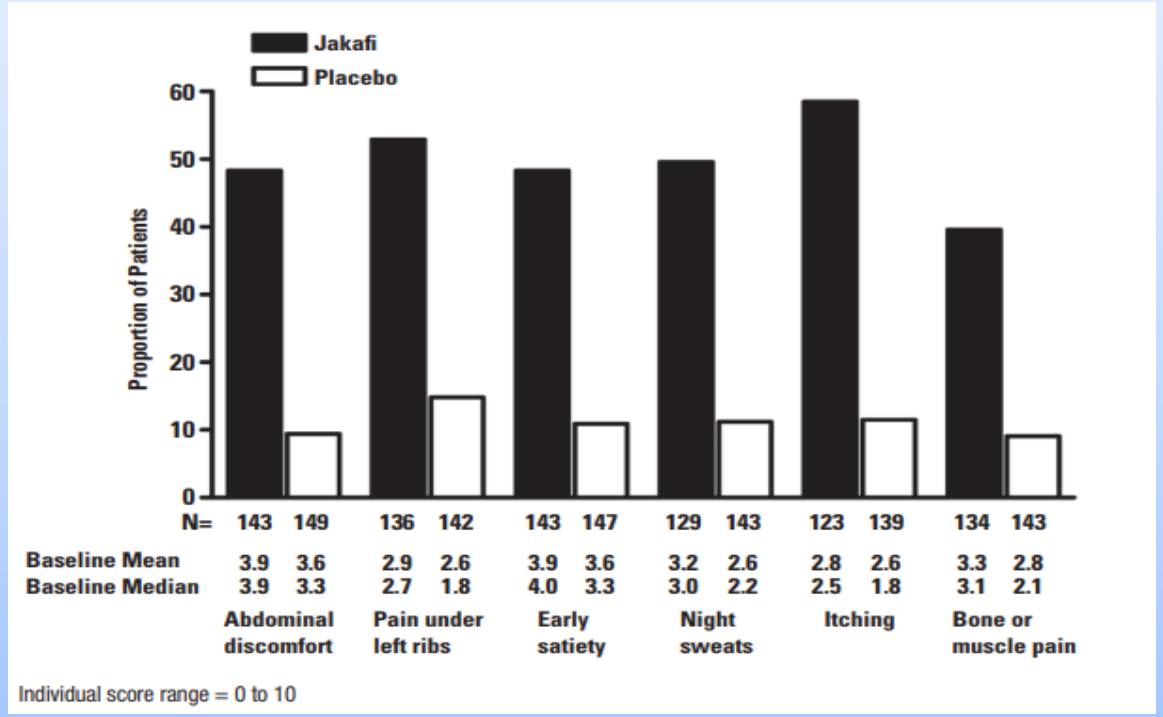
Data regarding any aspects of a patient's health condition reported by the patient him/herself

Complementary information



The Lancet Oncology 2016 17, 994-1003DOI: (10.1016/S1470-2045(16)30125-5)

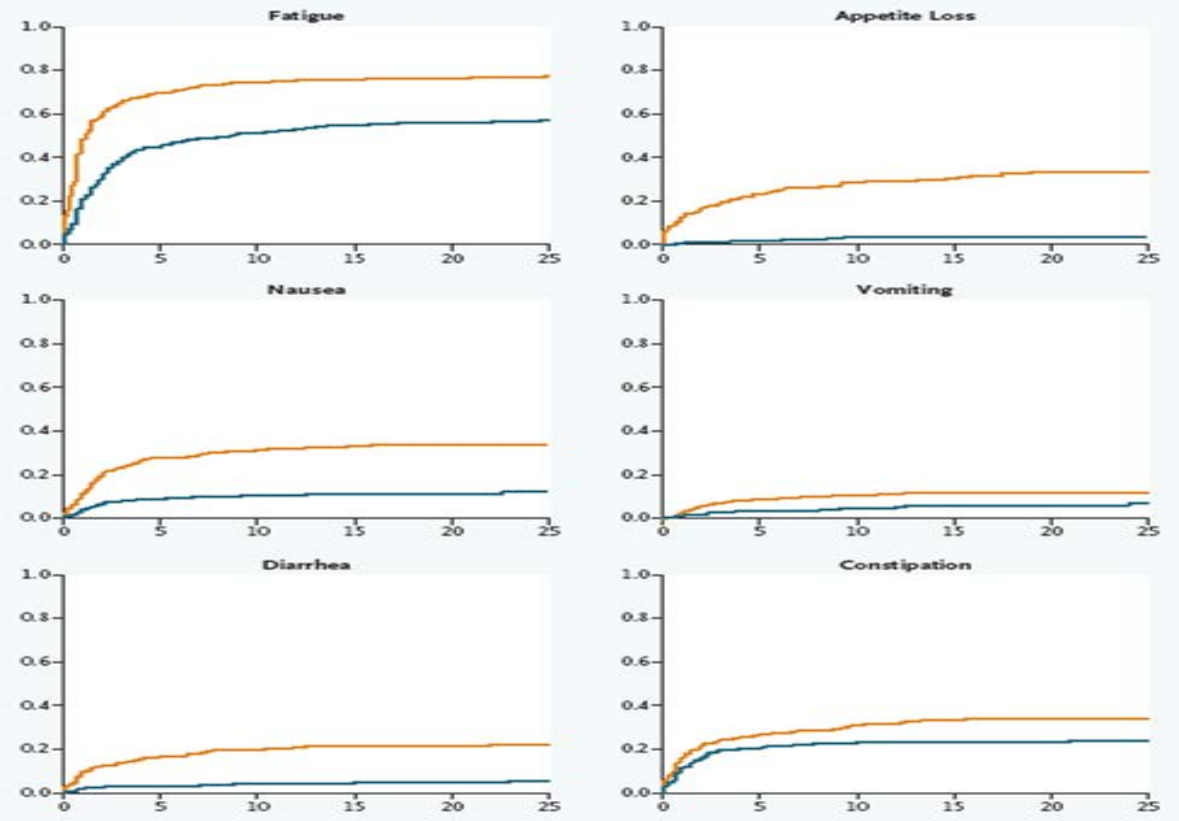
Complementary information



Jakafi, US Product Information

Complementary information

Orange: Patient- reported; Blue: Clinician reported



Basch: NEJM, 2010

Current situation

- Increasing amount of patient reported data is available in ongoing and legacy trials
  - OS and PFS primary endpoints, PRO are often other secondary or exploratory endpoints
- Novel drug development setting (accelerated approval, single arm studies, open-ended studies) is changing measurement paradigm and endpoints
  - OS increasingly challenging to document, PFS not necessary a validated surrogate
  - Accelerated drug approvals based on surrogate endpoints (e.g. PFS, PCR) that need conversion with clinical benefit data
  - Multiple stakeholders
    - FDA, EMA, HTA have (slightly) different expectations regarding PRO data and relative weight attributed in R:B assessment framework
    - Patients and Clinicians are requesting patient-relevant information in addition to additional quantity of life

[1] FDA PRO Guidance for industry, 2009; [2] EUNetHTA HRQoL, 2013; [3] EMA Appendix 2, 2016

Study Design

Daniel O'Connor, MB ChB, PhD
Expert Medical Assessor, MHRA, UK

Why include PRO assessment in your study?

- Provide a patient focused assessment of **the burden and impact of disease**, by understanding **how a treatment impacts on patient functioning and well-being**
- Add information on the clinical benefit of a therapy by **complementing** efficacy and safety data with patient-reported evaluation
- Assess the **relationship/ agreement between** clinical reported endpoints and other patient-reported endpoints
- Provide information to facilitate more **accurate** future patient-physician communication in terms of:
  - Quality of the survival time remaining for the patient
  - Burden of treatment-related morbidities and disease-related patient impacts

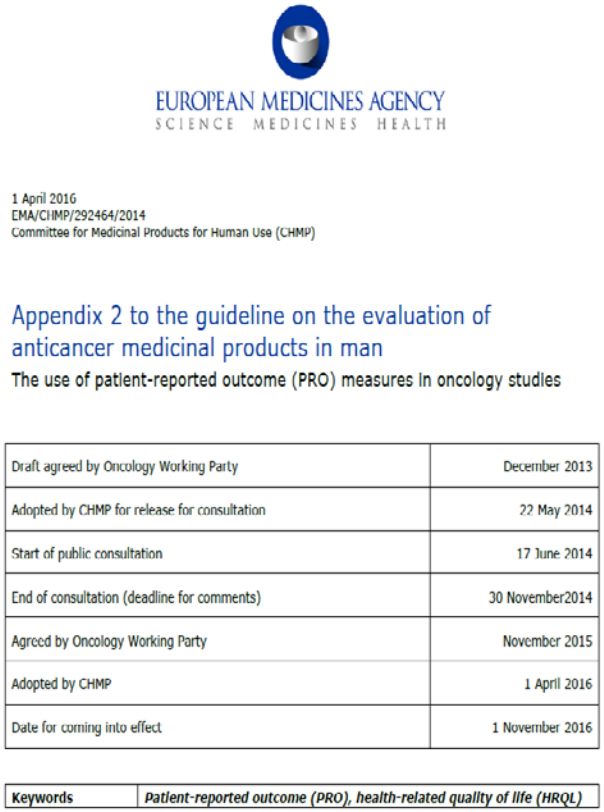
EMA Appendix 2: PRO measures (April 2016)

Key message

'The importance of the patient's point of view on their health status is fully acknowledged and such information may be used in drawing regulatory conclusions regarding treatment effects, in the benefit risk balance assessment or as specific therapeutic claims'

Key aim

'By outlining broad principles of scientific best practice rather than prescribing a particular approach to PRO selection and application, the appendix aims to encourage developments in the methods and application of PROs in the oncology regulatory setting'



EUROPEAN MEDICINES AGENCY
SCIENCE MEDICINES HEALTH

1 April 2016
EMA/CHMP/292464/2014
Committee for Medicinal Products for Human Use (CHMP)

Appendix 2 to the guideline on the evaluation of anticancer medicinal products in man
The use of patient-reported outcome (PRO) measures in oncology studies

Draft agreed by Oncology Working Party	December 2013
Adopted by CHMP for release for consultation	22 May 2014
Start of public consultation	17 June 2014
End of consultation (deadline for comments)	30 November2014
Agreed by Oncology Working Party	November 2015
Adopted by CHMP	1 April 2016
Date for coming into effect	1 November 2016

Keywords	*Patient-reported outcome (PRO), health-related quality of life (HRQL)*

Points to consider

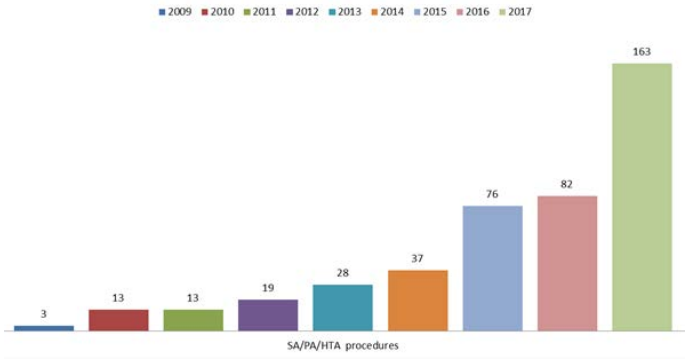
- Strongly consider scientific advice +/- joint with HTA bodies
- Assessment or rationale for the extent to which the inclusion of a PRO measure can provide added value in the clinical trial setting
- PRO measure should be considered early in the development program
- Consideration should be given to patient involvement in the study design process
- PRO endpoints should be stated as a specific clinical trial objective or hypothesis in the study protocol and statistical analysis plan
- PRO measures should be administered to study subjects at time points when there is a clear rationale for their use
- Measurements should not constitute an undue burden to the patient

Points to consider (Con'd)

- Blinding is not always possible and the priority should be capturing the patient experience through PRO data in all types of clinical trial designs
- Justify why certain timings of assessments were selected and why the instrument is sensitive to capture a range of anticipated and unanticipated effects
    - Most appropriate and valid PRO measures have involved patients in their development
    - Should be evidence based, shown to measure the concept it is intended to measure
    - Appropriate for the research objective, the disease and patient population
    - Practical trial considerations (respondent burden, feasibility)
- The study protocol should describe the principal data analysis features in the statistical section, with a detailed elaboration of the analysis and how to control and handle missing assessments

Recent initiatives, & involving patients

- New initiative from ONCWP on the development of high level principles on how and when to report PRO data in the assessment report and SmPC
- International collaborations in improving the standards of PRO in clinical trials
  - Setting International Standards in Analysing Patient-Reported Outcomes and Quality of Life Endpoints Data (SISAQOL)
  - Standard Protocol Items: Recommendations for Interventional Trials (SPIRIT) PRO extension
- EMA workshops on single arm studies, immunotherapy and histology independent indications – addressing evolving regulatory science
- The patient is a key partner in drug development, enhancing the understanding of the disease impact



Discussion

- What are the **core concepts** you need to see to evaluate the **clinical benefit of a treatment**?
- What are **the key issues for the study design** when considering PRO?
- What are the **key challenges in reviewing PRO data** collected in oncology trials?
- How to collect and interpret PRO in the **novel study designs** e.g. single arm, basket, umbrella?

Endpoints

Amylou Dueck, PhD
Associate Professor of Biostatistics
Mayo Clinic
Scottsdale, Arizona, USA

Key concept

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

- A precise definition of an **endpoint** typically specifies the type of assessments made, the timing of those assessments, the assessment measures used, and possibly other details, as applicable, such as how multiple assessments within an individual are to be combined.

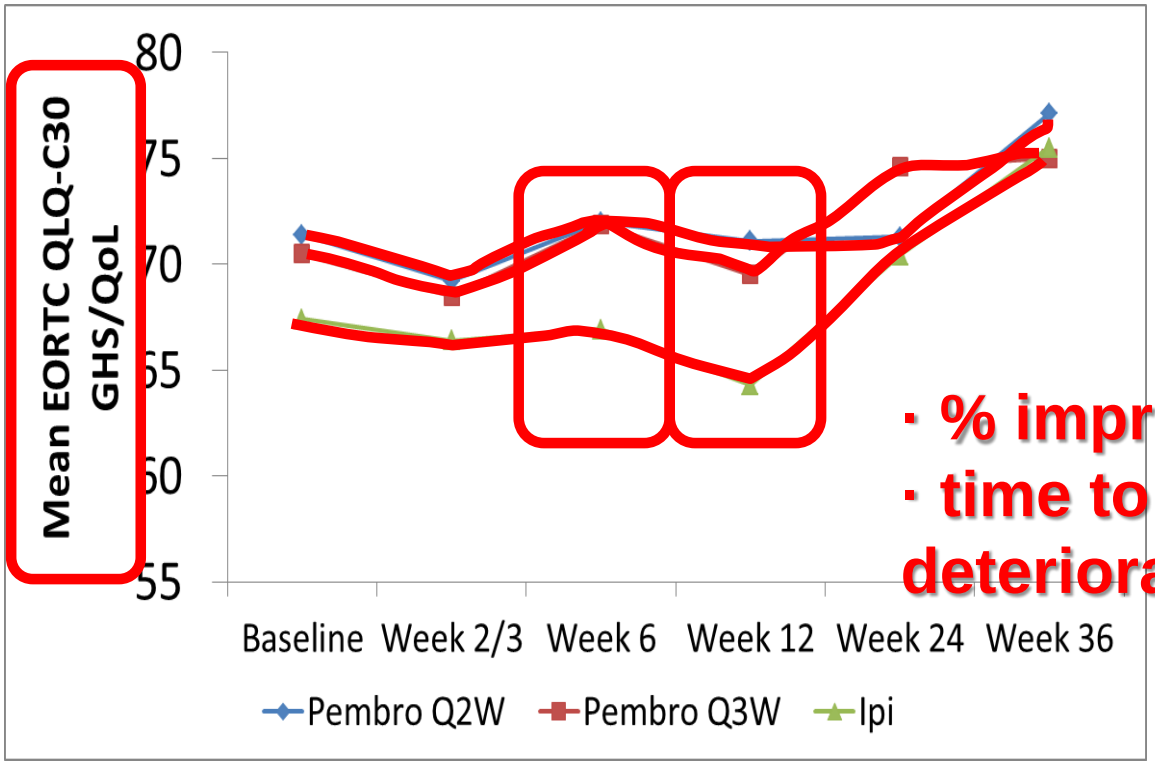
Defining a PRO endpoint

Let's assume that you already have:

- A research question of interest (i.e., hypothesis)
- Concept(s) for measurement (e.g., HRQoL improvement, symptom deterioration)
- A valid PRO measure
- Time points selected

Some PRO endpoint options – A picture

KEYNOTE-006: EORTC QLQ-C30 & EuroQoL EQ-5D @ baseline; weeks 2/3, 6, 12, 24, & 36; tx discontinuation; & 30-day f/u visit



Petrella TM, et al. Eur J Cancer. 2017; 86:115-124.

Points to consider

- Endpoint should be based on research question of interest
- Endpoint should be meaningful to patients
- Endpoint may be influenced by availability of data
  - KEYNOTE-006: >50% of the patients in the ipi arm were expected to have disease progression at week 12

Points to consider (con'd)

- Specificity vs. multiplicity (summary & composite endpoints)
  - Multidimensional instruments (total score vs. scale score**s**)
  - Composite endpoint (dyspnea, pain, cough) vs. individual symptom endpoint**s**
  - Combining data across time points

  For: aid interpretation, handle multiple testing, can allow some missing data

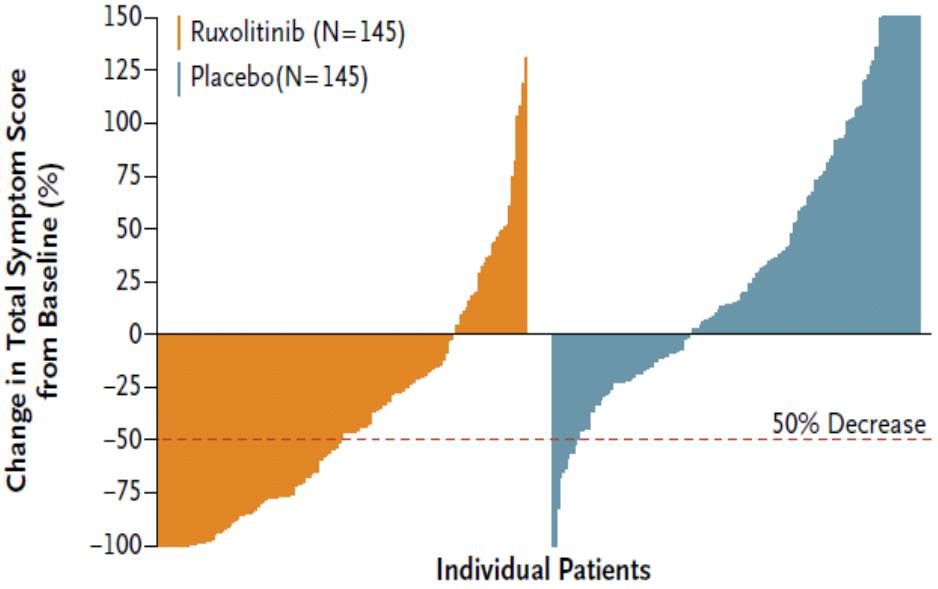
  Against: high-level summarization can obfuscate important individual differences, "off-the-shelf" measure may include concepts which are irrelevant in the current setting, bias can be introduced by missing data

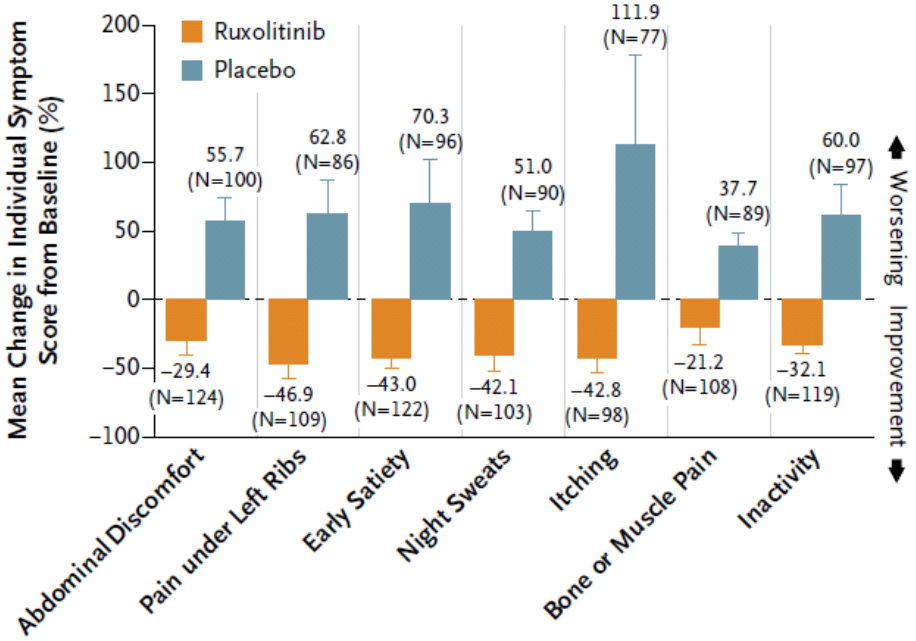
Points to consider (con'd)

- "Responder" and time to event type endpoints have their own challenges (e.g., need meaningful change threshold)
- Choice of endpoint directly impacts analytical approach and statistical power
  - Time to event, longitudinal models, proportion of patients, etc.
  - Some endpoints may also require particular supplemental analyses (e.g., a "responder" type endpoint typically requires cumulative distribution function analysis; a summary or composite endpoint typically requires separate analysis of its individual components)
- Choice of endpoint can impact population
  - If targeting a symptom palliation endpoint, require symptomatic @ BL?
  - For other endpoints: ITT, PRO-evaluable, other?

Example

COMFORT-I: Reduction in the total symptom score of 50% or more from baseline to week 24, as assessed with the modified Myelofibrosis Symptom Assessment Form (MFSAF) version 2.0





Verstovsek S, et al. N Engl J Med. 2012; 366(9):799-807.

Opportunities and challenges

Lack of standardization	• Is there really that much variability in PRO endpoints in drug development clinical trials? • If yes, can PRO endpoints be standardized within certain diseases, settings, PRO measures, etc.?
Pre-specification vs. post-hoc	• What would be a well-defined PRO endpoint? • If PROs are considered as exploratory in a clinical trial, do endpoints need to be specified? • Is post-hoc analysis ever appropriate (e.g., early phase clinical trials)?
Multiplicity	• Is the use of a total, summary, and/or composite score endpoint a reasonable approach to handling multiplicity? • Should PRO endpoints be considered within the hierarchy of clinical endpoints? • Is strict adjustment of alpha across all PRO endpoints necessary? What about PROs measuring tolerability?

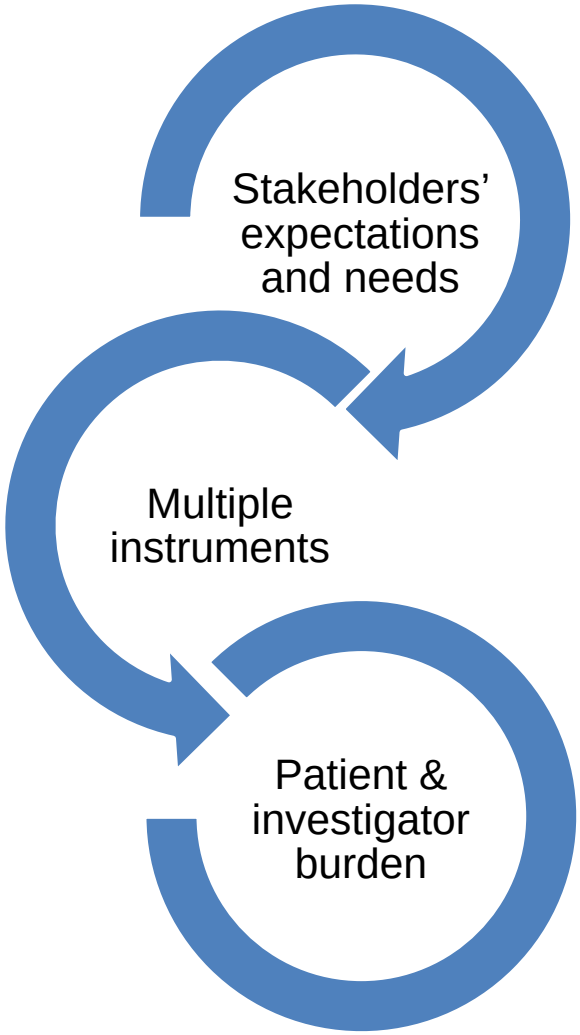
Discussion

- What would be a "**well-defined"** endpoint?
- What are the key challenges in interpreting **PRO endpoints** in oncology trials?
- What would be your recommendation to account for testing **multiplicity** inherent with multi-concept instruments/multiple analytical approaches?

Measure Validity

Andrew Bottomley, PhD
Assistant Director, Head of QOL Department
European Organisation for Research and Treatment of Cancer

Background & rationale



- How can we reduce **patient burden** while collecting **high quality data** using tools that meet the needs of our research questions?
- Are **static measures** sufficient to capture the experiences of patients undergoing novel (e.g., targeted) forms of treatment?
- How can state-of-the-art PRO tools help us address these challenges?

Multiple measures: Increased burden

Stakeholders now demand many measures, e.g., EQ-5D + EORTC QLQ-C30 + PROMIS + PRO-CTCAE

Patients	Investigators
• Longer completion time • Concept repetition & overlap • Higher cognitive load • Decreased motivation if perceived as irrelevant	• Higher attrition & missing data • Reduced quality of data • Difficult interpretation (e.g., similar domains across several measures)

Increased burden: How to respond?

- Careful selection and administration of instruments
  - Must be relevant to protocol and research hypothesis (but not excessive) → not simply done as a stakeholders preference
  - Patients must see the tools as meaningful (increased motivation)
  - Minimizing overlap between measures
  - Single vs. multi-items → will single items be sufficient in some trials?

- Use of dynamic, patient-centered tools
  - Item banks/libraries
  - Computerized adaptive testing (CAT)
  - Electronic PRO (ePRO)

Single vs. Multi-Items

- Some constructs require more items than others
  - Appetite loss vs. physical functioning
- Complex concepts can be interpreted in different ways
  - May lead to poor reliability
- **Reliability** (the extent to which a measure gives consistent results) is a necessary, but not sufficient, component of validity
  - Poor reliability undermines claims about validity and sensitivity
- Balance is needed to have the minimum number of items to capture the intended construct → typically, more items = higher reliability

Item Banks/Libraries

- New tools to create individualized measures and ad hoc checklists
- Several measurement systems with different scopes
  - PRO-CTCAE → symptom measurement
  - PROMIS (CAT available) → QOL and symptoms but not cancer specific
  - EORTC Item Library (CAT available) → cancer specific QOL and symptoms
- Can reduce burden by minimizing number of measures required
- Increased **content validity**
- Increased **flexibility** and **efficiency**
  - More tailored to the needs of specific treatments and populations
- Can identify important gaps and inspire future development of measures

Computerized Adaptive Testing

- Works in conjunction with item banks/libraries
- Relies on item response theory (IRT)
  - Shorter, more individualized measures
    - Reduced patient/investigator burden
  - Increased construct validity
  - Higher precision
    - Smaller sample sizes without loss of power
- Not yet systematically used in clinical trials

Ongoing Challenges and questions

Data Analysis	• Often limited to descriptive scores only for item banks/libraries • How to select the best single vs. multi-items & scales → are these valid for specific trials? • Implications of modifying scale structure
Comparability	• Limited data for new tools generated from item banks/libraries • Tradeoff between individualizing measures and our ability to compare results across groups and measures
Clinical Meaningfulness	• More data needed on predictive accuracy & validity of new tools • How to determine clinically meaningful thresholds for item banks/libraries → can these be compared to those from static measures?

Discussion

- What could be done to reduce **completion burden and concept redundancy** without decreasing the quality of PRO data?
- What are your expectations with regards to **demonstration of questionnaire/score/endpoint "validity**”?

Wrap Up

Eva Skovlund, PhD

Professor, Norwegian University of Science and Technology (NTNU)