







ISSUE PANEL: PROVE IT WITH PROS

Yvonne-Beatrice Böhler Stefan Holmstrom Finn Boerlum Kristensen Anke van Engen



November 12, 2018 ISPOR European Conference, Barcelona, Spain





Introduction

Anke van Engen

There is a drive to incorporate the patient experience of treatment into healthcare decision making













What do we mean by the patient experience?

Concepts measured in PROs

Concept	Description
Health-Related Quality of Life (HRQoL) (incl. health status)	HRQoL is multi-dimensional and represents the patient's evaluation of a health condition and its treatment on daily life: physical function, psychological function, social function, role function, emotional function, well-being, vitality, health status, etc
Health-Related Quality of end of life	Same as HRQoL at the end of life
Patient satisfaction	Evaluation of treatment, patients preference, health care delivery systems and professionals, patient education programs and medical devices
Physical functioning	Physical limitations and activity restrictions, e.g. self-care, walking, mobility, sleep, sexual disability
Psychological functioning (incl. coping)	Positive or negative affect and cognitive, e.g. anger, alertness, self-esteem, sense of well-being, distress, coping
Signs and symptoms	Reports of physical and psychological symptoms or sensations not directly observable, e.g. energy and fatigue, nausea, irritability
Social functioning (incl. work)	Limitations in work or school, participation in community
Treatment adherence	Reports or observations of actual use of treatments
Utility	Generic measures of HRQoL with societal reference weights for their classification systems that can help to inform health-care resource allocation. Utilities provide a useful summary index of overall QoL relative to full health (utility = 1) and death (utility =0)

- How a patient feels and functions in the context of their disease or condition
- How a patient feels and functions in the context of their treatment

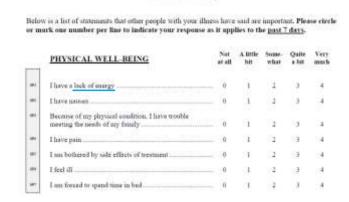
Source: ISPOR HTA Training material, 2018





PRO instruments should capture symptoms that are meaningful to patients

"I just overall didn't have a lot of energy for doing the kind of things that I was normally used to doing, like playing sports and things like that. Even just doing activities around the house, I got to be kind of a couch potato, just didn't have a lot of energy to do things." - Prostate cancer patient



FACT-P (Version 4)

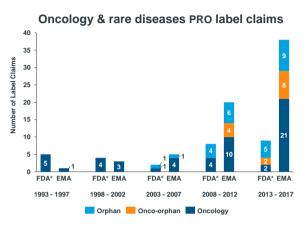
Source: IQVIA research; http://www.facit.org/LiteratureRetrieve.aspx?ID=42292



In regulatory decision-making, the science and requirements for PRO evidence have matured



Consequently, the last decade has shown a marked increase in PRO data included in oncology labels, particularly in Europe



*Note: The FDA labels only include PRO label claims, FDA descriptive labelling was excluded from the analysis

Sources: ePROVIDE PROLABELS.- analysis of oncology drugs and drugs with EMA orphan drug status tus%5D=Approved

Key drivers

- · Rise of patient-centricity
- Effect on overall survival may be difficult to detect
- Improved outcomes with standard of care raising the bar for traditional endpoints
- · Safety profile of chemotherapies
- · Technology enablers

EMA label claim: "Osimertinib improved patientreported lung cancer symptoms compared to chemotherapy by demonstrating a statistically significant difference in mean change from baseline vs. chemotherapy for all 5 pre-specified primary PRO symptoms.



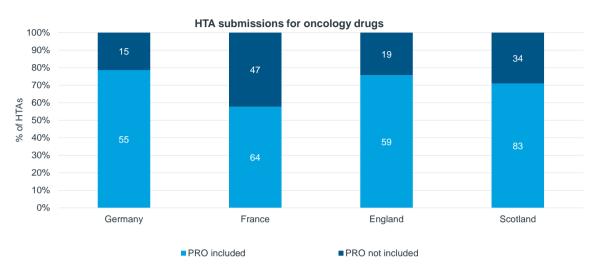
Although PRO evidence is increasingly collected, the use of PRO evidence varies across therapy areas

	Oncology	Rare diseases	Diabetes
HTA submissions with PRO data ¹	70%	48%	27%
Significant HRQoL burden			
Significant therapy AE burden			
"Hard" endpoints poorly defined			
"Hard" endpoints hard to achieve			
Demand for patient-relevant endpoints			
● High relevance			

Scope: Single drug assessments (original, extension of indication, resubmissions) in oncology and diabetes published by HAS, G-BA, NICE and SMC between Jan 2011 and Dec 2016

≣IQVIA ■

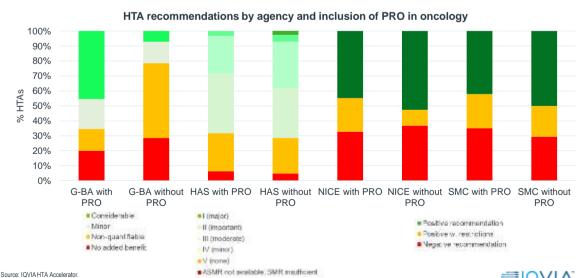
Submissions in France include PRO data less frequently than submissions to the other three HTA bodies



Source: IQVIA HTA Accelerator Scope: Single drug assessments (original, extension of indication, resubmissions) for oncology w/ a recommendation from Jan 2011 to Dec 2016 from 4 HTA bodies (G-BA, HAS, NICE, SMC)



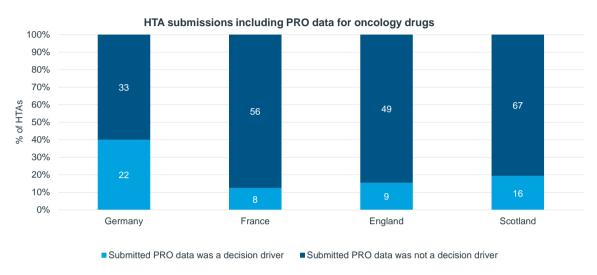
Products with PRO data do not necessarily receive a more favourable recommendation



Scope: Single drug assessments (original, extension of indication, resubmissions) for oncology w/ a recommendation from Jan 2011 to Dec 2016 from 4 HTA bodies (G-BA, HAS, NICE, SMC)

≣IQVIA ■11

Evidence that PRO data influenced the decision was most clear in **Germany**



Source: IQVIA HTA Accelerator

Scope: Single drug assessments (original, extension of indication, resubmissions) for oncology w/ a recommendation from Jan 2011 to Dec 2016 from 4 HTA bodies (G-BA, HAS, NICE, SMC)



The German perspective

Yvonne-Beatrice Boehler

12-NOV-2018 Page: 13

Yvonne-Beatrice Boehler - The German perspective

Agenda

- o Setting the scene General process
- o Facts Methodology to assess PROs
- o Case studies Data driven
- o In a nutshell Learnings

12-NOV-2018

Yvonne-Beatrice Boehler – The German perspective

Technology Arts Sciences TH Köln

Page: 14

Agenda

- o Setting the scene General process
- Facts Methodology to assess PROs
- o Case studies Data driven
- o In a nutshell Learnings



12-NOV-2018 Page: 15 Yvonne-Beatrice Boehler – The German perspective

Assessment 3 months 4 Appraisal 3 months 4 Appraisal 3 months 4 Appraisal 4 Appraisal 5 Committee (G-BA) National Association of Statutory Health Care (IQWIG) First 6 month – methodological viewpoints on PROs Focus Last 6 months – strategic viewpoints on PROs

Agenda

12-NOV-2018

Page: 16

Setting the scene – General process

- o Setting the scene General process
- o Facts Methodology to assess PROs
- o Case studies Data driven
- o In a nutshell Learnings

12-NOV-2018 Page: 17 Yvonne-Beatrice Boehler – The German perspective

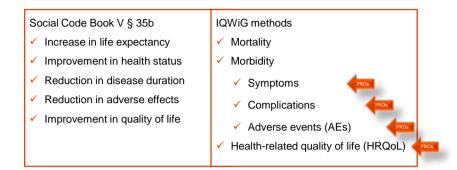
Yvonne-Beatrice Boehler – The German perspective

Technology Arts Sciences TH Köln

Technology Arts Sciences

TH Köln

Patient-relevant benefit - Outcome categories





PROs may qualify for an added benefit against the appropriate comparator therapy in several outcome categories

Inferential statistical thresholds forcelative effect measures

0.90

1.00

All-cause

mortality

0.85

0.95

1.00

b: Risk must be at least 5% for at least 1 of the 2 groups compared.

Outcome category

Non-serious (or non-

N/A

0.80

0.90

severe) symptoms (or

late complications) & AEs

Serious (or severe)

symptoms (or late

complications) &

0.75 & risk ≥5%b

AEs, HRQ6La

12-NOV-2018 Page: 18

Yvonne-Beatrice Boehler - The German perspective

Technology Arts Sciences TH Köln

Facts – Methodology to assess PROs

Dimensions of added benefit

Depending on.

Indicating certainty of Probability

Proof



- Indication
- ✓ Hint

Extent

- Major
- Considerable
- Minor
- ✓ Non-quantifiable added benefit
- No added benefit proven
- ✓ Benefit of drug smaller than benefit of appropriate comparator therapy



a: Precondition (as for all patient-reported outcomes):

use of a validated or established instrument, as well as a validated or established response criterion.

Benefit

Major

Minor

category

Considerable

References 7

12-NOV-2018 Page: 19

Yvonne-Beatrice Boehler - The German perspective

Facts - Methodology to assess PROs

Implications

- ✓ Validation of instruments applied & response criteria to be addressed in the dossier (module 4, methods section 4.2.5.2)
 - Requires validated or established response criterion (e.g. individual minimally important difference [MID])⁸
- ✓ If results are dichotomous (responders/non-responders, relative effect measures)
 - ✓ Clinical relevance of effects is addressed
 - ✓ Extent criteria for added benefit can be applied



MID is key...otherwise...

✓ Clinical relevance

- ✓ Use of standardized mean difference (SMD expressed as Hedges' g)
- ✓ Irrelevance threshold of 0.2, confidence interval of e.g. change from baseline effect estimate must lie completely above⁹
- ✓ Extent criteria
 - ✓ Only non-quantifiable



Examples in upcoming case studies

References 7-

12-NOV-2018 Page: 20 Yvonne-Beatrice Boehler – The German perspective

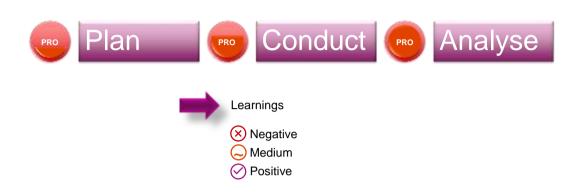
Technology Arts Sciences TH Köln

Agenda

- Setting the scene General process
- o Facts Methodology to assess PROs
- o Case studies Data driven
- In a nutshell Learnings

12-NOV-2018 Page: 21 Yvonne-Beatrice Boehler – The German perspective

Case studies - Data driven



12-NOV-2018

Yvonne-Beatrice Boehler – The German perspective

Page: 22

Technology Arts Sciences TH Köln

Case studies - Data driven



Dulaglutide (2015)

Disease-specific instruments (APPADL, IW-SP, LBSS), IQWiG/G-BA: Questionnaires not accepted, validated populations did not correspond to the target population (e.g. diabetes type 1)

Sofosbuvir/Velpatasvir/Voxilaprevir (2018)

- Used disease-specific instruments (CLDQ-HCV, FACIT-F), IQWiG/G-BA: Questionnaires not accepted, validated populations did not correspond to the target population, CLDQ-HCV content validity questionable
- However SF-36 was accepted, MCS with statistically significant advantage (mean difference, no responder analysis), Hedges'g not completely above irrelevance threshold: No added benefit for this endpoint

Abiraterone (03/2018)

- ✓ A real treasure of examples, negatively & positively
- ✓ EQ-5D VAS health status, FACT-P & one of many BPI-SF Item response criteria accepted
- 🗴 BFI and many BPI-SF response criteria & sensitivity analyses not accepted mean difference applied: Hedges'g not completely above irrelevance threshold: No added benefit for these endpoints, before addendum

Learnings

Negative Medium

Positive

References 10-15

12-NOV-2018

Yvonne-Beatrice Boehler - The German perspective

Page: 23

ISPOR BARCELONA - ISSUE PANEL: PROVE IT WITH PROS

Case studies - Data driven



Bosutinib (08/2018)

EQ-5D (same response criteria as for abiraterone?!), FACT-Leu response criteria & sensitivity analyses not accepted (for both very long IQWiG argumentation lines) - mean difference applied: Hedges'g not completely above irrelevance threshold: No added benefit for these endpoints

Cariprazine (2018 - indication field: relapse prevention)

- ClinRO/PRO PANSS response criteria & sensitivity analyses not accepted mean difference applied: Hedges'g not completely above irrelevance threshold: No added benefit for these endpoints
- √ PSP, response criteria & sensitivity analyses not accepted, mean difference applied: Hedges'g completely above irrelevance threshold: Driver of added benefit

Learnings

- Negative
- Medium

Positive

References 16-19

12-NOV-2018

Yvonne-Beatrice Boehler – The German perspective

Page: 24

Technology Arts Sciences

TH Köln

Case studies - Data driven



Regorafenib (2015)

EORTC QLQ-C30 (symptom scales – morbidity – non-serious) & EORTC QLQ-C30 (functional scales – HRQoL), ITT problem in 2013 assessment (>30% missing), in re-assessment 2015 MMRM still not accepted

Learnings

Negative Medium

Positive

References 20-21

12-NOV-2018

Yvonne-Beatrice Boehler - The German perspective

Page: 25

ISPOR BARCELONA - ISSUE PANEL: PROVE IT WITH PROS

Case studies - Data driven



Abiraterone (03/2018)

- EQ-5D VAS health status responder analysis was assessed as "non-serious" & statistically significant difference did therefore not qualify for an added benefit according to IQWiG methods
- After addendum: BFI Item 3 response criteria (many staggered sensitivity analyses, robust effects) accepted
- FACT-P, BFI Item 3 & BPI-SF Item 3 responder analysis resulted in added benefit, however overall survival data was available

Cariprazine (2018 – indication field: relapse prevention)

- PSP mean difference data was assessed as "non-serious" & not as HRQoL-measure (difference in added benefit thresholds), however, non-quantifiable anyway due to Hedges'g
- After addendum: ClinRO/PRO BARS response criteria accepted no statistically significant differences, no change of initial assessment

Dabrafenib (2015)

- EORTC QLQ-C30 (symptom scales morbidity non-serious) & EORTC QLQ-C30 (functional scales HRQoL), time to deterioration of 10 points responder analyses accepted - added benefit based on HRQoL
- After addendum: EQ-5D VAS responder analyses & sensitivity analyses (also for subgroups) accepted positive change for added benefit for men

Learnings

Negative

Medium

Positive

References 12, 15, 17, 19, 22, 23

12-NOV-2018

Yvonne-Beatrice Boehler - The German perspective

Page: 26

Technology Arts Sciences TH Köln

Agenda

- o Setting the scene General process
- Facts Methodology to assess PROs
- Case studies Data driven
- o In a nutshell Learnings



12-NOV-2018

Yvonne-Beatrice Boehler - The German perspective

Page: 27

In a nutshell - Learnings



Plan

- Choose or develop validated instruments (content validity & psychometric properties)
 - Validated in Disease/subgroup
 - ✓ Subscale/Item validation
 - MID available
- Consider validated generic &/or diseasespecific instruments
- Consider blinded endpoint assessors for ClinROs in open-label trials
- Pre-specify analyses, e.g.
 - Responder-Analyses with different cut-offs
 - Subscale & Item-Analyses according to validation
- Consider early dialogues to inform your viewpoints on your PROs



Conduct

- Implement the data collection correctly
 - Close monitoring avoiding missing data & collecting robust data
 - Consider effective ePRO implementation
- Consider amendments of analysis plans once more validation data is published, available etc. (blinded!)



Analyse

- Follow ITT principle
- Deliver (even if staggered and posthoc) responder & sensitivity analyses, also for relevant subgroups
- Deliver an argumentation on severity of measured symptoms/concepts (relevant for extent category severe vs. nonsevere)
- Support these with change from baseline data
- Address sources of bias
- Prepare for potential data submissions in hearing procedure - and therewith for possible addendum



Be ready to: PROVE it with PROs!

12-NOV-2018 Page: 28

Yvonne-Beatrice Boehler - The German perspective

Technology Arts Sciences TH Köln



The industry perspective

Stefan Holmstrom

HTA GUIDANCE IS NOT DETAILED AND CONSISTENT **ENOUGH FOR THE INDUSTRY TO BE ABLE TO IMPLEMENT IT**

PRO guidance available

PRO guidance in development

No published PRO guidance



In Germany, HTA bodies use specific criteria to assess PRO evidence and it can be very influential in the reimbursement decision (positively and negatively)



EUnetHTA's 2017-2020 work plan includes the development of a joint position on the principles for development and validation of patient reported outcomes



Ex-HAS: "Quality of life is a niceto-have but it would only have a very minor impact on the ASMR rating"



Scottish Medicines

🥭 astellas

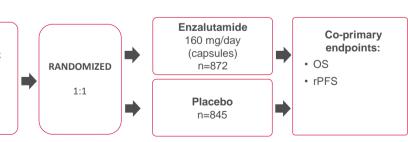
EQ5D is preferred as a utility measure for cost-effectiveness analysis, but no guidance on e.g. data collection

CASE STUDY: XTANDI IN MEN WITH MCRPC NOT YET INDICATED FOR CHEMOTHERAPY

PREVAIL study overview: A Phase 3 trial of enzalutamide after progression on ADT in men with mCRPC

Patient population:

- · 1717 men with progressive mCRPC
- Asymptomatic/ mildly symptomatic
- · Chemotherapy-naïve
- · Steroids allowed but not required



ADT=androgen-deprivation therapy: mCRPC=metastatic castration-resistant prostate cancer: OS=overall survival; rPFS=radiographic progression-free survival

astellas

Source: Beer TM. et al. ASCO-GU 2014: Oral presentation: ClinicalTrials dov. identifier: NCT01212991

MULTIPLE PRO INSTRUMENTS WERE INCLUDED IN THE PREVAIL STUDY

PRO instruments assessment schedule

PRO instrument	Screening	Week 1	Week 5	Week 13	Week 25	Week 37 and every subsequent 12 weeks
BFI		Х				
BPI	Х	Х		Х	Х	
EQ-5D		Х		X	Х	Х
FACT-P		Х	Х	Х	Х	Х

BFI: Brief Fatigue Inventory
BPI: Brief Pain inventory
EQ-5D: European Quality of Life 5-Domain Scale
FACT-P: Functional Assessment of Cancer Therapy - Prostate



THE PRO RESULTS WERE GENERALLY POSITIVE

Adjusted mean change from baseline in FACT-P and EQ-5D at week 61 and BPI-SF worst pain at week 25

	Enzalutamide	Placebo	Treatment difference*	pvake	
FACT-P					
FACT-P total score	-508(-687tn-328)	-30-87 (-13-49 to -8-25)	580 (3.18 to 8.41)	+0-0001	
Physical wellbeing	-261(-368 to -214)	-353(-424to-282)	0.92 (0.21 to 1.63)	0.011	
Functional wellbeing	-195(-249 to-141)	-311(-392to-230)	1 16 (0-35 to 1 97)	0.0050	
Emotional wellbeing	0-19 (-0.18 to 0.57)	-1-03 (-1-63 to -0-43)	1.23 (0.62 to 1.84)	+0-0001	
Social or family wellbeing	0-57 (0-08 to 1-07)	-036 (-098 to 046)	0-84 (0.13 to 1-54)	0.021	
Prostate cancer subscale	-149 (-261to-145)	-318 (-417 to -218)	149 (049 to 349)	9.020	
Prostate cancer subscale pain related	-1-37 (-1-77 to -0-98)	-1.87 (-2-49 to -1.26)	0-50 (-0-12 to 1-12)	0-11	
EQ-50					
Utility index	-0.07 (-0.09 to -0.05)	-010 (-014to-006)	0-03 (-0-00 to 0-07)	0.080	
VAS score	-519(-714to-323)	-976 (-12 61 to -6 92)	4-58 (1-85 to 7-31)	0.0010	
BPLSF					
Worst pain	0.90 (0.64 to 1.17)	130 (100 to 161)	-0.40 (-0.66 to -0.15)	0.0022	

FACT-P
The between-group differences regarding decreases in most scores at week 61 were significantly in favour of enzalutamide

EQ-5D
Enzalutamide had a beneficial effect versus placebo on general health utilities measured by EQ-5D visual analogue scale

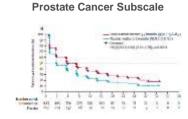
BPI-SF
BPI-SF worst pain deteriorated to a lesser extent in the enzalutamide

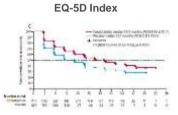


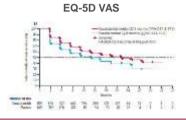
Source: Loriot et al, 2015, Effect of enzalutamide on health-related quality of life, pain, and skeletal-related events in asymptomatic and minimally symptomatic, chemotherapy-naive patients with metastatic castration-resistant prostate cancer (PREVAIL): results from a randomised, phase 3 trial

XTANDI SHOWED A DELAY IN THE TIME TO DETERIORATION IN HRQOL











Source: Loriot et al, 2015, Effect of enzalutamide on health-related quality of life, pain, and skeletal-related events in asymptomatic and minimally symptomatic prostate cancer (PREVAIL): results from a randomised, phase 3 trial

THE PRO EVIDENCE PACKAGE SUBMITTED DIFFERED BY **HTA BODY**

Submitted PRO data

	Endpoint	HAS	G-BA	NICE	SMC
FACT P	Median time to deterioration in FACT P total score				
	FACT P total score (w13 + w25)				
	Change from baseline FACT P total score (w13 + w25)				
	Change from baseline FACT P total score (w61)				
EQ5D	Median time to deterioration in EQ-5D index				
	Median time to deterioration in EQ-5D VAS				
	EQ-5D VAS (w13 + w25)				
	Change from baseline EQ-5D VAS (w13 + w25)				
	Change from baseline EQ-5D index (w61)				
	Change from baseline EQ-5D VAS (w61)				
BPI	Median time to progression of pain at its worst				
	Pain severity and pain interference at (w13 + w25)				
	Change in pain severity and pain interference (w25)				
	Progression of pain at its worst (w25)				



Data submitted in dossier

ACCEPTANCE OF PRO DATA VARIED BY HTA BODY AND OFTEN NO CLEAR ASSESSMENT WAS REPORTED

Acceptance and impact of PRO data in HTAs

	Endpoint	HAS	G-BA	NICE	SMC	
FACT P	Median time to deterioration in FACT P total score	+	✓	?	+	1
	FACT P total score (w13 + w25)		?			
	Change from baseline FACT P total score (w13 + w25)		?			1
	Change from baseline FACT P total score (w61)			+	?	1
EQ5D	Median time to deterioration in EQ-5D index			?	?	1
	Median time to deterioration in EQ-5D VAS	?		?	?	1
	EQ-5D VAS (w13 + w25)		x			1
	Change from baseline EQ-5D VAS (w13 + w25)		x			1
	Change from baseline EQ-5D index (w61)			?	?	1
	Change from baseline EQ-5D VAS (w61)			?	?	1
BPI	Median time to progression of pain at its worst			?	?	Data submitted in dossier
	Pain severity and pain interference at (w13 + w25)		x			? Data not mentioned in repo
	Change in pain severity and pain interference (w25)		X	?	?	x Data not accepted
	Progression of pain at its worst (w25)	+				+ Data acknowledged ✓ Data was a decision driver

THE CRITIQUE OF THE PRO DATA WAS MIXED

There are indications of significant added benefit for serious / severe symptoms / adverse events and health-related quality of life



Scottish Medicines Consortium The median time until decline in the FACT-P global score was also significantly extended by 5.7 months relative to placebo. These outcomes may have particular importance to patients

The difference in collection of BPI data was more than 15% between the two treatments, so no valid statement could be derived



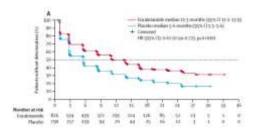
Available data are inconclusive as to the effect of the treatment

It is challenging to convince HTA bodies of patient-relevant improvements, and better guidance from HTA bodies is needed



HOWEVER, THE INDUSTRY MUST ALSO PRESENT EVIDENCE IN A MORE USER-FRIENDLY FORMAT

Time to deterioration in HRQoL based on FACT-P total score



A **total score** can indicate an overall trend but can be confusing and hard to interpret

Transformed scores for items of the physical well-being domain of the FACT-P



Item-level analysis can demonstrate where patients may be performing well or if there are any areas with a significant deterioration





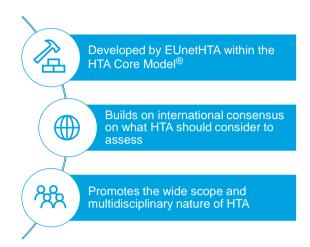
PROs, the HTA Core Model and European HTA

Finn Børlum Kristensen Professor, University of Southern Denmark

EUnetHTA developed the HTA Core Model, which contains nine **HTA** domains

Domains of HTA

Health problem and current use of technology
Technical characteristics
Safety
Clinical effectiveness
Costs and economic evaluation
Ethical analysis
Organizational aspects
Patient and social aspects
Legal aspects



Source: European Network for Health Technology Assessment, EUnetHTA, www.eunethta.eu

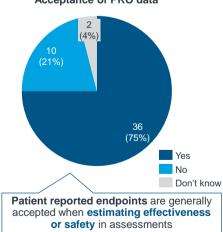
www.scienceandpolicv.dk

PROs may be used to assess the clinical value of new technologies in HTAs

Domains of HTA



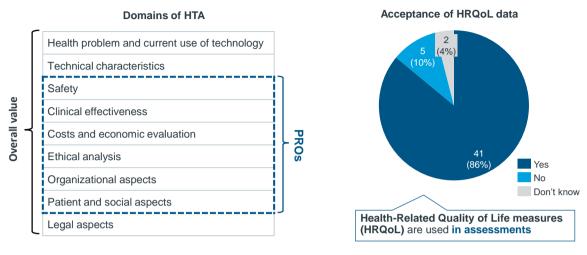
Acceptance of PRO data



Source: Kristensen FB. Mapping of methodologies in EU and Norway, 2018

Finn Børlum Kristensen | Science & Policy | www.scienceandpolicy.dk

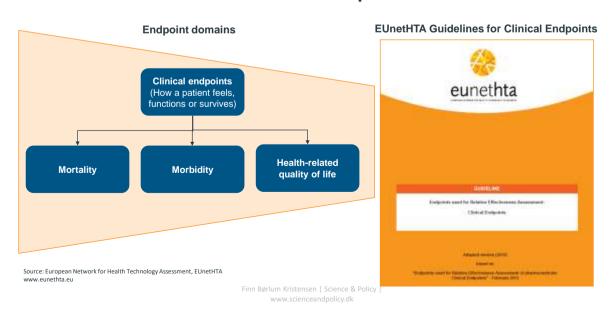
And in assessing the overall value, HRQoL plays an even larger role



Source: Kristensen FB. Mapping of methodologies in EU and Norway, 2018

Finn Børlum Kristensen | Science & Policy www.scienceandpolicy.dk

HRQoL is one of the main categories of endpoints in the EUnetHTA Guidelines for Clinical Endpoints



EUnetHTA guidelines also touch upon the need for HRQoL in cost-effectiveness analyses

EUnetHTA guideline on HRQoL for REA and utility measures

A general recommendation applicable to all types of REA irrespective of their particular purpose, is to require the inclusion of a **disease- or population specific and a generic HRQoL measure** for most adequately capturing the impact of a disease on daily life. In case there is a need for the calculation of QALYs, a utility measure (Time Trade-Off or Standard Gamble) or generic HRQoL, instrument associated with a reference set of utility values (generic utility instrument) is recommended.

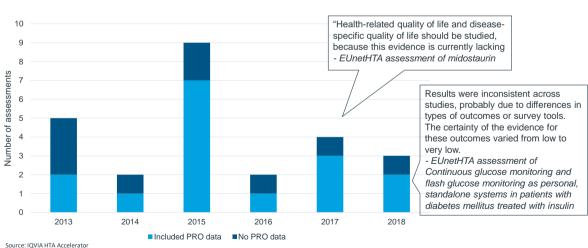
CARRELANT

Source: European Network for Health Technology Assessment, EUnetHTA

Finn Børlum Kristensen | Science & Policy www.scienceandpolicy.dk

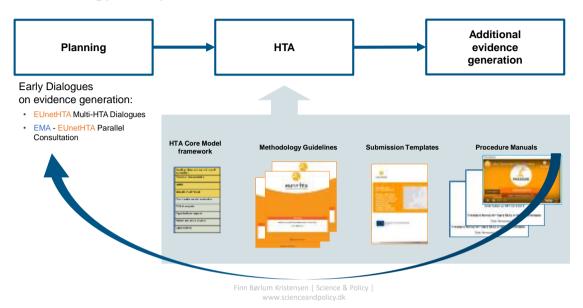
The majority of recent EUnetHTA assessments included PRO data and lack of PRO data was regretted

Inclusion of PRO data in EUnetHTA assessments



Finn Børlum Kristensen | Science & Policy | www.scienceandpolicy.dk

Evidence generation for HTA should take place throughout the technology lifecycle



PROVE IT WITH PROS

- PROs are generally accepted by HTA bodies in assessment of clinical value the degree to which they are decisive differs by HTA body and therapeutic value
- A sound PRO strategy is needed to generate PRO evidence with impact: PROs are not consistently included as endpoint in clinical trials, or data is not adequately collected
- Guidance from HTA bodies should be more clear on the distinction between PROs as outcomes and as utility measures for health economic evaluation
- · Industry must present PRO evidence in a more insightful way
- The new EU joint HTA structure provides an opportunity for more consistency and more guidance for collecting PRO data and inclusion of PROs in HTA submissions

Thank you for your attention! Any comments or questions?











