Patient Reported Outcomes

INTRODUCTION TO CLINICAL RESEARCH
A TWO-WEEK INTENSIVE COURSE, 2010

Milo Puhan, MD, PhD, Associate Professor
Patient-reported outcomes (PRO) is a broad group of outcomes that directly reflect the patients’ perspective on symptoms, functional impairment and health-related quality of life.

Patient reported outcomes for comparative effectiveness research should be selected based on:

- Good responsiveness
- Established test-retest reliability
- Knowledge about their minimal importance difference
- Existing literature
What do patient report outcomes (PRO) measure?
Information about a feature of health/disease that is obtained directly from the patient (with no interpretation by anyone else)
PROs: from simple to very complex measurements

- **Multi-item**
  - **Multi domains**
  - **Measurement Complexity**
  - **Single Item**
    - Pain
    - Fatigue
    - Dyspnea during daily activities
    - Pain during exercise
    - Dyspnea when walking

- **Abstractness of outcome**
  - Health-related quality of life
  - Social Activities
  - Activities of Daily Living
  - Physical Function
  - Symptom Index
Patient-reported outcomes: Symptoms

Uni-dimensional

Dyspnea associated with physical activity

Medical Research Council dyspnea scale

<table>
<thead>
<tr>
<th>Grade</th>
<th>Degree of breathlessness related to activities</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Not troubled by breathlessness except on strenuous exercise</td>
</tr>
<tr>
<td>2</td>
<td>Short of breath when hurrying or walking up a slight hill</td>
</tr>
<tr>
<td>3</td>
<td>Walks slower than contemporaries on level ground because of breathlessness, or has to stop for breath when walking at own pace</td>
</tr>
<tr>
<td>4</td>
<td>Stops for breath after walking about 100m or after a few minutes on level ground</td>
</tr>
<tr>
<td>5</td>
<td>Too breathless to leave the house, or breathless when dressing or undressing</td>
</tr>
</tbody>
</table>
### Patient-reported outcomes: Symptoms

#### Multi-dimensional

**Dyspnea in different situations**

**Chronic Respiratory Questionnaire dyspnea domain**

<table>
<thead>
<tr>
<th></th>
<th>Extremely short of breath</th>
<th>Very short of breath</th>
<th>Quite a bit short of breath</th>
<th>Moderate shortness of breath</th>
<th>Some shortness of breath</th>
<th>A little shortness of breath</th>
<th>Not at all short of breath</th>
<th>Not Done</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Feeling emotional such as angry or upset</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td>2 Taking care of your basic needs (bathing, showering, eating or dressing)</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td>3 Walking</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td>4 Performing chores (such as housework, shopping, groceries)</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td>5 Participating in social activities</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
</tr>
</tbody>
</table>

→ **Summary dyspnea domain score**
Patient-reported outcomes: Health-related quality of life

- Security
- Health
- Infrastructure
- Social environment
- Political stability

Example: Mercer-Studies

- Symptoms
- Physical functioning
- Mental functioning

Example: SF-36
Patient-reported outcomes: Health-related quality of life

Questions do not address directly “quality of life”

Abstract construct derived from underlying questions
Patient-reported outcomes: Health-related quality of life

Uni-dimensional

Please indicate on the scale from 0 to 100 how your health state was during last 7 days.
Patient-reported outcomes: Health-related quality of life

**SF-36**

36 questions

8 domains

- Physical function
- Role physical
- Bodily pain
- General Health
- Vitality
- Social function
- Role emotional
- Mental Health

2 summary scores

Physical component score
Mental component score

Patient-reported outcomes: Health-related quality of life
Disease-specific instruments: *Chronic Respiratory Questionnaire*

20 questions

- Dyspnea
- Fatigue
- Emotional function
- Mastery

4 domains

Total score
PROs: from simple to very complex measurements

- **Multi-item**
  - Multiple domains

- **Measurement Complexity**
  - **Single Item**
    - Pain during exercise
    - Dyspnea when walking
  - **Multi-item**
    - Pain
    - Fatigue
    - Dyspnea during daily activities

- **Abstractness of outcome**
  - Pain
  - Fatigue
  - Dyspnea during daily activities
  - Physical Function
  - Activities of Daily Living
  - Social Activities
  - Health-related quality of life
Development of PROs
The development of PROs is iterative

i. Identify Concepts and Develop Conceptual Framework
   Identify concepts and domains that are important to patients.
   Determine intended population and research application.
   Hypothesize expected relationships among concepts.

ii. Create Instrument
   Generate items.
   Choose administration method, recall period, and response scales.
   Draft instructions.
   Format instrument.
   Draft procedures for scoring and administration. Pilot test draft instrument. Refine instrument and procedures.

iii. Assess Measurement Properties
   Assess score reliability, validity, and ability to detect change.
   Evaluate administrative and respondent burden. Add, delete, or revise items.
   Identify meaningful differences in scores. Finalize instrument formats, scoring, procedures, and training materials.

iv. Modify Instrument
   Change concepts measured, populations studied, research application, instrumentation, or method of administration.
Is there a specific conceptual framework?

A valid PRO should be based on a conceptual framework, which includes:

- **What is the target population?**
  - Any disease?
    - Specific disease?
    - Disease severity?
    - Clinical or general population?

- **What should be measured?**
  - Symptoms?
  - Functional impairment?
  - Health-related quality of life?

- **What’s the purpose of the instrument?**
  - Treatment evaluation?
  - Comparison of patients?
  - Prediction of outcomes?
  - ➔ **evaluative**
  - ➔ **discriminative**
  - ➔ **predictive**
**Conceptual framework**

**Disease:** COPD  
**Severity:** Moderate to severe  
**Population:** Clinical  
**Purpose:** Evaluative

- **Concept:** COPD-specific health-related quality of life  
- **Domains:** Dyspnea, Fatigue, Emotional function, Mastery

**Items:** “Dyspnea when doing chores”  
Answer options: Likert-type 1-7  
Scores: 1-7 per domain  
Recall period: 2 weeks
Steps in the development of PROs

1. Conceptual framework
2. A priori considerations
3. Item and domain identification
4. Item selection
5. Questions
6. Answer options
7. Pilot testing
8. Revision
9. Testing of measurement properties (validation)
10. Adaptation
11. Translations
12. Validation on other populations
Translation of PROs

Testing of measurement properties (validation)

Translations

Forward Translation 1
Forward Translation 2

Professional interpreters

Professional interpreters
Clinicians
Methodologists

Consensus meeting

Translated version 1

Pilot testing

Back translation

Revision

Translated version 2

Validation

Pilot testing
Measurement formats and properties
Reliability and validity

Poor reliability and validity

Good reliability but poor validity

Good reliability and validity
**Test-retest reliability**

Typically 1-2 weeks)

First administration

Second administration

No change (treatments, exacerbations)

- Inappropriate statistic
- Not bad, but does not correct for systematic differences
- Adequate

**Intra-class correlation coefficient**

Between person variance

Between + within person variance

\[ r = 0.9 \]
The validity challenge: No gold standard

- Face validity (expert knowledge) considered insufficient
- Known-group validity: Expected versus observed distributions

![Box plot showing dyspnea severity]

- Construct validity: Expected versus observed correlations with external validation instruments
### Construct validity: Expected versus observed correlations

<table>
<thead>
<tr>
<th>Instrument</th>
<th>Chronic Respiratory Questionnaire</th>
</tr>
</thead>
<tbody>
<tr>
<td>Validation instruments</td>
<td>Dyspnea</td>
</tr>
<tr>
<td>MRC Dyspnea scale</td>
<td>Fatigue</td>
</tr>
<tr>
<td>Self-reported physical activity</td>
<td>Mastery</td>
</tr>
<tr>
<td>Six-minute walk distance</td>
<td>Emotional function</td>
</tr>
<tr>
<td>Exacerbations</td>
<td></td>
</tr>
<tr>
<td>SF-36 physical functioning</td>
<td></td>
</tr>
<tr>
<td>SF-36 mental functioning</td>
<td></td>
</tr>
<tr>
<td>Hospital Anxiety and Depression Scale</td>
<td></td>
</tr>
</tbody>
</table>
# Construct validity: Expected correlations

<table>
<thead>
<tr>
<th>Instrument</th>
<th>Chronic Respiratory Questionnaire</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Validation instruments</strong></td>
<td>Dyspnea</td>
</tr>
<tr>
<td>MRC Dyspnea scale</td>
<td>Moderate (0.3-0.5)</td>
</tr>
<tr>
<td>Self-reported physical activity</td>
<td>Moderate (0.3-0.5)</td>
</tr>
<tr>
<td>Six-minute walk distance</td>
<td>Moderate (0.3-0.5)</td>
</tr>
<tr>
<td>Exacerbations</td>
<td>Moderate (0.3-0.5)</td>
</tr>
<tr>
<td>SF-36 physical functioning</td>
<td>Moderate (0.3-0.5)</td>
</tr>
<tr>
<td>SF-36 mental functioning</td>
<td>Low (&lt;0.3)</td>
</tr>
<tr>
<td>Hospital Anxiety and Depression Scale</td>
<td>Low (&lt;0.3)</td>
</tr>
</tbody>
</table>
Construct validity: Observed correlations

<table>
<thead>
<tr>
<th>Instrument</th>
<th>Chronic Respiratory Questionnaire</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Validation instruments</strong></td>
<td>Dyspnea</td>
</tr>
<tr>
<td>MRC Dyspnea scale</td>
<td>0.62 (&gt;0.5)</td>
</tr>
<tr>
<td>Self-reported physical activity</td>
<td>0.40 (0.3-0.5)</td>
</tr>
<tr>
<td>Six-minute walk distance</td>
<td>0.31 (0.3-0.5)</td>
</tr>
<tr>
<td>Exacerbations</td>
<td>0.38 (0.3-0.5)</td>
</tr>
<tr>
<td>SF-36 physical functioning</td>
<td>0.43 (0.3-0.5)</td>
</tr>
<tr>
<td>SF-36 mental functioning</td>
<td>0.20 (&lt;0.3)</td>
</tr>
<tr>
<td>Hospital Anxiety and Depression Scale</td>
<td>0.13 (&lt;0.3)</td>
</tr>
</tbody>
</table>
Responsiveness: Expected versus observed change

First administration -> Second administration

Treatment of known effectiveness

Dyspnea
- Extreme
- No

Baseline Follow-up

Exacerbations
- Beginning
- Peak
- Recovery

Exacerbation (recovery)
Measures of responsiveness

Treatment of known effectiveness

Dyspnea

Extreme

No

Baseline Follow-up

Comparison baseline and follow-up

Examples
t-tests
Wilcoxon

Effect sizes

Cohen’s effect size

Mean change
SD baseline

Standardized response mean

Mean change
SD change

0.2 to <0.5 = small effect
0.5 to <0.8 = moderate effect
≥ 0.8 = large effect

Many other measures…
Generic vs. disease-specific instruments

281 COPD patients from four sites

Respiratory rehabilitation of 12 weeks

179 patients at follow-up
(102 lost because of SARS outbreak and other reasons)

Generic instruments
- SF-36
- 3 utility instruments

COPD-specific instruments
- Chronic Respiratory Questionnaire
- St. Georges Respiratory Questionnaire

Resp Med 2007; 101, 308
Disease-specific instruments are more responsive

<table>
<thead>
<tr>
<th></th>
<th>Standardized response mean</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>COPD-specific instruments</strong></td>
<td></td>
</tr>
<tr>
<td>CRQ, 4 domains</td>
<td>0.56-0.84</td>
</tr>
<tr>
<td>SGRQ, 3 domains</td>
<td>0.33-0.51</td>
</tr>
<tr>
<td><strong>Generic instruments</strong></td>
<td></td>
</tr>
<tr>
<td>SF-36, 8 domains</td>
<td>0.07-0.37</td>
</tr>
<tr>
<td>3 utility instruments</td>
<td>0.20-0.28</td>
</tr>
</tbody>
</table>
Conclusions responsiveness

Look at responsiveness data

Responsiveness disease-specific
  > disease-specific
    >> generic
    > generic
Minimal important difference
How should we interpret these data?

Respiratory rehabilitation vs usual care for COPD

Study

- Simpson K 1992
- Goldstein 1994
- Güell 1995
- Wijkstra PJ 1996
- Cambach 1997
- Hernandez MT 2000
- Troosters T 2000
- White 2002
- Oh Eg 2003
- Singh V 2003
- Faager 2004
- Combined

Difference CRQ (95% CI)

- 0.69 (-0.2-1.57)
- 0.52 (0.14-0.91)
- 1.14 (0.44-1.83)
- 0.66 (-0.03-1.34)
- 0.95 (0.28-1.62)
- 0.69 (-0.14-1.52)
- 0.70 (0.3-1.1)
- 0.17 (-0.1-0.44)
- 0.95 (0.36-1.54)
- 0.83 (0.04-1.61)
- 0.34 (-0.36-1.34)
- 0.54 (0.391-0.702, p<0.001)
Imagine the following results

![Graph showing comparison between usual care and rehabilitation]

- Combined difference: 0.54 (0.391-0.702, p<0.001)
- Combined difference: 0.44 (0.291-0.602, p<0.001)
- Combined difference: 0.34 (0.191-0.502, p<0.001)
- Combined difference: 0.24 (0.091-0.402, p=0.01)
- Combined difference: 0.14 (0.00-0.302, p=0.05)
- Combined difference: 0.04 (-0.191-0.202, p=0.25)

We need a quantitative interpretation
The minimal important difference

“the smallest difference in the outcome of interest that informed patients or their proxies perceive as important and that may lead to a change in the management”

Related terms:
Minimal clinically important difference (MCID)
Clinically important difference
The MID of the CRQ is 0.5

**Effects of inhaled drugs**

<table>
<thead>
<tr>
<th>Study</th>
<th>Difference CRQ (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simpson K 1992</td>
<td>0.69 (-0.2-1.57)</td>
</tr>
<tr>
<td>Goldstein 1994</td>
<td>0.52 (0.14-0.91)</td>
</tr>
<tr>
<td>Güell 1995</td>
<td>1.14 (0.44-1.83)</td>
</tr>
<tr>
<td>Wijkstra PJ 1996</td>
<td>0.66 (-0.03-1.34)</td>
</tr>
<tr>
<td>Cambach 1997</td>
<td>0.95 (0.28-1.62)</td>
</tr>
<tr>
<td>Hernandez MT 2000</td>
<td>0.69 (-0.14-1.52)</td>
</tr>
<tr>
<td>Troosters T 2000</td>
<td>0.70 (0.3-1.1)</td>
</tr>
<tr>
<td>White 2002</td>
<td>0.17 (-0.1-0.44)</td>
</tr>
<tr>
<td>Oh Eg 2003</td>
<td>0.95 (0.36-1.54)</td>
</tr>
<tr>
<td>Singh V 2003</td>
<td>0.83 (0.04-1.61)</td>
</tr>
<tr>
<td>Faager 2004</td>
<td>0.34 (-0.36-1.34)</td>
</tr>
<tr>
<td>Combined</td>
<td>0.54 (0.391-0.702, p&lt;0.001)</td>
</tr>
</tbody>
</table>
**Methods to determine the MID**

<table>
<thead>
<tr>
<th>Consensus-based</th>
<th>Distribution-based</th>
<th>Anchor-based</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experts reach a consensus (e.g. Delphi methods)</td>
<td>Statistical criteria: How strong is signal compared to noise</td>
<td>External instrument (anchor) with known MID used to estimate MID</td>
</tr>
<tr>
<td></td>
<td>Standardized response mean $\geq 0.5$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cohen’s effect size $\geq 0.5$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Standard error of measurement</td>
<td></td>
</tr>
</tbody>
</table>
Table 1: Changes in HADS and CRQ and Feeling Thermometer scores and correlations of changes

<table>
<thead>
<tr>
<th></th>
<th>HADS depression domain</th>
<th>HADS anxiety domain</th>
<th>HADS total score</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-2.44 (2.79)</td>
<td>-2.02 (2.65)</td>
<td>-2.23 (2.34)</td>
</tr>
<tr>
<td>Changes from baseline to</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>follow-up</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CRQ dyspnea</td>
<td>1.25 (1.17)</td>
<td>-0.24</td>
<td>-0.17</td>
</tr>
<tr>
<td>CRQ fatigue</td>
<td>0.94 (1.25)</td>
<td>-0.43</td>
<td>-0.37</td>
</tr>
<tr>
<td>CRQ emotional function</td>
<td>0.96 (1.07)</td>
<td>-0.42</td>
<td><strong>-0.55</strong></td>
</tr>
<tr>
<td>CRQ mastery</td>
<td>0.94 (1.28)</td>
<td>-0.28</td>
<td><strong>-0.51</strong></td>
</tr>
<tr>
<td>CRQ total</td>
<td>1.02 (0.99)</td>
<td>-0.41</td>
<td>-0.48</td>
</tr>
<tr>
<td>Feeling Thermometer</td>
<td>11.16 (15.82)</td>
<td>-0.23</td>
<td>-0.21</td>
</tr>
</tbody>
</table>
Between 1.41 and 1.68 with anchor-based method

Between 1.17 and 1.40 with distribution-based method

→ 1.5 on scale from 0-21 seems fair estimate

→ Always use different methods since none of them is perfect
Use of the MID

- Interpretation of treatment effects for superiority AND non-inferiority trials

Non-inferiority analysis for interval exercise trial

CRQ Domain
- Dyspnea
- Fatigue
- Emotional function
- Mastery
- Total score

Differences between improvements

<table>
<thead>
<tr>
<th>CRQ Domain</th>
<th>Favours continuous exercise (n=41)</th>
<th>Favours interval exercise (n=44)</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dyspnea</td>
<td></td>
<td></td>
<td>0.07 (-0.53 to 0.38)</td>
</tr>
<tr>
<td>Fatigue</td>
<td></td>
<td></td>
<td>0.02 (-0.41 to 0.32)</td>
</tr>
<tr>
<td>Emotional function</td>
<td></td>
<td></td>
<td>0.08 (-0.48 to 0.32)</td>
</tr>
<tr>
<td>Mastery</td>
<td></td>
<td></td>
<td>0.01 (-0.42 to 0.45)</td>
</tr>
<tr>
<td>Total score</td>
<td></td>
<td></td>
<td>-0.05 (-0.42 to 0.32)</td>
</tr>
</tbody>
</table>

Sample size calculations → detection of important difference

Annals of Internal Medicine 2006;145(11):816
Selection of PRO instruments
PROs: What are you interested in?

**Multi-item**
- Multiple domains

**Measurement Complexity**
- Pain
- Fatigue
- Dyspnea during daily activities

**Single Item**
- Pain during exercise
- Dyspnea when walking

**Abstractness of outcome**
- Health-related quality of life
- Social Activities
- Activities of Daily Living
- Physical Function
- Symptom Index

**Measurement Complexity**
PROQOLID, the Patient-Reported Outcome and Quality of Life Instruments Database

About ProQolid | Demo | Subscribe to ProQolid | Authors | Newsletter | Contact us | Links

Search

Alphabetical list | Generic | Pathology / Disease | Population | Author’s name | Search engine

Highlights

- ProQolid Demo
- Health and Quality of Life Outcomes’s article on ProQolid
- ISOQOL members discount on ProQolid subscriptions
- We’re proud to include the US FDA, the Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWIG) and the National Institute of Health (NIH) among our growing list of ProQolid subscribers

Free access

This level is available to all ProQolid visitors at no charge. For each instrument in the database, you will find 14 categories of basic information (e.g., author, objective, mode of administration, original language, existing translations, pathology, number of items, etc.).

Advanced access (members only)

This level presents a greater degree of practical information on each instrument, most notably the author’s details and contact information, conditions of use, psychometric properties, etc., and, when available, a review copy of the original instrument, its translations, and a user manual.

Database

ProQolid currently contains:
- Descriptions of 647 instruments
- 500 review copies of original instruments
- Review copies of 1010 translations
- Review copies of 170 user manuals
- Descriptions of 82 databases

On-going developments

Include:
- 26% of the seventh update of the instruments described in the database in collaboration with their developers completed
- New information added for each instrument: Complete description of the methodology of development
- 6 new instruments added in 2009

Instruments recently added:
- Standard Evaluation Questionnaire on Pain (SEQ Pain)
- Urinary Symptom Profile (USP)
- QUALIVEEN 30 items (QUALIVEEN-30)
- More

NB: ProQolid content is based on information collected in the literature and/or validated by the authors of the instruments.

The adequacy of study methodology and psychometric properties is not evaluated.

New in ProQolid:

Now available to members, it is mentioned:
- If the questionnaire has obtained a PRO labelling claim from FDA or EMEA
- If the use of the questionnaire is recommended in a clinical research guidance (published by FDA, EMEA, NICE, etc.)
Have standard steps for the development been followed?

Conceptual framework → A priori considerations → Item and domain identification → Item selection → Questions → Answer options → Pilot testing → Revision → Testing of measurement properties (validation) → Adaptation → Translations → Validation on other populations
Choice of patient-reported outcomes for CER

Symptoms, functioning or health-related quality of life?

Ideally matches measures of existing trials $\rightarrow$ Meta-analysis!

Can be measured reliably

Responsive to change $\rightarrow$ disease-specific instruments

Established $\text{MID}$ using valid methods

Is reasonably efficient to measure (budget considerations)
Trial of rehabilitation in patients with COPD

<table>
<thead>
<tr>
<th>Our considerations</th>
<th>MRC dyspnea scale</th>
<th>SF-36</th>
<th>CRQ</th>
<th>St Georges Resp Q</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interested in health-related quality of life</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Matches measures of existing trials</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Can be measured reliably</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Responsive to change</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Established MID</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Efficient to measure</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
# Planning of a CER study in COPD

## Trial of rehabilitation in patients with COPD

<table>
<thead>
<tr>
<th>Our considerations</th>
<th>MRC dyspnea scale</th>
<th>SF-36</th>
<th>CRQ</th>
<th>St Georges Resp Q</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interested in health-related quality of life</td>
<td></td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Matches measures of existing trials</td>
<td>-</td>
<td>+/-</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Can be measured reliably</td>
<td>+/-</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Responsive to change</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+/-</td>
</tr>
<tr>
<td>Established MID</td>
<td>+/-</td>
<td>+/-</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Efficient to measure</td>
<td>+</td>
<td>-</td>
<td>+/-</td>
<td>-</td>
</tr>
</tbody>
</table>
Key messages

- Patient-reported outcomes (PRO) is a broad group of outcomes that directly reflect the patients’ perspective on symptoms, functional impairment and health-related quality of life.

- Patient reported outcomes for comparative effectiveness research should be selected based on:
  - Good responsiveness
  - Established test-retest reliability
  - Knowledge about their minimal importance difference
  - Existing literature

Journals:
Health and Quality of Life Outcomes
Journal of Clinical Epidemiology
Quality of Life Research