



Patient-Reported Outcomes

1- Added value

**2- Instruments : Development
and validation**

3- Application to clinical research

4- Is it useful for something ?



Weak correlation between HRQL and clinician-reported and physiological endpoints

<i>(n = 96)</i>	<i>r</i>	<i>BPQ</i>	<i>CRQ</i>
6-min walk test		0.17	0.07
Pre SaO ₂		0.14	0.17

Symptoms BPQ : Breathing Problems Questionnaire

HRQL CRQ : Chronic Respiratory Disease Questionnaire

→ Variability in exercise capacity contributed to only 3% of the variability in BPQ score

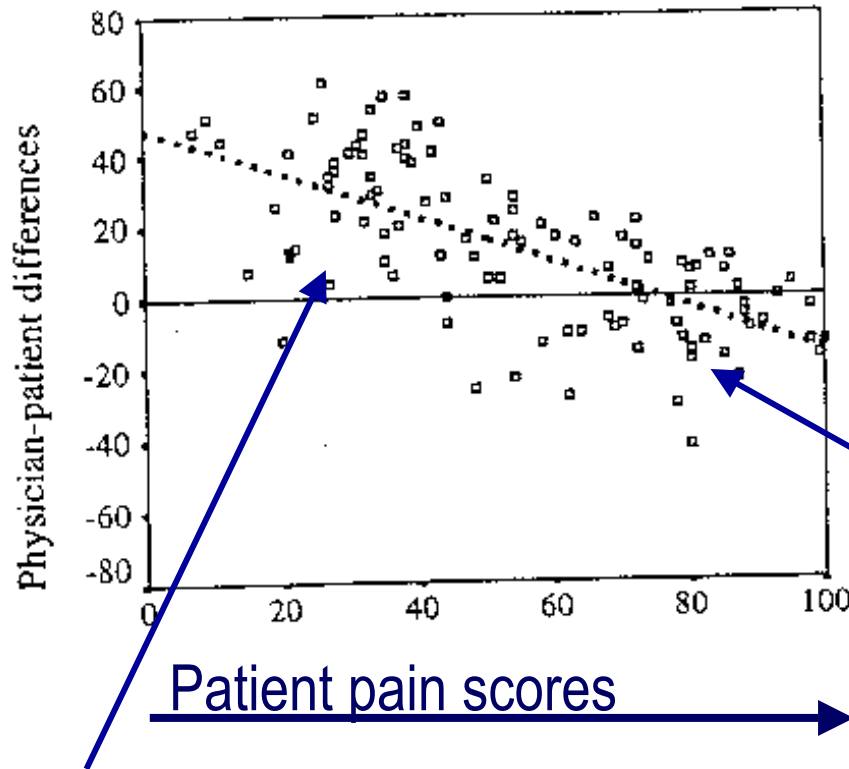


Weak correlation between HRQL & symptoms

- e.g. **Irritable Bowel Syndrome (IBS)**
- The absence of abdominal pain (e.g. during a consultation with a physician) may not be linked with a good HRQL.
The patient :
 - May be anxious not to know when the next bout will occur
 - May be limited in his inter-personal life and his leisure's
 - Constrained to take drugs and to pay attention to food
- The same is true in asthma, migraine, osteoarthritis, acne, heart failure, HIV (e.g. impact of lipodystrophia induced by antiretroviral therapy, even in patients who have not yet the side effect) ...



Weak correlation between patients & physicians



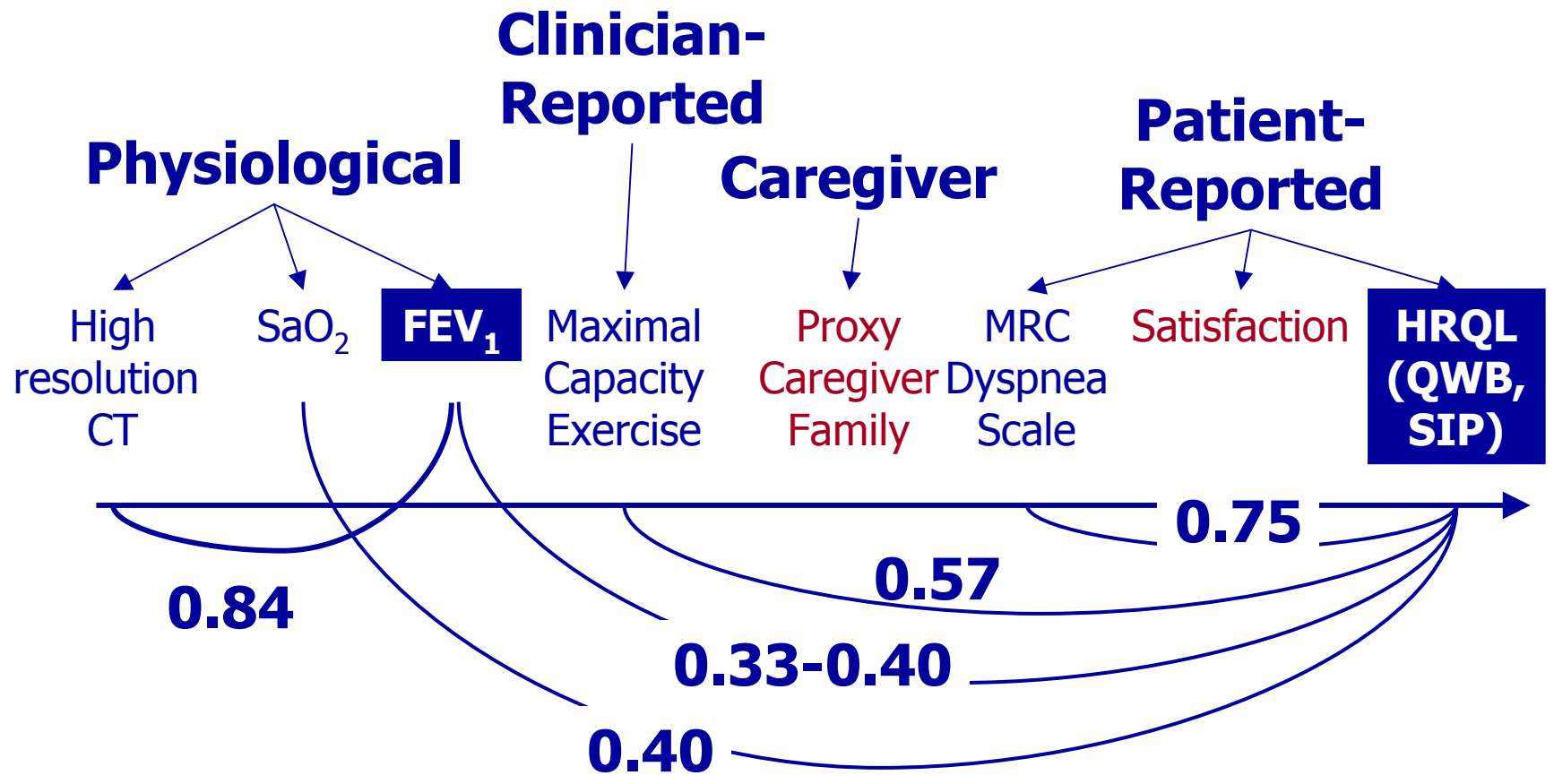
The physician is more disposed to bear the pain of his patient than the patient himself

Tendency of physician to overestimate the pain

Tendency of physician to underestimate the pain



Correlation between different outcomes



High-resolution computer tomography in cystic fibrosis. Dorlochter L, et al. Tidsskr Nor Laegeforen 2002
HRQL as a treatment outcome in patients with cystic fibrosis. Munzenberger PJ, et al. Pharmacotherapy 1999
Quality of life in patients with cystic fibrosis. de Jong W, et al. Pediatr Pulmonol 1997
QWB before and after antibiotic treatment of pulmonary exacerbation in patients with cystic fibrosis. Orenstein DM. Chest 1990



Patient-Reported Outcomes

1- Added value

**2- Instruments : Development
and validation**

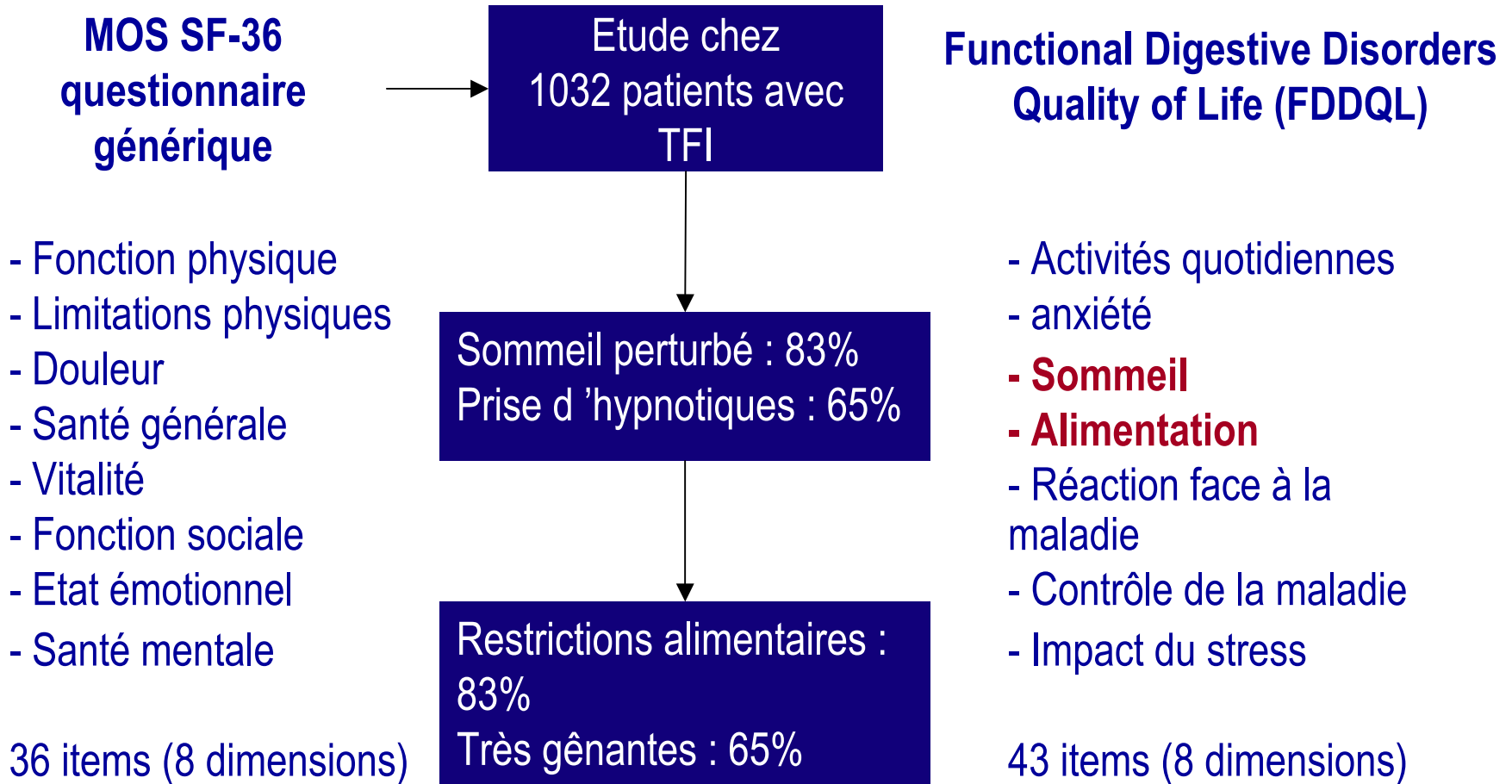
3- Application to clinical research

4- Is it useful for something ?



Which are the arguments in favour of HRQL ?

Generic vs specific questionnaires





How measuring fatigue ?

Are we sure that the questionnaire really measures fatigue ?

Multiple causes

- Lack of rest or exercise
- Improper or inadequate diet
- Psychological stress (depression, anxiety)
- Use of recreational substances
- Anemia
- Abnormalities of the thyroid gland and hypogonadism
- Infections
- Side effects of medications
- Sleep disturbances
- Fever

Fatigue description

- Lack of energy
- Sleepiness
- Tiredness
- Exhaustion
- Inability to get enough rest
- Weakness

Specific
fatigue
questionnaire

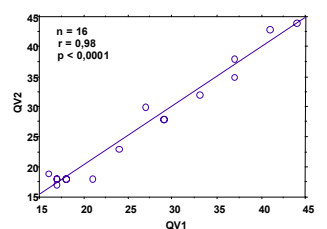
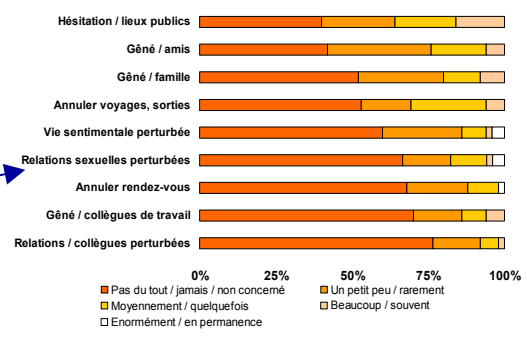
HRQL questionnaire :
must have items related
to fatigue



How to increase the credibility of the HRQL ?

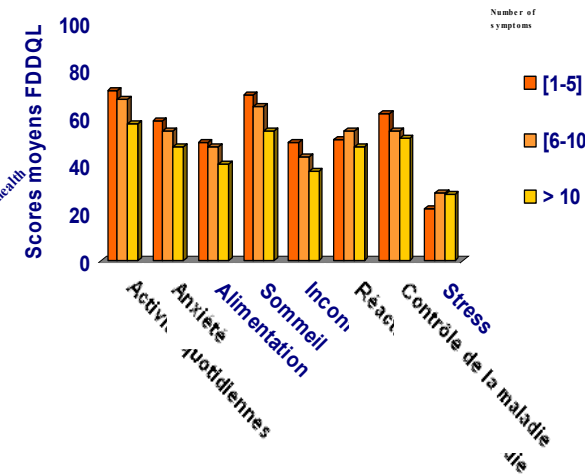
To follow the rigorous procedures of development of HRQL or PRO questionnaires

- Item generation
- Scaling
- Item reduction
- Reproducibility
- Content validity
- Construct validity
- Discriminant validity
- Criterion validity
- Responsiveness
- Cultural adaptation



	Physical functioning	Role physical	Body pain	General health	Vitality	Social functioning	Role emotional	Mental health
Daily activities	0.51	0.63	0.63	0.48	0.52	0.60	0.43	0.48
Anxiety	0.34	0.54	0.46	0.44	0.45	0.43	0.35	0.45
Diet	0.32	0.37	0.43	0.37	0.35	0.50	0.28	0.28
Sleep	0.40	0.41	0.48	0.30	0.39	0.36	0.33	0.36
Discomfort	0.35	0.34	0.49	0.42	0.46	0.44	0.31	0.39
Coping	0.51	0.51	0.54	0.69	0.51	0.54	0.43	0.50
Control	0.24	0.27	0.33	0.40	0.40	0.35	0.25	0.36
Stress	0.06	0.08	0.20	0.15	0.18	0.21	0.20	0.35

Items	Factor	I	II	III	IV
1. Emotional distress					
Discouraged or distressed		0.74			
Frustrated		0.69			
Anxious or upset		0.74			
Worry/fear about health		0.77			
Irritable		0.64			
Worry serious disease		0.75			
2. Sleep disturbance					
No good night sleep			0.83		
Tired-lack of sleep			0.75		
Wake up at night			0.84		
Not waking fresh/rested			0.61		
Trouble getting to sleep			0.73		
3. Food/drink problems					
Discomfort due to eating/drinking				0.71	
Eat smaller meals				0.65	
Unable eat food one likes				0.78	
Food seems unappealing				0.63	
Intolerance to food				0.73	
Avoid certain food/drink				0.74	
4. Physical/social functioning					
Avoid bending over					0.42
Kept from doing things with family/friends					0.68
Difficulty socializing					0.61
Unable carry out daily activities					0.72
Unable to carry out physical activities					0.78





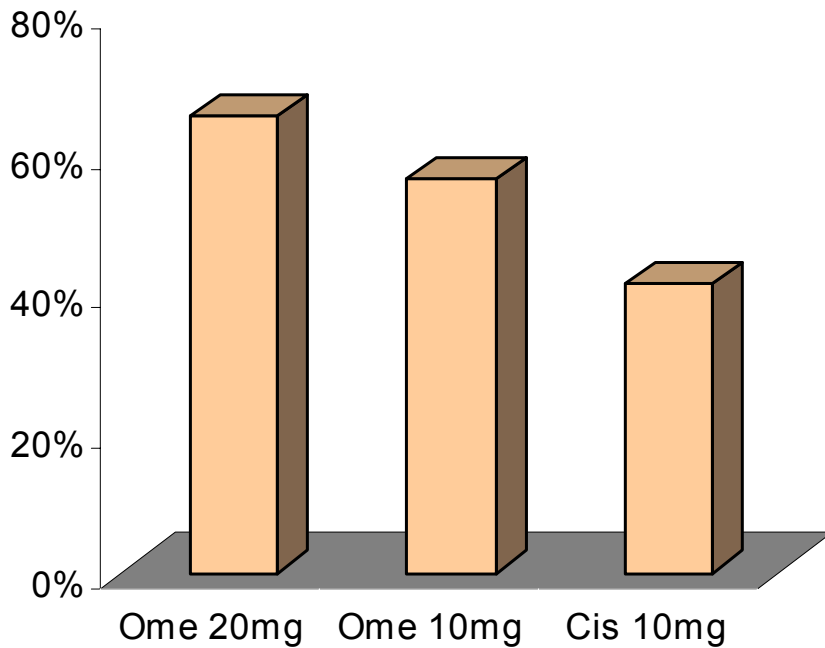
How to increase the credibility of the HRQL ?

Generic questionnaire : responsiveness ?

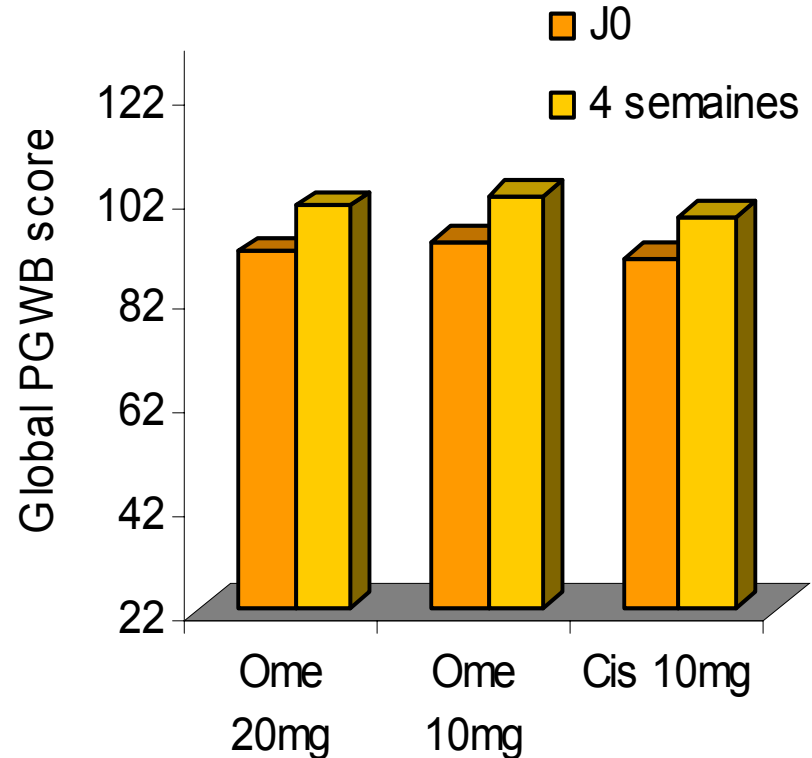
Psychological General Well-Being (PGWB) & GERD

HRQL is not improved by gastro-esophageal reflux disease drugs ?

Résolution des symptômes à 4 semaines



24% difference in pyrosis relief



No difference in PGWB score



How to increase the credibility of the HRQL ?

Cultural adaptation

Forward-backward translation

- **US** **Shoveling the snow**
- **Japan**
- **Norwegian**



How to increase the credibility of the HRQL ?

Cultural adaptation

Forward-backward translation

- **US** **Shoveling the snow**
- **Japan** **Beat futons**
- **Norwegian**



How to increase the credibility of the HRQL ?

Cultural adaptation

Forward-backward translation

- **US** **Shoveling the snow**
- **Japan** **Beat futons**
- **Norwegian** **Going fishing**



Patient-Reported Outcomes

- 1- Added value
- 2- Instruments : Development and validation
- 3- Application to clinical research
- 4- **Is it useful for something ?**



How to increase the credibility of the HRQL ?

Interpreting the results

Beyond the statistical significance ?

	<i>Zk vs PI</i>	<i>p</i>
<i>Daytime symptoms (0 to 3 (severe))</i>	- 0.14	< 0.001
<i>Nighttime awakening (per wk)</i>	- 0.63	< 0.001
<i>β 2 agonist use (puffs/day)</i>	- 0.64	< 0.001
<i>FEV1</i>	0.05	0.331
<i>Morning PEF (BL : 362)</i>	+ 13,1 L/min	< 0.001
<i>Evening PEF (BL : 398)</i>	+ 11,5 L/min	< 0.001
<i>Global AQLQ score (BL : 4.28)</i>	+ 0.26	0.004



How to increase the credibility of the HRQL ?

Interpreting the results

Effect size (Distribution-based approach)

- Dividing a difference between 2 groups or the change over time in one group by the SD at baseline (or the SD of the difference : Standardized Response Mean)

<i>Effect Size</i>	<i>Small</i>	<i>Moderate</i>	<i>Large</i>
Benchmark	> 0.20	> 0.50	> 0.80



How to increase the credibility of the HRQL ?

Interpreting the results

Effect size

- Treatment in claudication (Peripheral Arterial Occlusive Disease)
- Phase III, randomized, double-blind, vs placebo

<i>Criteria</i>	<i>Effect Size</i>
Walking distance	2.13
Specific HRQL questionnaire (CLAUS)	0.48



How to increase the credibility of the HRQL ?

Interpreting the results

Effect size

* Within-subject difference (6-wk minus baseline) in salmeterol group divided by the pooled within-subject SD of change

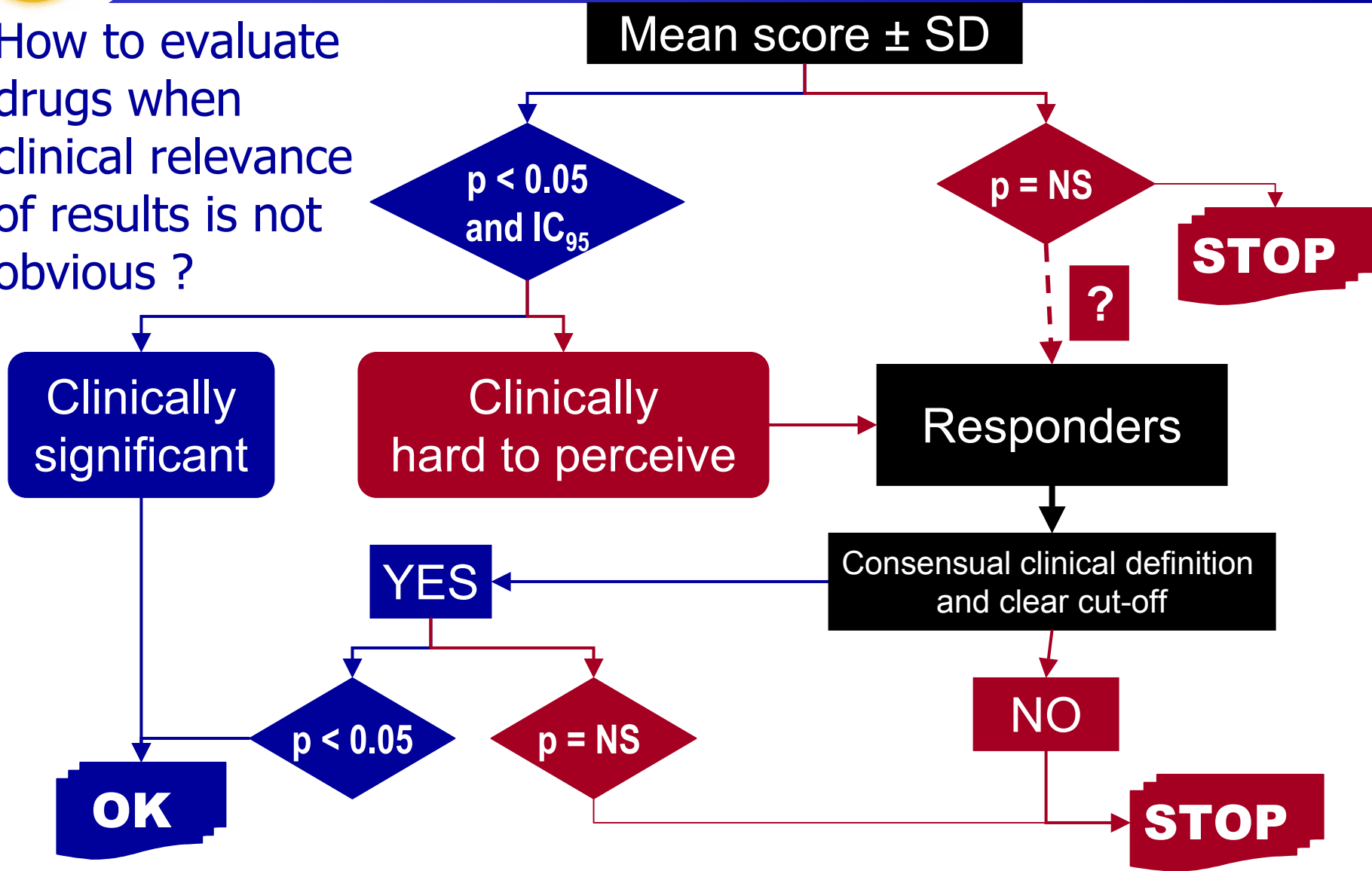
- Randomized, double-blind - 6-wk
- Salmeterol 50 µg or Salbutamol 400 µg bid
- 120 patients randomized
- FEV₁ % predicted : 59%

	<i>Effect size*</i>
AQLQ (32)	0.820
<i>Activities (11)</i>	0.860
<i>Symptoms (12)</i>	0.723
<i>Environment (4)</i>	0.550
<i>Emotions (5)</i>	0.302
	<i>Effect size</i>
LWAQ	0.694
<i>Health knowledge</i>	0.625
<i>Health appraisal</i>	0.333
	<i>Effect size</i>
SIP	0.320



How to increase the credibility of the HRQL ?

How to evaluate drugs when clinical relevance of results is not obvious ?





How to increase the credibility of the HRQL ?

Interpreting the results

Minimal Important Difference (MID) - vs global rating

Answer to the global rating change*	Worse	Better	Interpretation of change	Mean change in HRQL scale (range 1-7)
A very great deal	- 7	+ 7	Large	1.5
A great deal	- 6	+ 6	Moderate	1.0
A good deal	- 5	+ 5		
Moderately	- 4	+ 4		
Somewhat	- 3	+ 3	Small	0.5
A little	- 2	+ 2		
Almost the same	- 1	+ 1		
About the same				

* "Overall, has there been any change in your shortness of breath during your daily activities since the last time you saw us ?"



How to increase the credibility of the HRQL ?

MCID : DEPENDS ON WORDING

Changes in AQLQ symptom-domain anchored to global	Asthma control global		Asthma change global	
Global category	Average	n	Average	n
Worse	- 0.04	3	- 1.05	3
Minimally worse	0.13	49	0.18	11
No change	0.35	102	0.33	45
Minimally improved	0.78	135	0.42	86
Improved	1.48	18	0.85	121

343 patients with mild to moderate asthma

Global asthma control question : “ How well is your asthma controlled ? ”

Global asthma change question : “ Overall has there been any change in your asthma since the beginning of the study ? ”

Response from 0 to 6 (poorly controlled / much worse)



How to increase the credibility of the HRQL ?

Interpreting the results

Minimal Important Difference (MID) - vs global rating

- Wording of the Global Rating
- Characteristics of patients (age, gender...)
- Characteristics of disease (severity ...)
- Setting of the trial, type of intervention
- Cross-cultural differences
- Baseline level of scores ...
- MID may be different for patients improving or worsening

Currently, there is no consensus, whether to be relevant,
MID should be **> 0.5 on a range score from 1 to 7**



How to increase the credibility of the HRQL ?

Interpretation of scores

Minimal Important Difference (MID)

- **The St George's Respiratory Questionnaire (SGRQ)**
- Used in COPD and asthma
- 50 items
- Overall score and 3 subscores (symptoms, activity and impact)
- Range score from 0 (best) to 100 (worst)

*"The minimally clinically important response to treatment is defined as an improvement of **4%** on the separate domains and the total score."*

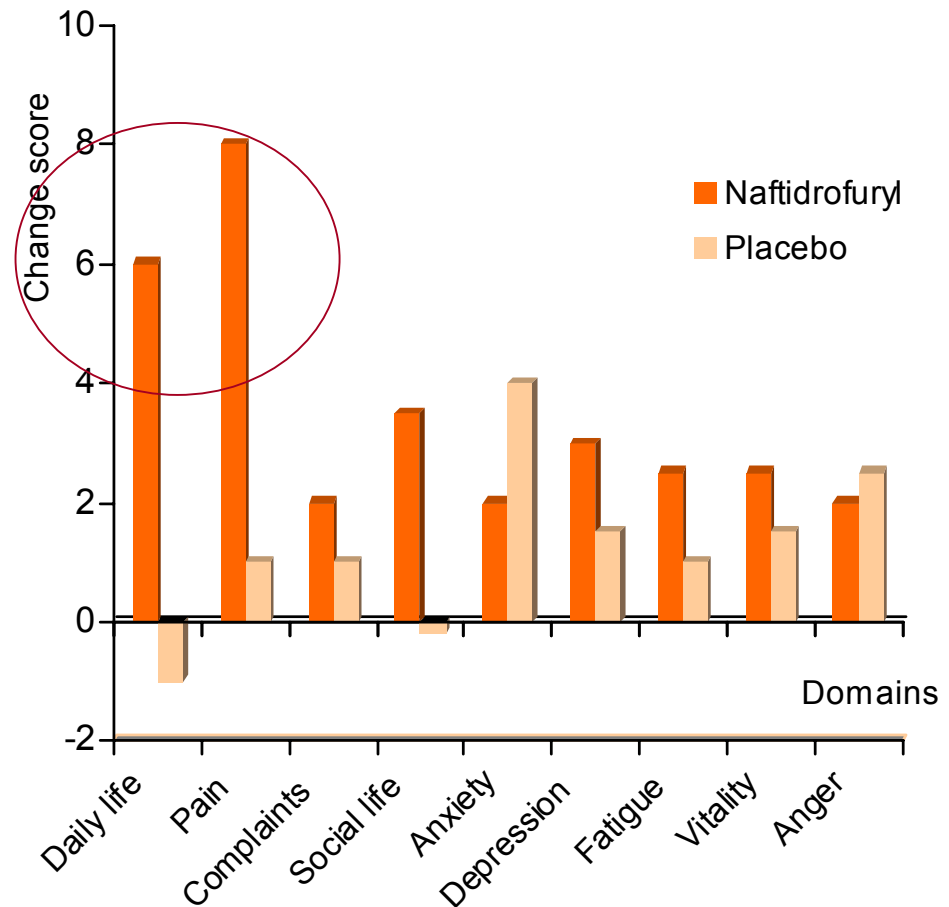


How to increase the credibility of the HRQL ?

Interpreting the results

How many and which domains should improve for a claim ?

- 234 Patients with PAOD
- **HRQL primary endpoint** using the specific questionnaire : CLAU-S (9 domains, 80 items)
- **Results** : 2 domains significantly improved with drug (daily life, $p=0.004$; pain, $p=0.001$)
- **Should the planners have hypothesized that only these 2 domains would improve?**





Example of EMEA PRO CLAIM

VANIQA

eflornithine HCl Cream

Indication : treatment of facial hirsutism in women



Summary of Products Characteristics

5.1 Pharmacodynamic properties

Patient **self-assessment** demonstrated a significantly reduced psychological **discomfort** with the condition, as measured by responses to 6 questions on a visual analogue scale.

Vaniqa significantly reduced how **bothered** patients felt by their facial hair and by the time spent removing, treating, or concealing facial hair.

Patient comfort in various social and work settings was also improved.

Patient **self-assessments** were found to correlate with physician observations of efficacy.

These patient-observable differences were seen 8 weeks after initiating treatment....



CONCLUSION



How to increase the credibility of the HRQL ?

Checklist for designing, conducting and reporting HRQL - PRO in clinical trials

HRQL / PRO objectives

- Added value of HRQL / PRO
- Choice of the questionnaires
- Hypotheses of HRQL / PRO changes

Study design

- Basic principles of RCT fulfilled ?
- Timing and frequency of assessment
- Mode and site of administration...

HRQL / PRO measure

- Description of the measure (items, domains...)
- Evidence of validity
- Evidence of cultural adaptation

Statistical analysis plan

- Primary or secondary endpoint
- Superiority or equivalence trial
- Sample size
- ITT, type I error, missing data

Reporting of results

- Participation rate, data completeness
- Distribution of HRQL / PRO scores

Interpreting the results

- Effect size,
- Minimal Important Difference
- Number needed to treat...