March 2-4 2004, Sigtuna, Sweden

Regulatory issues
What does a regulator look for in the submission files with regard to PROs ?

Olivier CHASSANY, MD, PhD
Medical Manager, Clinical Research Dept (institutional sponsor)
Assistance Publique - Hôpitaux de Paris, France

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…

Patient Reported Outcomes (PRO) and Regulatory Issues : A European Guidance Document for the improved integration of health-related quality of life assessment in the drug regulatory process. Chassany O et ERIQA Working Group. Drug Information Journal 2002.

Statistical analysis plan : PRO multiplicity

Salmeterol / COPD

- Open label
- Salmeterol 50 μg
- or SR Theophylline bid
- Randomized (n = 178)
- Completers (n = 145)
- HRQL (secondary) : SF-36
- Mean changes between baseline and the 4 assessments over time, for each dimension : Student t test

SF-36	*Assessment*
8 (+1) dimensions	3 months
"	6 months
"	9 months
"	12 months
Number of tests	36

(n = ???) in favor of Salmeterol	*Assessment*	*p*
Physical Functioning (PF)	3 months	0.02
Change in Health Perception (HT)	9 months	0.03
Social Functioning (SF)	12 months	0.04

Efficacy, tolerability and effects on HRQL of inhaled Salmeterol in COPD. Di Lorenzo G et al. Clin Ther 1998.

Report of results - full disclosure

A Double-blind comparison of lamotrigine and carbamazepine in newly diagnosed epilepsy with health-related quality of life as an outcome measure. Gillham R et al. Seizure 2000; 9: 375-379.

- More a Side effects than a Quality of life questionnaire : Side Effect and Life Satisfaction
- No description of the content of the 5 domains
- No description of scoring (min-max)
- Evidence of validation ?
- No disclosure of domain scores (Baseline, 48wk)
- Only total score presented (on a graph)
- Relevance of a 4-point difference ?

Interpreting PRO results ?

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

Zafirlukast improves asthma symptoms and HRQL in patients with moderate reversible airflow obstruction. Nathan RA et al. J Allergy Clin Immunol 1998.

Marquis P, Chassany O, Abetz L. A comprehensive strategy for the interpretation of quality of life data based on existing methods. Value in Health 2004 ; 7 : 93-104.

Consistency with other endpoints

- Randomized
- Double-blind
- 437 patients randomized (FEV1 % predicted : 74%)

Fluticasone > zafirlukast	p
FEV1	0.001
Morning PEF	0.004
Evening PEF	0.002
% of symptom-free days	0.007
% of rescue-free days	0.001
Albuterol use	0.001
Comined symptom scores	0.001
Awakening-free nights	0.001
Asthma exacerbation	0.035
AQLQ	**0.001**

Fluticasone propionate versus zafirlukast: effect in patients previously receiving inhaled corticosteroid therapy. Kim KT et al. Ann All Asthma Immunol 2000; 85: 398-406.

Interpretation of results - Effect size

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)

Effect Size	*Small*	*Moderate*	*Large*
Benchmark	> 0.20	> 0.50	> 0.80

Interpretation of results - Effect size

- Randomized, DB, placebo-controlled, parallel groups trial (n = 367)
- Chronic heart failure

HRQoL assessment (primary) : SIP →

1- **POMS** : profile of mood states
2- Inability of patients to carry out regular activities
3- Number of hobbies and whether treatment interfered on them
4- **HSI** : health status index
5- Mahler index of dyspnea-fatigue

	Cilazapril vs placebo		Captopril vs placebo	
	Mean ± SD	**ES**	Mean ± SD	**ES**
Total SIP	0.08 ± 6.6	-0.01	0.56 ± 6.5	0.09
Physical dim.	0.73 ± 6.1	**0.12**	0.87 ± 6.1	**0.14**

Bulpitt et al Quality of life in chronic heart failure: cilazapril and captopril versus placebo. Heart 1998; 79: 593-8.

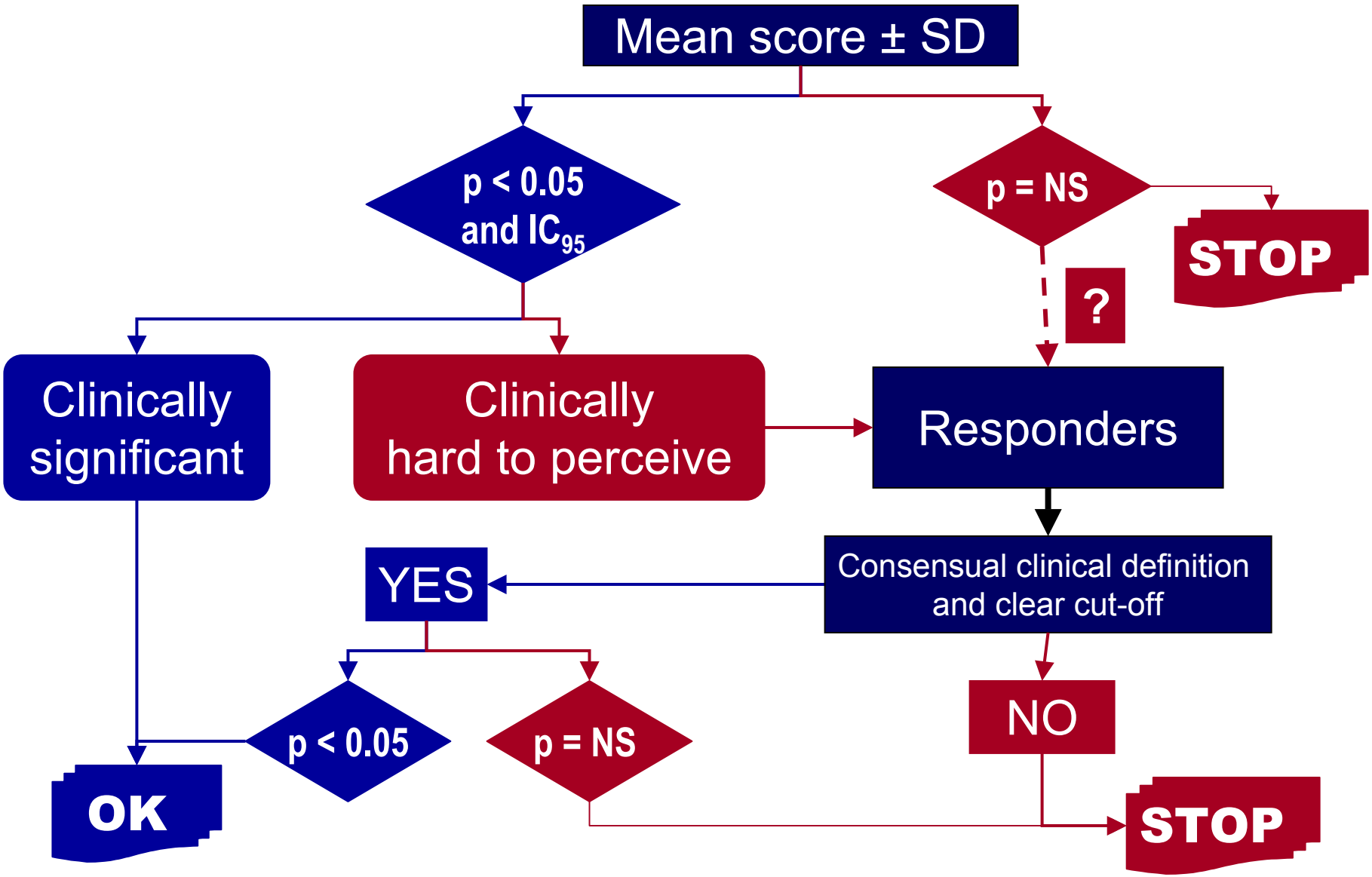
Interpretation of results - Effect size

GERD	Treatment group		Difference	ES
PGWB global score*	OME	RAN		
(Revicki, Dig Dis 1998)	82.5	78.8	3.7	0.22
PGWB global score*	OME	PLA		
(Havelund, Am J Gastro 1999)	103.9	100.6	4.5	0.26
GSRS global score**	OME	RAN		
(Festen, Am J Gastro 1999)	12.3	10	3.3	0.30

range score : * (22-132), ** (5-35)

Effect Size	No change	Small change (non pertinent)	Moderate change	Large change
	< 0.20	0.20-0.50	0.50-0.80	> 0.80

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



Minimal Important Difference (MID) or change

MID obtained from comparison with a 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 A good deal Moderately	- 6 - 5 - 4	+ 6 + 5 + 4	Moderate	1.0
Somewhat A little	- 3 - 2	+ 3 + 2	Small	**0.5**
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 ?"

Guyatt GH, Juniper EF. Several publications

Minimal Important Difference (MID) or change

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

n = 343 (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 ?"**

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

Barber BL et al. Qual Life Res 1996.

Minimal Important Difference (MID)

MID obtained from comparison with a Global Rating may be different according to :

- Wording of the Global Rating
- Improvement vs. worsening
- Characteristics of patients (age, gender…)
- Characteristics of disease (severity …)
- Setting of the trial, type of intervention
- Cross-cultural differences
- Baseline level of scores …

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

Impact of the global on patient perceivable change in an asthmatic specific QOL questionnaire. Barber BL et al. Qual Life Res 1996.

Number needed to Treat (NNT)

- derived from the difference of responders (patients who improved their HRQL score > MID) between groups

Chronic Respiratory Questionnaire	***mean Δ***	NNT to have 1 patient receive at least a small benefit	NNT to have 1 patient receive at least a moderate or large benefit
Dyspnea	0.61	4.1	5.8
Fatigue	- 0.63	4.4	6.9
Emotional function	- 0.64	3.3	6.3
Mastery	0.05	2.5	2.8

Prospective randomized controlled trial of rehabilitation

84 subjects completed

Intervention : 2 months of impatient rehabilitation followed by 4 months of outpatient supervision

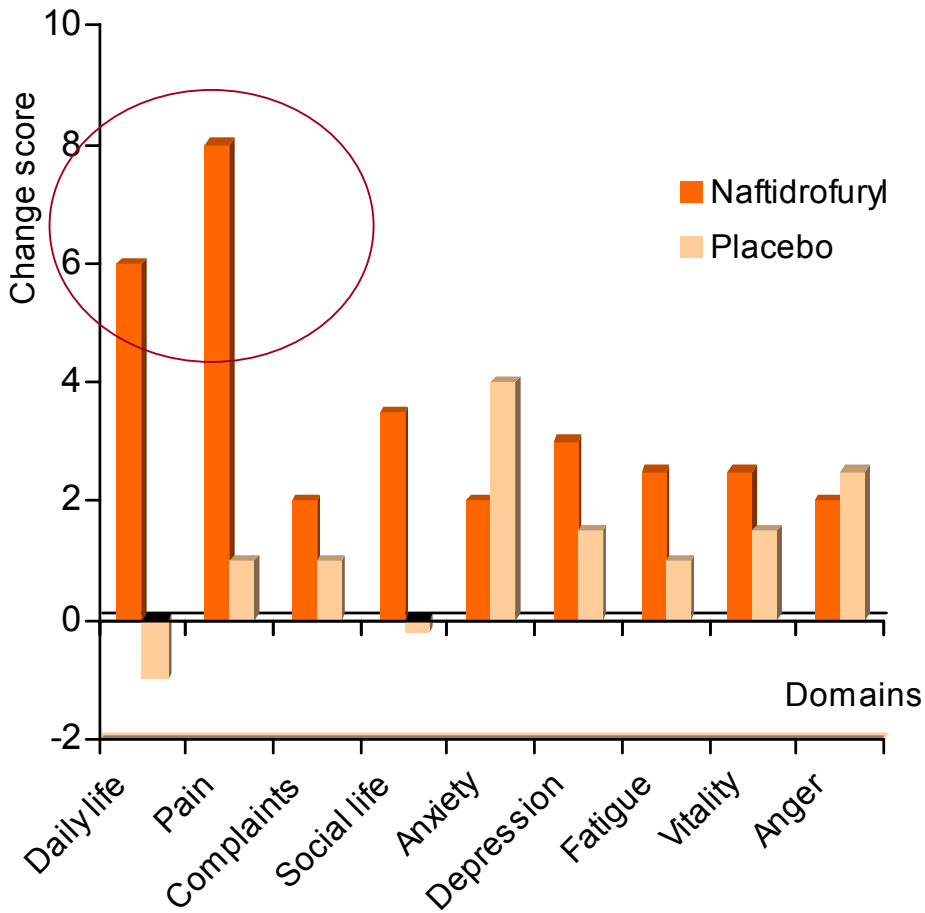
Economic analysis of respiratory rehabilitation. Goldstein RS et al. Chest 1997; 112: 370-9.

Number needed to Treat (NNT)

Study		Drug	Treatment duration	Criterion	NNT
Woscops	NEJM 1995	Pravastatin	5 yrs	Mortality (primary prevention)	**111**
4S	Lancet 1994	Simvastatin	5.4 yrs	Morality (secondary prevention)	**30**
LIPID	NEJM 1998	Pravastatin	6.1 yrs	Morality (secondary prevention)	**32**
Left Ventricular dysfunction	J Am Coll Cardio 1994	Enalapril	41 wks	Mortality	**22**
MIRACL	JAMA 2001	Atorvastatin	16 wks	Composite score	**38**
CAPRIE	Lancet 1996	Clopidogrel	1 yr	Composite score	**196**
MUCOSA	Ann Intern Med 1995	Misoprostol	6 months	Severe gastrointestinal complications (NSAID)	**263**

How many and which PRO domains should improve for a claim ?

- 234 Patients with Peripheral Arteriopathy Occlusive Disease (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?**



The effects of naftidrofuryl on quality of life. Liard F et al. Dis Manage Health Outcomes 1997.

How many and which PRO domains should improve for a claim ?

- 90 (6 x 15) statistical tests
- Difference of **0.2** (range 1-7) at 3 months
- No difference at 12 months

Abstract "Aerobic group-training of elderly patients recovering from an acute coronary event beneficially influences physical fitness and several parameters expressing quality of life"

	J3	J12
Symptoms		
- Chest pain	NS	NS
- Shortness of breath	<0.05	NS
- Dizziness	NS	NS
- Palpitation	<0.05	NS
- Cognitive ability	NS	NS
Alertness	NS	NS
Quality of sleep	NS	NS
Physical ability	NS	NS
Daily ability	NS	NS
Depression	NS	NS
Self perceived health	NS	NS
Ladder of life: future	NS	NS
Fitness	<0.05	NS
Physical activity	<0.01	NS

Stahle A et al. Improved physical fitness and HRQL following training after acute coronary events. Eur Heart J 1999.

How many and which PRO domains should improve for a claim ?

Interpretation - $CI_{95\%}$

RQLQL	*Δ*	*IC95%*	*p*
Sleep	0.3	- 0.7 - 1.7	0.30
Non-hay fever symptoms	0.7	0.0 - 1.6	**< 0.05**
Practical problems	1.0	- 0.001 - 2.3	0.07
Nasal symptoms	1.0	**0.25 - 1.75**	**0.01**
Eye symptoms	1.3	**0.25 - 2.3**	**0.008**
Activities	1.0	0.0 - 1.7	**0.03**
Emotions	0.8	0.0 - 1.75	**< 0.05**
Overall HRQL	0.8	0.18 - 1.5	**0.02**

Randomized - double-blind - placebo-controlled - parallel group

44 patients randomized (37 completed)

Rhinoconjunctivitis Quality of Life Questionnaire (RQLQ) : 0 - 6 (severe)

0.5 : minimal clinically relevant difference (Juniper …)

Walker SM et al. J All Clin Immunol 2001.

How many and which PRO domains should improve for a claim ?

Broad HRQL claim ?

Unlikely, unless

- most of scales of HRQL questionnaires improved
- consistency with standard criteria

Specific-domain claim ?

- If pre-specified
- If consistency with standard criteria
- If evidence of clinical relevance

Where ? Indications or Pharmacodynamic properties chapter ?

ADVAIR experience in asthma

- Juniper-Guyatt's "Asthma Quality of Life Questionnaire"
- AQLQ was administered at day 1 and week 12 (or endpoint, for patients terminating early)
- **Minimal Important Difference = 0.5 for overall score and for individual domains**

Results - Change from Baseline to Endpoint

	Placebo	Advair	Salm	FP
AQLQ global	-0,33	0,99	-0,03	0,56
activity	-0,13	0,99	-0,06	0,74
symptoms	-0,51	1,04	-0,08	0,55
Emotion	-0,45	1,07	0	0,42
Environ.	-0,14	0,87	0,14	0,45

- Results appeared consistent, were not driven by any single domain and were replicated in another trial

ADVAIR : combination Salmeterol + Fluticasone Propionate

5 key issues for Drug Approval Process

HRQL (and PRO) to be considered as a credible criterion if there is enough evidence (in the file) about the :

1- Added-value of HRQL/PRO with respect to other criteria
2- Psychometric properties of the HRQL/PRO instruments
3- International validation of the HRQL/PRO instruments
4- Adequacy of the statistical analysis plan
5- Clinical significance of observed changes

Meeting with representatives of AFSSAPS, EMEA and ERIQA Working Group, Paris, 1999

olivier.chassany@sls.ap-hop-paris.fr

Why there are so few HRQL mention in labelling ?

- In a recent past and overall completed, the poor quality of the clinical trials having evaluated PRO and especially HRQL, which left a persistent feeling of mistrust

- The problem of the exact place of PRO as an endpoint :
  - **Is it an efficacy, tolerance, or utility endpoint ?**
  - **Can PRO be a primary endpoint, and in which diseases ?**
  - **Or shall PRO always be relegated as a secondary endpoint and thus be considered by some regulators as inevitably less rigorous ?**

Why there are so few HRQL mention in labelling ?

- The lack of experience and training of the reviewers and regulators

- The fears (legitimate) of the regulatory authorities to officially acknowledge the PRO and to take into account a subjective criterion by nature :
  - **Whose clinical interpretation remains difficult**
  - **Whose good practices of advertising remain to be specified in a market where competition is rough**
  - **Without counting the possibility for a drug which would have shown a substantial benefit on HRQL/PRO, to have claim in terms of rate of refunding, or price**

What can one wish for the future ?

- Training of reviewers and regulators to HRQL & PRO

  WORKMAT : Educational Program for Reviewers

- Appropriation and adaptation by regulatory agencies of the published recommendations

  Guidelines FDA

  European Position Paper (EWP) ?

- Questionnaires constantly in adequacy with the beneficial and harmful effects of the new treatments

- Choice among the various questionnaires, of those which have the best psychometric properties (responsiveness)

- That HRQL and PRO be part of the daily medical-decision making