

# Introduction to Patient-Reported Outcomes (PROs)

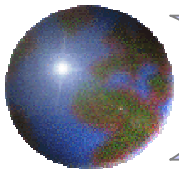
March 2-4 2004, Sigtuna, Sweden

## **Patient-Reported Outcomes in Phase III Regulatory Requirements**

Regulatory issues and best practice

**Olivier CHASSANY, MD, PhD**

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

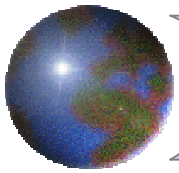


# Why there are so few HRQL mention in labelling ?

## Drug Approval Process

### Major biases encountered in reviewing dossiers

- No justification of HRQL choice
- No evidence of questionnaire validation
- No objective of HRQL changes
- No justification of sample size
- No description of the follow up of patients
- No clear handling of missing data
- Not all patients are analysed
- No correct presentation of results
- No adjustment for multiple comparisons
- No interpretation of results



# 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

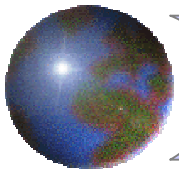
- Primary or secondary endpoint
- Superiority or equivalence trial
- Sample size
- ITT, type I error, missing data

## Reporting of results

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

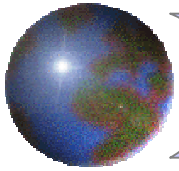
## Interpreting the results

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



## Define the conditions for which the measurement of HRQL/PRO in clinical trial is useful

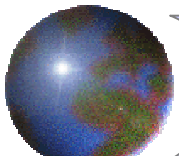
- Patient's self-report is the primary or sole indicator of disease activity, e.g. dermatological disorders (psoriasis, acne), erection dysfunction
- No objective marker or several possible markers of disease activity (migraine, osteoarthritis, asthma, menopause, heart failure)
- Disease expressed by many symptoms (IBS)
- To ensure that treatments prolonging survival (AIDS), do not adversely affect patients' lives due to morbidity, functional or psychological impairments or side effects
- The treatment does not seem to improve survival (cancer, rheumatoid arthritis, Parkinson's disease), but it could improve HRQL, by reducing pain, anxiety, level of stress or by improving the functional status.



# What is not quality of life ?

The abuse of the term HRQL in some clinical trials, whereas the questionnaire measured anything else

- A listing of symptoms or of **side effects** cannot claim to measure HRQL
- The following concepts cannot alone explore all HRQL:
  - **physical or intellectual performance scale**
  - **handicap or functional incapacity scale**
  - **anxiety or depression scale**
  - **tiredness or pain scale**
  - **symptom bother scale.**



# Choice of instrument - What are we measuring ?

## Different PROs : example of HIV/AIDS

### Fatigue

Items related to intensity, circumstances, and consequences of fatigue.

*Psychometric properties of the HIV-related fatigue scale. Barroso J et al. J Assoc Nurses AIDS Care 2002.*

### Cognitive function

Dutch four-item MOS-HIV cognitive functional status subscale.

*The importance of cognitive self-report in early HIV-1 infection: validation of a cognitive functional status subscale. Knippels HM et al. AIDS 2002.*

### Treatment satisfaction

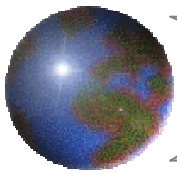
Weak correlation between severity of side-effects and score of satisfaction concerning these side-effects ( $r = 0.18$ )

*Validation of the HIV treatment satisfaction questionnaire. Woodcock A et al. Qual Life Res 2001.*

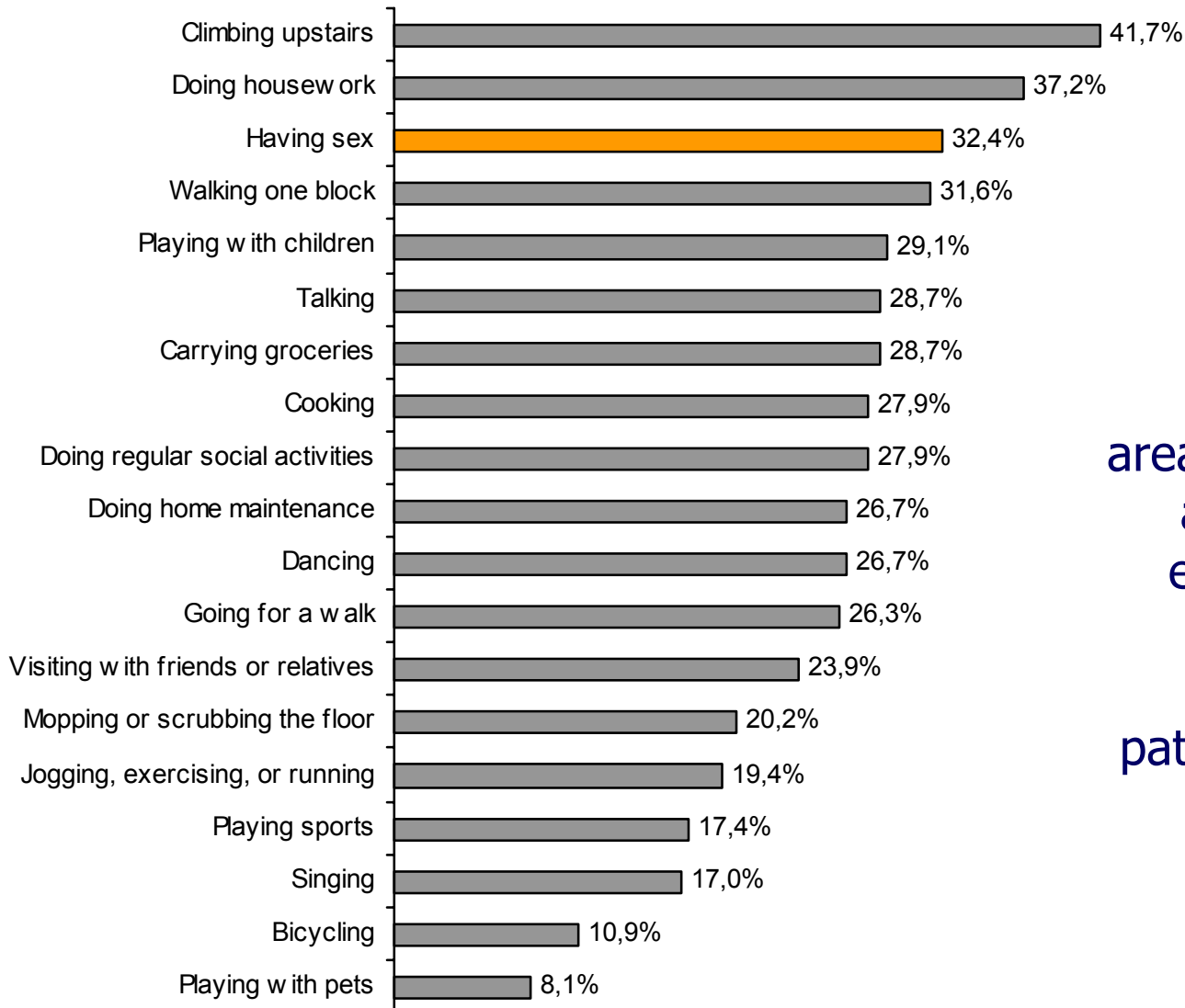
### Doctor-Patient satisfaction

Satisfaction among HIV-infected patients was not associated with QOL

*The doctor-patient relationship and HIV-infected patients' satisfaction with primary care physicians. Sullivan LM et al. J Gen Intern Med 2000.*

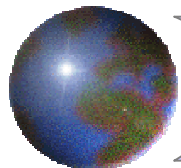


# Choice of a PRO questionnaire - Importance of the sample included during the validation process



Importance of various areas of limitations due to asthma among Harlem emergency department users (n =247) mostly Afro-american patients with a low socio-economic status and a lower compliance

Asthma-related limitations in sexual functioning: an important but neglected area of quality of life.  
Meyer IH, et al. Am J Public health 2002; 92: 770-772.



# Who measures Well-Being ? Clinicians ??

**N = 30**

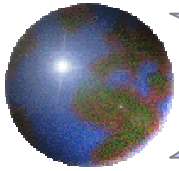
	Indapamide		Captopril	
	before	after	before	after
Personal presentation	0	0	0.12	0.12
Communication	0	0	0	0
Muscular or sensory asthenia	1.21	1.07	1.62	1.06
Functional disturbances	0.36	0.28	0.62	0.25
Sleep disorders	1.43	1.00	1.37	1.31
Mental asthenia	0.14	0.43	0.37	0.31
Emotional lability	0.50	0.28	0.56	0.50
Anxiety	1.86	1.50	1.24	0.81
Will	0.28	0.28	0.43	0.31
Personality	0.50	0.43	0.87	0.50
Well-being	0.72	0.86	0.87	0.87

Figure 1. General well-being assessment by the investigator. Differences in scoring from baseline to the final visit.

Analysis of Well-Being between indapamide and captopril.

Lacourciere Y. Am J Med 1988; 84: 47-51.





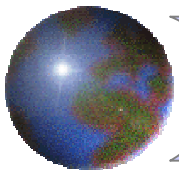
# Who should fill-in questionnaire ?

In studies evaluating sexual impairment induced by antihypertensive treatment in male patients, the answers given to nurses, by patients themselves and by their spouses were quite different...

## Rate of sexual dysfunction

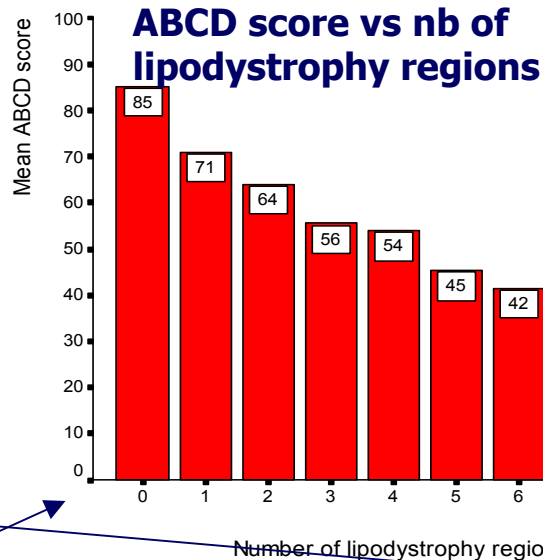
---

Nurses	Low
Patients	Moderate
Patients (palm pilot)	Higher
Spouses	<b>Very high</b>



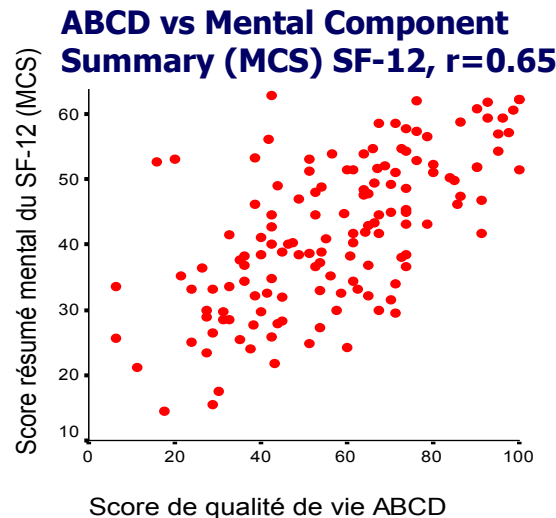
# To follow the rigorous procedures of development of HRQL or PRO questionnaires

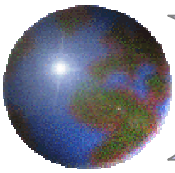
- Item generation
- Scaling
- Item reduction
- Reproducibility
- Content validity
- Construct validity
- Discriminant validity
- Convergent validity
- Responsiveness
- Cultural adaptation



### Factorial analysis ABCD Score

ABCD 20 items	Factor			
	1	2	3	4
a	,723	,084	,284	,177
b	,529	,067	,427	,293
c	,696	,359	,152	,290
d	,580	,488	,149	,318
e	,625	,143	,471	,096
f	,684	,118	,347	-,105
g	,609	,195	,381	,125
h	,767	,417	-,050	,089
i	,181	,323	<b>,728</b>	,132
j	,387	<b>,697</b>	,369	,104
k	,110	,293	<b>,740</b>	,119
l	,174	<b>,732</b>	,317	,000
m	,181	<b>,775</b>	,298	,121
n	<b>,542</b>	<b>,611</b>	-,078	,358
o	,195	<b>,731</b>	,265	,249
p	,378	<b>,490</b>	,123	,478
q	<b>,778</b>	,412	-,101	,290
r	,149	,136	<b>,505</b>	,221
s	,241	,247	,339	<b>,662</b>
t	,100	,089	,166	<b>,821</b>





# How measuring fatigue ? Identification of concepts

## Multiple causes

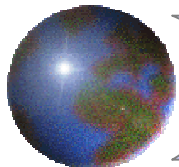
- Lack of rest or exercise
- Improper or inadequate diet
- Psychological stress (depression, anxiety)
- Use of recreational substances
- Anemia
- Abnormalities of the thyroid gland and hypogonadism
- Infections
- Side effects of medications
- Sleep disturbances
- Fever

## Fatigue description

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

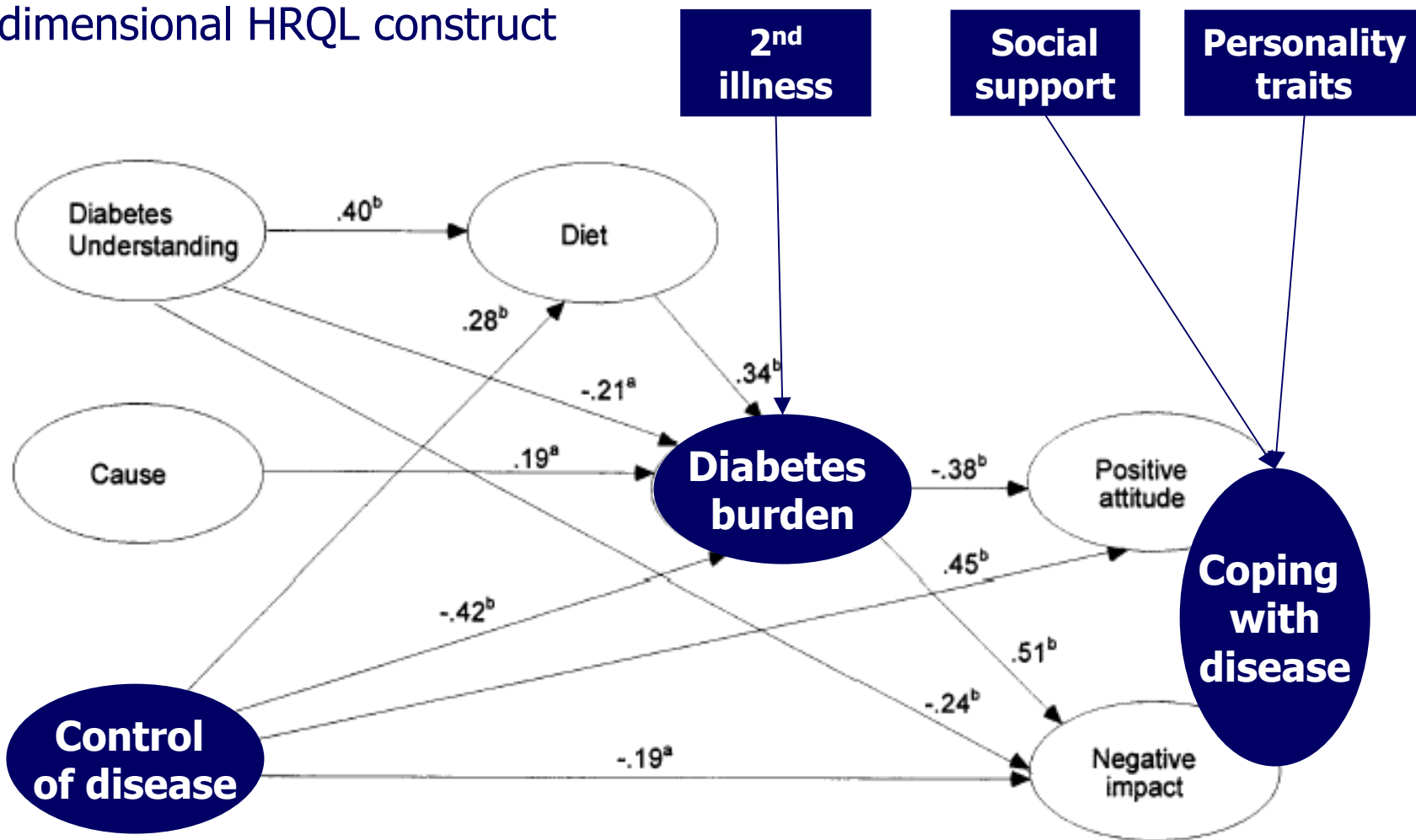
Specific  
fatigue  
questionnaire

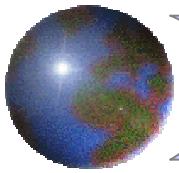
HRQL questionnaire :  
must have items related  
to fatigue



# Determinants of the Quality of Life

Various factors involved in the multidimensional HRQL construct





# Items about DIET can express different concepts

## Input of patients in item generation is critical

**Diabetes** --> Cause --> **Food** --> consequence --> **DIET**

I am able to keep my diet regimen under control

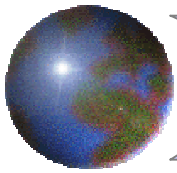
**Control of disease / self-management**

My diabetes and its treatment (e.g. diet) keeps me going out with friends / to restaurant / as much as I want

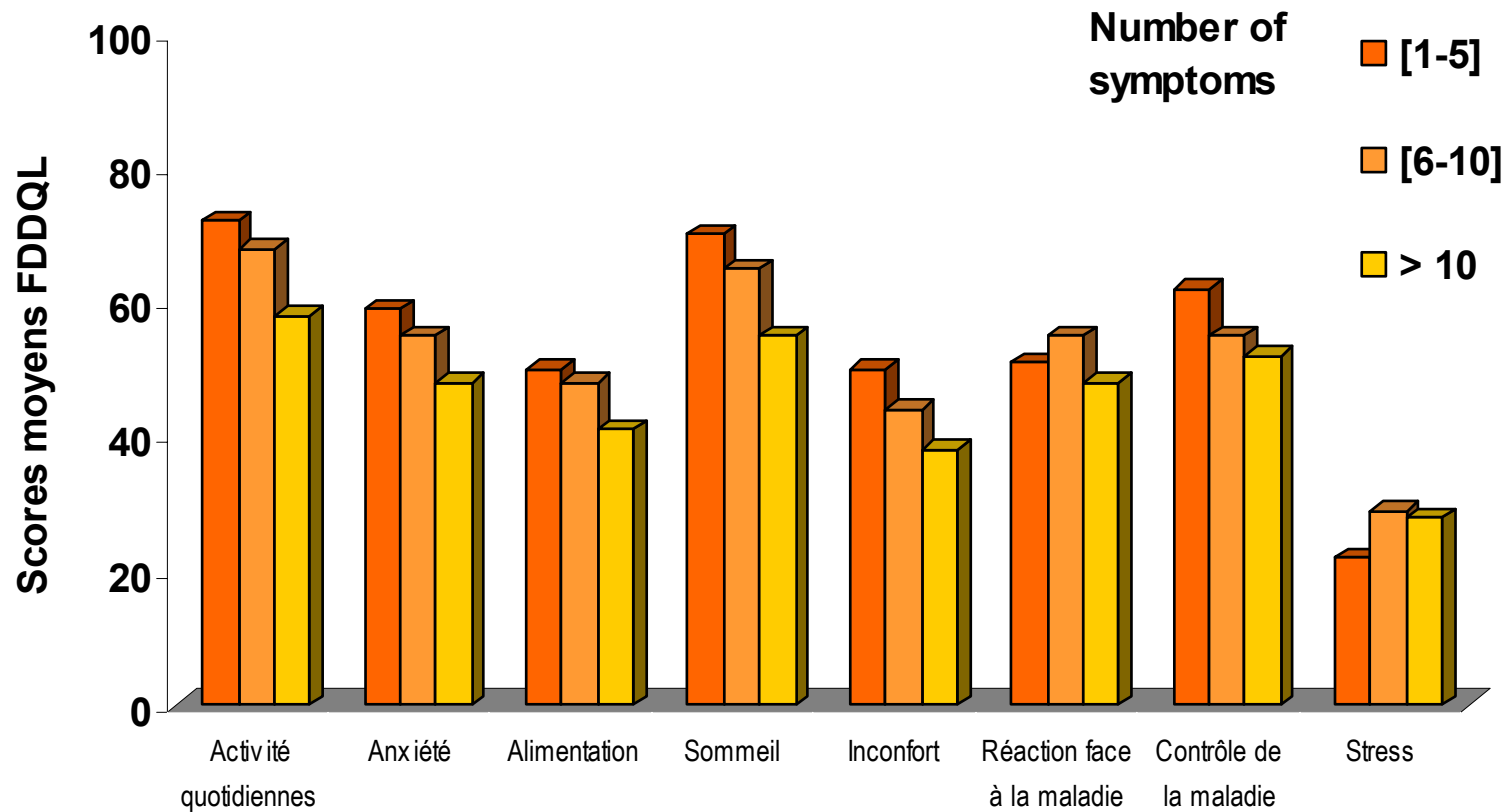
**Interference with social and personal relationships**

I find it hard to do all the things (e.g. diet) I have to do for my diabetes

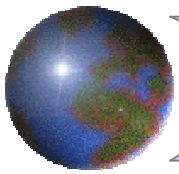
**Coping with disease**



# Discriminant validity of the Functional Digestive Disorders Quality of Life questionnaire (FDDQL)



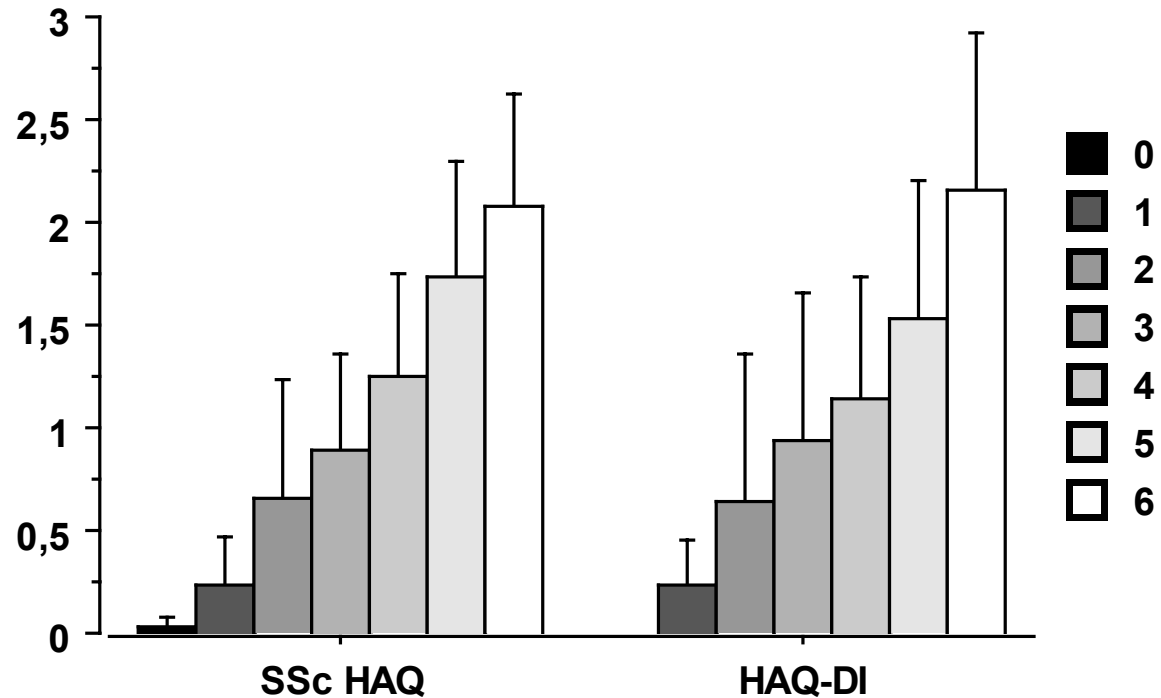
International study : France, Germany, Great Britain  
391 IBS and dyspeptic patients



# Discriminant validity of the Health Assessment questionnaire adapted to Scleroderma (SSc HAQ)

Score values ( $m \pm SD$ ) of the global SSc HAQ and HAQ-DI, according to the number of the following organ involvements (n=6):

- Raynaud's phenomenon
- Digital ulcers
- Gastro-intestinal
- Pulmonary
- Musculoskeletal
- Hand contracture



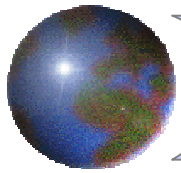
Comparison using ANOVA ( $p < 0.0001$  for both scores) (n=100 patients)

HAQ-DI: Health Assessment Questionnaire – Disability Index;

**Global SSc HAQ = (8 HAQ-DI domains + 5 VAS)/13.**

Validation of French version of the scleroderma health assessment questionnaire (SSc HAQ).

Georges C, Chassany O et al. Clinical Rheumatology, Under press.



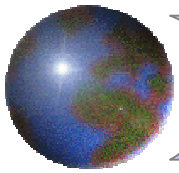
## Cultural adaptation - forward/backward

Disease: Asthma - Original version developed in Canada

Item: *Here is a list of activities in which some people with asthma are limited, among them: « shoveling SNOW »*

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



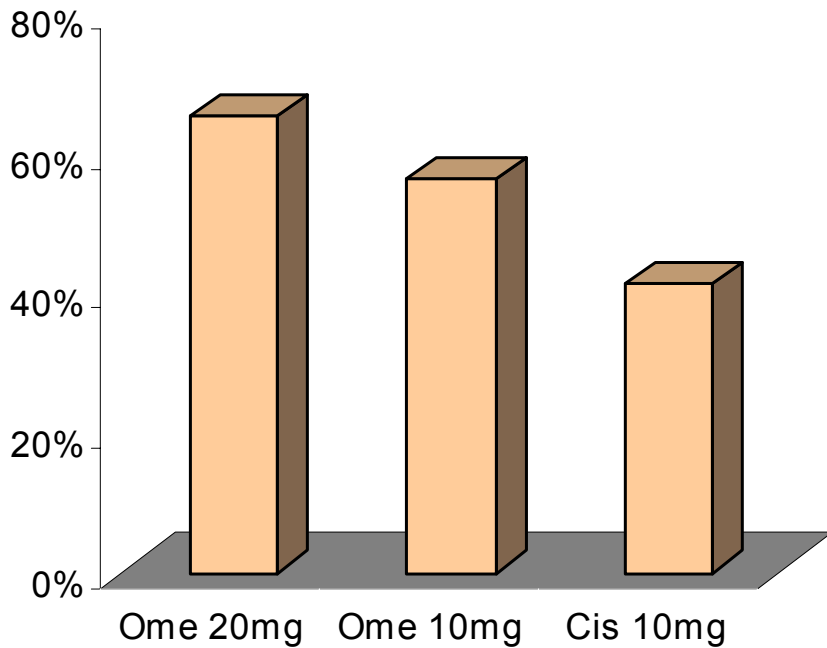


# Responsiveness - generic questionnaires

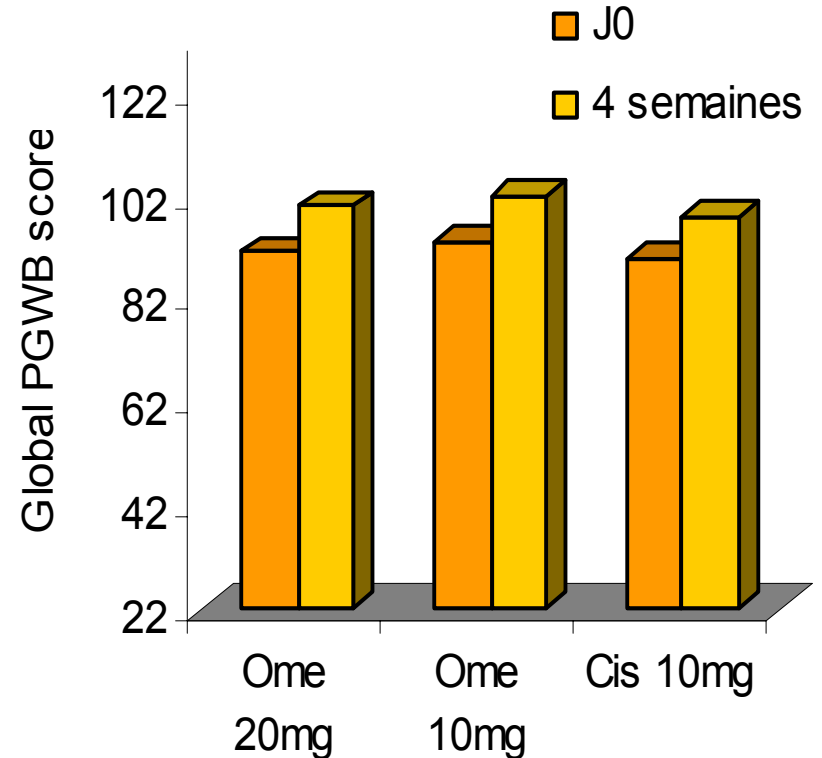
Psychological General Well-Being (PGWB) & GERD

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

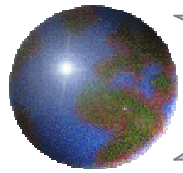
Résolution des symptômes à 4 semaines



24% difference in pyrosis relief

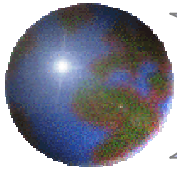


No difference in PGWB score



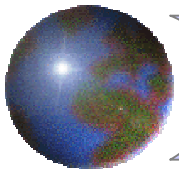
# Cross-cultural adaptation of questionnaires is not enough ?

- Specific CFQ-14 developed in France
- Translated in German
- Studies in  $n = 197$  and  $n = 103$  adolescents/adults
- **Construct validity** : same 9 HRQL domains as in the French original CFQ-14
- **Internal consistency** : ranged from 0.71 to 0.94
- **Clinical validity** : supported by severely ill patients reporting lower HRQL than less ill patients



## Study Design : specific issues related to HRQL / PRO measure

- **Eligibility criteria** : if HRQL primary endpoint, set a minimal impairment of HRQL (as for other criteria, e.g. pain, asthma onset... )
- **Timing and frequency** of HRQL assessment :
  - At baseline, at the end of the study or at withdrawal
- **Mode and site** of HRQL administration :
  - Self-administered whenever possible
  - Assure the confidentiality
  - Before the medical consultation
- **Data monitoring** and quality assurance
- **Procedures** for prevention and handling of missing data



# Study Design : Comparative randomized trials are a pre-requisite and double-blinded...

HRQL claims cannot be based on **non-comparative** and **non-blind** clinical trials.

They generally lead to a **higher rate** of positive results.

No causal link can be established between the therapeutic intervention and the HRQL change.

HRQL claim in Benign Hypertrophy Prostate

**1- Cohort study** (n = 7093) - specific scale.

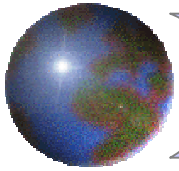
- HRQL score improving from **91 ± 32** (J0) to **109 ± 31** (J3) [**75% patients**] : **+ 29%**

- Improvement by 50% of symptoms (similar to the one observed in a study versus placebo, they forget to say that placebo leads to 40% improvement)

**2- Cohort study** (n = 5849)

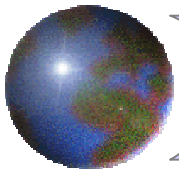
**3- Cohort study** (n = 4951) abstract

⇒ **Why not only 1 trial vs placebo ?**



## Statistical analysis plan : Estimating the adequate sample size

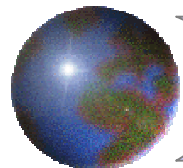
- HSQ (Health Status Questionnaire)
- before / after scores on 1300 patients
- All p values  $< 0.0001$
- Conclusion: all HRQL domains were significantly different across treatment groups
- Problem: 1300 provide 80% power to detect a change of **1 unit** on a **0-100** point scale



# Statistical analysis plan : Estimating the adequate sample size

Ranitidine vs. Placebo among > 500 patients with gastro-esophageal reflux disease

SF-36	Ran	Pla		p
PF	82.6	80,0	2.6	0.019
RP	77,0	74.6		NS
BP	73.8	69.1	4.7	0.003
GH	69.7	68.7		NS
VT	58,0	54.4	3.6	0.005
SF	85.5	83.7		NS
RE	81.9	78.2		NS
MH	72.5	71.7		NS



# Importance of withdrawals and missing data

TABLE 6. QUALITY OF LIFE CHANGE SCORES (MEAN ± SE) FOR PATIENTS COMPLETING 20 WEEKS OF ACTIVE THERAPY AND PATIENTS WITHDRAWING FROM STUDY

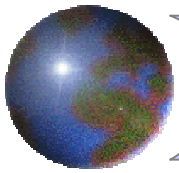
Quality-of-Life Scale	Treatment			
	Nifedipine GITS		Atenolol	
	Complete (n = 119)	Withdrawn (n = 68)	Complete (n = 131)	Withdrawn (n = 47)
Mean Psychosocial Score	+0.149 ± 0.03*	-0.156 ± 0.06	+0.020 ± 0.03	-0.063 ± 0.06
Mental Health Index	+0.134 ± 0.04*	-0.151 ± 0.06	+0.029 ± 0.03	-0.079 ± 0.09
Psychological Well-Being	+0.195 ± 0.04*	-0.117 ± 0.07	+0.041 ± 0.05	-0.018 ± 0.08
Psychological Distress	+0.098 ± 0.04*	-0.173 ± 0.06	+0.022 ± 0.03	-0.113 ± 0.10
General Perceived Health	+0.104 ± 0.04*	-0.148 ± 0.06	+0.016 ± 0.04	-0.084 ± 0.06

\* P < .001, significant difference between completers and withdrawals.

Positive (+) sign denotes improvement, negative (-) sign worsening.

**N = 365**  
(394 randomized)

**Poorer HRQL scores**



→ When the meaning of range score of different questionnaires is opposite

TABLE 1. Comparison of geometric mean scores (range limits)

	CAL	Controls	<i>P</i>
CRQ	75.2 (42–131)	115.8 (50–140)	<0.0001
BPQ	39.5 (9–83)	4.8 (0–41)	<0.0001

BPQ low scores = good QoL; CRQ low scores = severely disabled.