ISOQOL - Barcelona, Spain - 03 November 99

Health-Related Quality of Life & regulatory issues in Europe

ERIQA project

REFERENCE DOCUMENT

Olivier Chassany, MD

Agence Française du Médicament

Senior Lecturer in Therapeutics, University Lariboisière Hospital, France

for the European Regulatory Issues on QoL Assessment (ERIQA) Group

Major biases encountered in the review of a dossier in the drug approval process

- No **justification of QoL choice** (relevance, instruments)
- No evidence of quality of life questionnaire **validation**
- No **objective** of QoL changes
- No justification of sample size
- No description of the follow up of patients during the study
- No clear handling of **missing data**
- Not all patients are analyzed
- No correct presentation of results
- No adjustment for **multiple comparisons**
- **No interpretation of results**

Chassany O et al. Reporting on quality of life in randomised controlled trials. Authors are creating database of quality of life questionnaires. BMJ 1999; 318: 1142.

It is time to follow some guidelines

- HRQoL seems to be frequently added at the last moment without clear objective neither data quality procedures
- **Protocol, study report and clinical expert report do not document enough HRQoL assessment for a critical review**
- Reporting on quality of life should follow some **guidelines**.

Sanders C et al. Reporting on quality of life in randomised controlled trials: bibliographic study. BMJ 1998; 317: 1191-1194.

European regulatory authorities

- To convince them — that HRQoL is a relevant key outcome
- To make them confident — in the quality of the HRQoL results
- To help them — in reviewing and interpreting HRQoL results

Checklist on reporting on HRQoL in RCT

1- Study design clearly described ?

- Are basic methodological principles of RCT fulfilled and clearly reported

2- Scope and definition of the HRQoL component adequately described ?

- Relevance for assessing HRQoL for this trial
- **Justification for the choice of the HRQoL questionnaires**
- **Research objectives** of the HRQoL component clearly stated
- HRQoL a primary or secondary endpoint in the trial

3- Clear description of the study design elements as they related to the HRQoL component of the trial ?

- Sampling of patients and centres
- Eligibility criteria
- Timing and frequency of HRQoL assessment
- Mode and site of HRQoL administration
- Data monitoring and quality assurance

Where is HRQoL assessment ?

- Proton pump inhibitor / oesophagitis
- Phase III : 2 studies in USA, 1 study in Europe
- > 700 patients included

HRQoL claim : no clear definition of HRQoL in the study report, or in the protocol

- Overall physical well being (0 to 4)
- Time lost from usual activities of daily living

(less time lost in placebo group !)

No justification of choice of instruments

- Randomized, DB, placebo-controlled study of GH replacement in 40 patients with acquired GH deficiency
- **Assessment at baseline and 18 months :**
  - NHP (Nottingham Health Profile)
  - PGWB (Psychological General Well-being)
  - GHQ (General Health Questionnaire)
  - MMPI-2 (Minnesota Multiphasic Personality inventory)
- **Selection made on what ?**
  - Psychometrics properties ?
  - Prior use in a similar population ?
  - Cover several different concepts ?
- **What where the hypotheses of score changes ?**

Baum HBA et al. Effects of physiological growth hormone therapy on cognition and quality of life in patients with adult-onset GH deficiency. J Clin Endocrinol Metab 1998; 83: 3184-9.

Measuring HRQoL or symptom bothering ?

- **GSRS : Gastrointestinal Symptom Rating Scale**
- Disease-specific instrument to evaluate common symptoms of gastrointestinal disorders.
- 15 items, rated on a 7-point Likert scale from no discomfort to very severe discomfort into 5 scales :

1- abdominal pain
2- reflux syndrome (2 items)
3- diarrhoea syndrome
4- indigestion syndrome
5- constipation syndrome

- Have you been bothered by acid reflux during the past week ?
- Have you been bothered by heartburn during the past week ?

Reliability and validity of the gastrointestinal symptom rating scale in patients with gastroesophageal reflux disease. Qual Life Research 1998; 7: 75-83.

Checklist on reporting on HRQoL in RCT

4- Adequate description provided for the characteristics of the HRQoL measure(s) employed in the study ?

- Number of items and domains
- Instrument scaling and scoring
- Reliability
- **Validity**
- Responsiveness
- Respondent burden
- Cultural adaptation
- These properties have been shown in a population similar to the trial

HRQoL questionnaire validated ?

- Proton pump inhibitor / dyspepsia
- Phase III, RCT, DB, vs comparator & placebo
- n = 810, 2 weeks duration

HRQoL assessment (secondary)

Some unknown QoL index was used

➜ no description of validation data, no reference

➜ 10 additional items concerning gastro-intestinal symptoms were approved by Pr X !

Dossier for drug approval

Checklist on reporting on HRQoL in RCT

4- Adequate description provided for the characteristics of the HRQoL measure(s) employed in the study ?

- Number of items and domains
- Instrument scaling and scoring
- Reliability
- Validity
- Responsiveness
- Respondent burden
- **Cultural adaptation**
- These properties have been shown in a population similar to the trial

International trials and linguistic validation of questionnaire

- Treatment in claudication
- Phase III, RCT, DB, vs placebo, n=422, 6 months
- Setting in France and Italy

Results : initial change distance : Δ 32% vs placebo

HRQoL assessment (secondary)

- PQVS French generic questionnaire

➜ Italian version : not a word about linguistic validation

European mutual recognition procedure

Checklist on reporting on HRQoL in RCT

5- Clear description of the statistical analysis plan of the HRQoL component of the trial ?

- Efficacy or equivalence trial
- Intent to treat analysis (ITT)
- Procedures for type I error
- Sample size and statistical power
- Descriptive and inferential statistics
- Imputation of missing data

6- **Reporting of Results :**

- **Is the following information provided on the results of the HRQoL investigation ?**
  - Participation rate (at study entry and during follow-up)
  - Characteristics of the final study population
  - **Data completeness** (i.e. missing questionnaires and missing items)
- **Are the results presented in accordance with the original statistical analysis plan ?**

Intent to treat analysis and missing data

- Treatment in claudication
- Phase III, RCT, DB, vs placebo, n = 422, 6 months
- Results : initial change distance : Δ 32% vs placebo

HRQoL assessment (secondary)

- PQVS French generic questionnaire
- First factor of principal component analysis : global satisfaction (p = 0.049, t-test)

➜ Analysis performed on 324 patients

➜ How many and how were handled missing data ?

Multiple test comparisons

- Treatment in claudication
- Phase III, RCT, DB, vs comparator, n=324, 3 months
- Results : NO difference in walking in ITT

HRQoL assessment (secondary) : PQVS

➜ univariate analysis : statistical difference for 5 items among 19 (Per protocol analysis, n = 268)

➜ A statistical difference is likely to appear by random in about 5 items (type I error = 25%) !

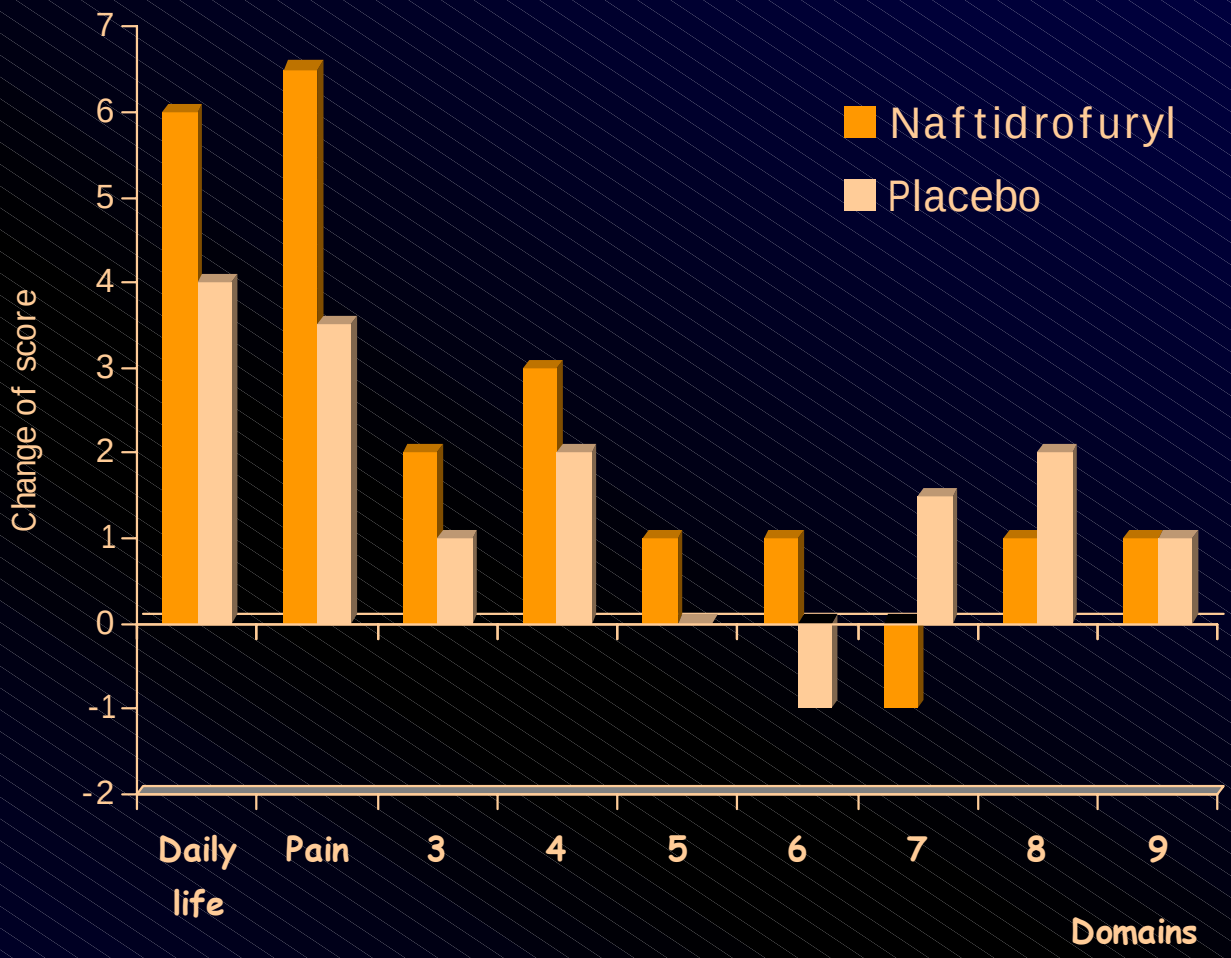
Checklist on reporting on HRQoL in RCT

7- Is there an attempt to interpret the statistical results in terms of clinical significance ?

- Description of the content of domains
- **Distribution of HRQoL scores within- and between groups**
- $_{95\%}$IC of the difference and/or odds ratio of the difference
- Effect size
- Comparisons of scores with norm scores (if available) and/or scores obtained in other studies in similar population to estimate a MID
- Comparison with external criteria to estimate a MID
- Number needed to treat

Distribution of HRQoL scores : no value

- Intermittent claudication
- RCT, DB, versus placebo
- Family practice
- 6 months
- HRQoL : primary end-point
- CLAU-S : 9 domains, 80 items
- ITT (234/250 included)
- Statistical significance set at 0.05/9
- Missing data : LVCF
- Not a value, just p-value and a graph



Liard F et al. The effects of naftidrofuryl on quality of life. Dis Manage Health Outcomes 1997; 2 (Suppl. 1): 71-78.

Checklist on reporting on HRQoL in RCT

7- Is there an attempt to interpret the statistical results in terms of clinical significance ?

- Description of the content of domains
- Distribution of HRQoL scores within- and between groups
- $_{95\%}$IC of the difference and/or odds ratio of the difference
- Effect size
- Comparisons of scores with norm scores (if available) and/or scores obtained in other studies in similar population to estimate a MID
- **Comparison with external criteria to estimate a MID**
- Number needed to treat

Clinical relevance of a difference

- Treatment in rheumatoid arthritis
- Phase III, RCT, DB, vs comparator, n=99, 6 months
- Results : Large improvement of ACR criteria (> 20% improvement : 71% (new drug) vs 27% (comparator)

HRQoL assessment (secondary)

- Health assessment questionnaire (HAQ)
- Differences on “disability”, “vitality” and “mental health” domains

➔ e.g.”disability” score (range : 0 to 3) at 6 months : 1.2 (comparator) vs 0.9 (new drug), how to interpret ?

Does it improve HRQoL or not ?

- Treatment in claudication
- Phase II, RCT, DB, dose-ranging vs placebo, n=340

Results : NO difference (absolute change distance)

HRQoL assessment (secondary) :

- SF36 : significant differences on "social function" and "mental health" scales
- NO difference on "pain", "physical function"

➔ Improvement of HRQoL ? (i.e. how many domains)

Checklist on reporting on HRQoL in RCT

7- Is there an attempt to interpret the statistical results in terms of clinical significance ?

- Description of the content of domains
- Distribution of HRQoL scores within- and between groups
- $_{95\%}$IC of the difference and/or odds ratio of the difference
- **Effect size**
- Comparisons of scores with norm scores (if available) and/or scores obtained in other studies in similar population to estimate a MID
- Comparison with external criteria to estimate a MID
- Number needed to treat

➔ **Try at least something !**

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	95% CI	**ES**	Mean ± SD	95% CI	**ES**
Total SIP	0.08 ± 6.6	-1.71, 1.44	-0.01	0.56 ± 6.5	-1.41, 2.53	0.09
Physical. dimension	0.73 ± 6.1	-0.86, 2.34	**0.12**	0.87 ± 6.1	-0.96, 2.7	**0.14**

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

DHEA : a fountain of youth (1)

- Randomized, DB, placebo-controlled, cross-over (24 women)
- Each treatment period : 4 months
- Washout period : one month
- Assessment after one and 4 month of each period and one month after the 2nd period

1- SCL-90 : revised version of the 90-item Symptom Checklist
2- Multidimensional Mood Questionnaire
3- Von Zerssen Symptom list
4- Short form of the Giessen Complaint list
5- German version of the Hospital Anxiety and Depression
6- Sexual functioning

Alrt W et al. Dehydroepiandrosterone replacement in women with adrenal insufficiency. N Engl J Med 1999; 341: 1013-20.

DHEA : a fountain of youth (2)

- **SCL-90 : revised version of the 90-item Symptom Checklist**
  - Psychological well-being
  - Psychometric properties verified prior to the trial in similar population ?
- **Multidimensional Mood Questionnaire**
  - How reliable is the questionnaire ?
- **Von Zerssen Symptom list**
  - Content not provided
- **Short form of the Giessen Complaint list**
  - More a symptom listing
- **Sexual functioning**

No evidence of validation provided

Alrt W et al. Dehydroepiandrosterone replacement in women with adrenal insufficiency. N Engl J Med 1999; 341: 1013-20.

DHEA : a fountain of youth (3)

- No primary end-point defined
- No sample size justified (i.e. no hypotheses)
- Choice of multiple instruments not justified (i.e. SCL-90, MMQ, Zerssen and HAD reflect the same concept)
- Multiple statistical test comparisons :
  - **At least 23 tests repeated 3 times (α > 32%)**
  - No adjustment for multiple comparisons
  - No selection of the most important measures
- Double-blind respected ? most women receiving DHEA experienced skin-related androgenic effects

Alrt W et al. Dehydroepiandrosterone replacement in women with adrenal insufficiency. N Engl J Med 1999; 341: 1013-20.

DHEA : a fountain of youth (4)

- **No data about completeness of assessments of the 2 periods:** About **204 items** completed at least **6 times**.
- Significant differences between groups at 4 months in favour of DHEA for several questionnaires :
  - SCL-90 (e.g. global score : **Δ of 0.11**, 0-4 range)
  - Clinical pertinence ? **(Effect size = 0.28)**
- **Conclusion: DHEA improves well-being**
- **Editorial:** "it is now justifiable, to prescribe DHEA as long-term treatment in patients with adrenal insufficiency, **provided that they are monitored for breast or prostatic cancer**".

Alrt W et al. Dehydroepiandrosterone replacement in women with adrenal insufficiency. N Engl J Med 1999; 341: 1013-20.

Conclusion

- Regulatory authorities will accept more easily HRQoL statistical significant results if :
  - they have confidence in the quality of the trial itself
  - and protocol, study report and clinical expert report document enough HRQoL assessment for a critical review
- Thus, the clinical relevance of results will appear less important.
- Whether the endpoint is considered primary or secondary, the scientific principles of clinical trial design must apply to Health-Related Quality of Life.