Quality of life assessment in clinical oncology research and practice: Current state-of-the-art

Neil K. Aaronson

DIA Workshop: Assessing Treatment Impact Using PROs: Challenges in Study Design, Conduct and Analysis

Paris, May 10-11, 2004

fact or fiction?

The term "(health-related) quality of life," is well defined and widely understood.

Fact – if you keep things simple
Fiction – if you dig deeper

"Quality of life is a vague and ethereal entity, something that many people talk about, but which nobody clearly knows what to do about."

Campbell et al., 1976

"The idea has become a kind of umbrella under which are placed many different indexes dealing with whatever the user wants to focus on."

Feinsten, 1987

4 criteria for evaluating clinical effectiveness of chemotherapeutic agents in lung cancer

D.A. Karnofsky et al., *Cancer* 1:634 , 1948

- subjective improvement
- objective improvement
- performance status
- length of survival

Subjective improvement

"The patient's subjective improvement is measured or described in terms of:

- improvement in his mood and attitude

- his general feeling of well-being,

- his activity, appetite, and the alleviation of distressing symptoms such as pain, weakness, and dyspnea."

WHO definition of health, 1949

"A state of complete physical, mental and social well-being, and not merely the absence of disease and infirmity."

Key dimensions of quality of life as defined by ASCO, 1995

Physical   Symptoms commonly caused by cancer and the toxicities of treatment

Psychologic   Effects of cancer and its treatment on cognitive function and emotional state

Social   Effects of cancer and its treatment on interpersonal relationships, school, work and recreation

Examples of QL definitions

"The difference between the hopes and expectations of the individual and the individual's present experience."

Calman, 1987

"The functional effect of an illness and its consequent therapy upon a patient, as perceived by the patient."

Schipper et al. 1996

Attributes of QL definitions

- non-specific vs. health-related

- health states (or status) versus personal evaluation of those states (e.g., expectations, discrepancies, satisfaction)

- scope of concerns (e.g., spirituality or existential issues)

- polarity of concerns (well-being vs. dysfunction and its resolution)

Does it matter?

- Yes, because the content of QL questionnaires reflects the underlying definition.

- It may be less important in clinical trials, where group comparisons will be internally valid, regardless of the definition used.

- It is more important in comparing results across trials and in descriptive (e.g., prevalence) studies.

Covinsky et al. Am J Med 1999; 106:435-440

- study of 493 older patients

- QL rated as good/excellent by 43% of those with worst physical functioning and 47% with highest levels of psychological distress

- QL was rated as poor by 15% of those with the best physical functioning and 21% with the lowest levels of psychological distress
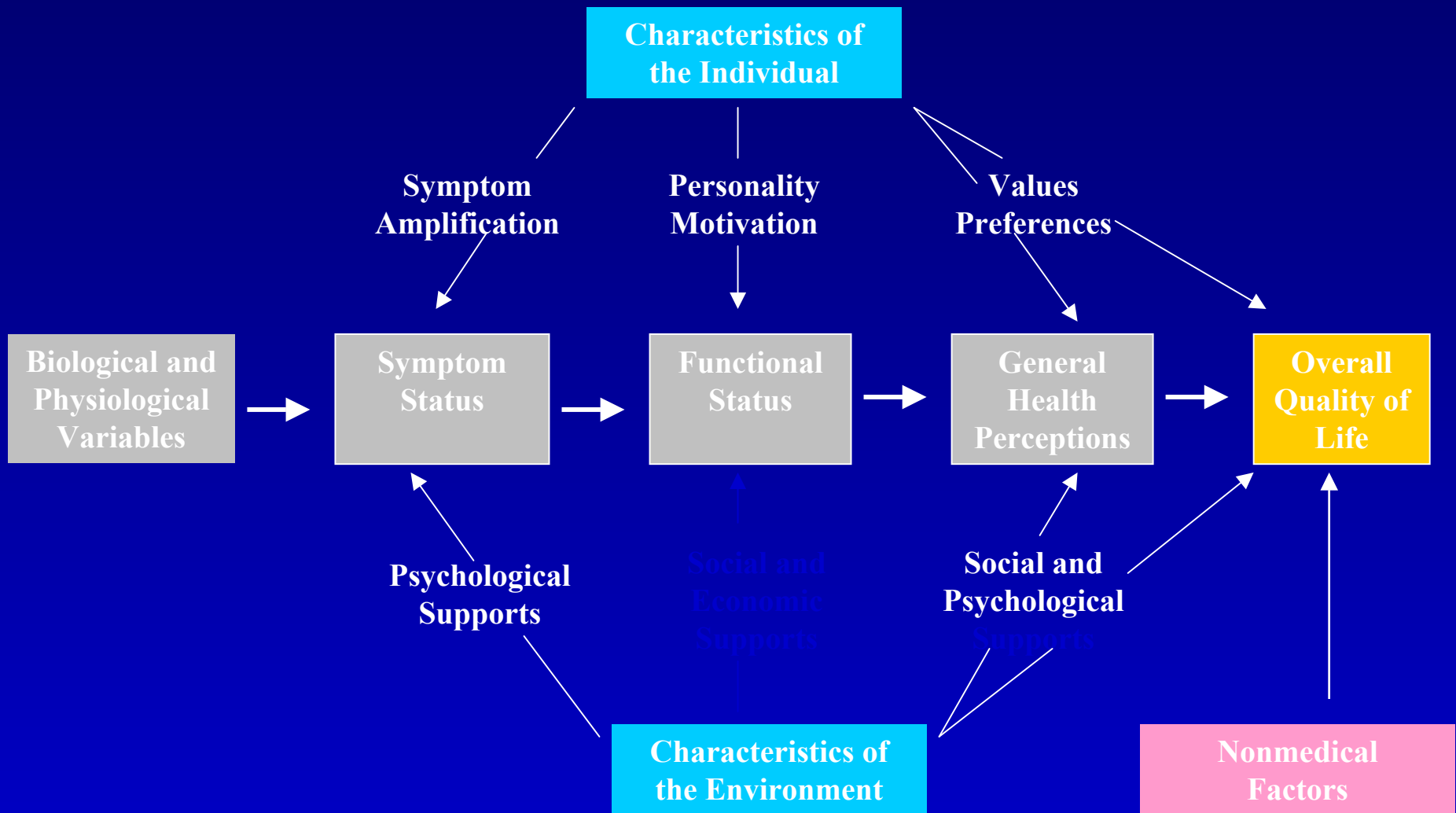
fact or fiction?

We have adequate conceptual models for studying the underlying associations between QL domains, and the factors influencing those associations.

reasonably factual

multidimensional health-related quality of life assessment

existential/spiritual

social

psychological

physical functioning and symptoms
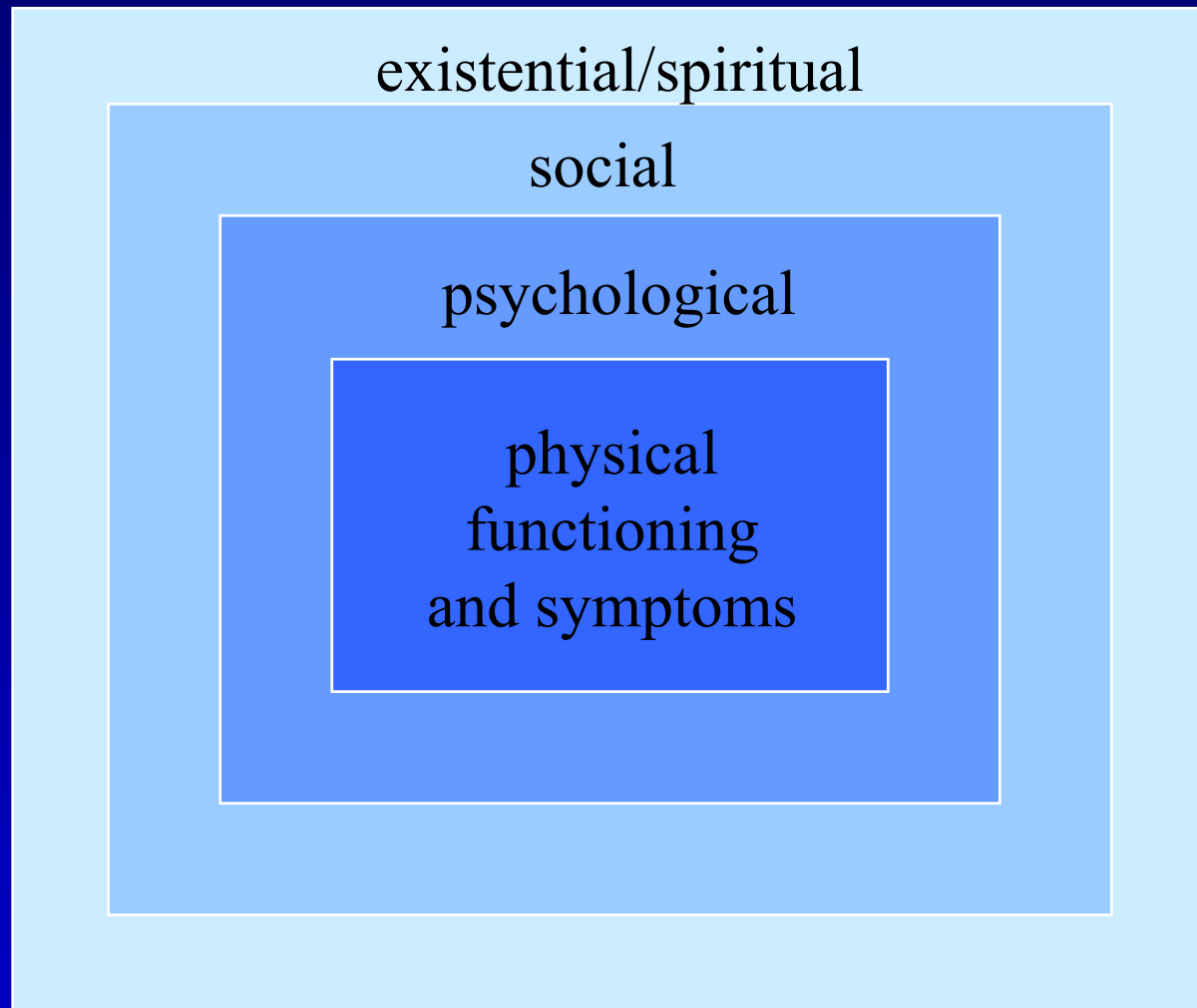
Relationships among measures of patient outcome in a health-related quality of life conceptual model
(Wilson and Cleary, JAMA 1995; 273(1): 59-65)

fact or fiction?

The patient is the sole legitimate source of information about his/her QL. Other "proxy" raters (e.g., family members, health care providers) are, at best, poor substitutes.
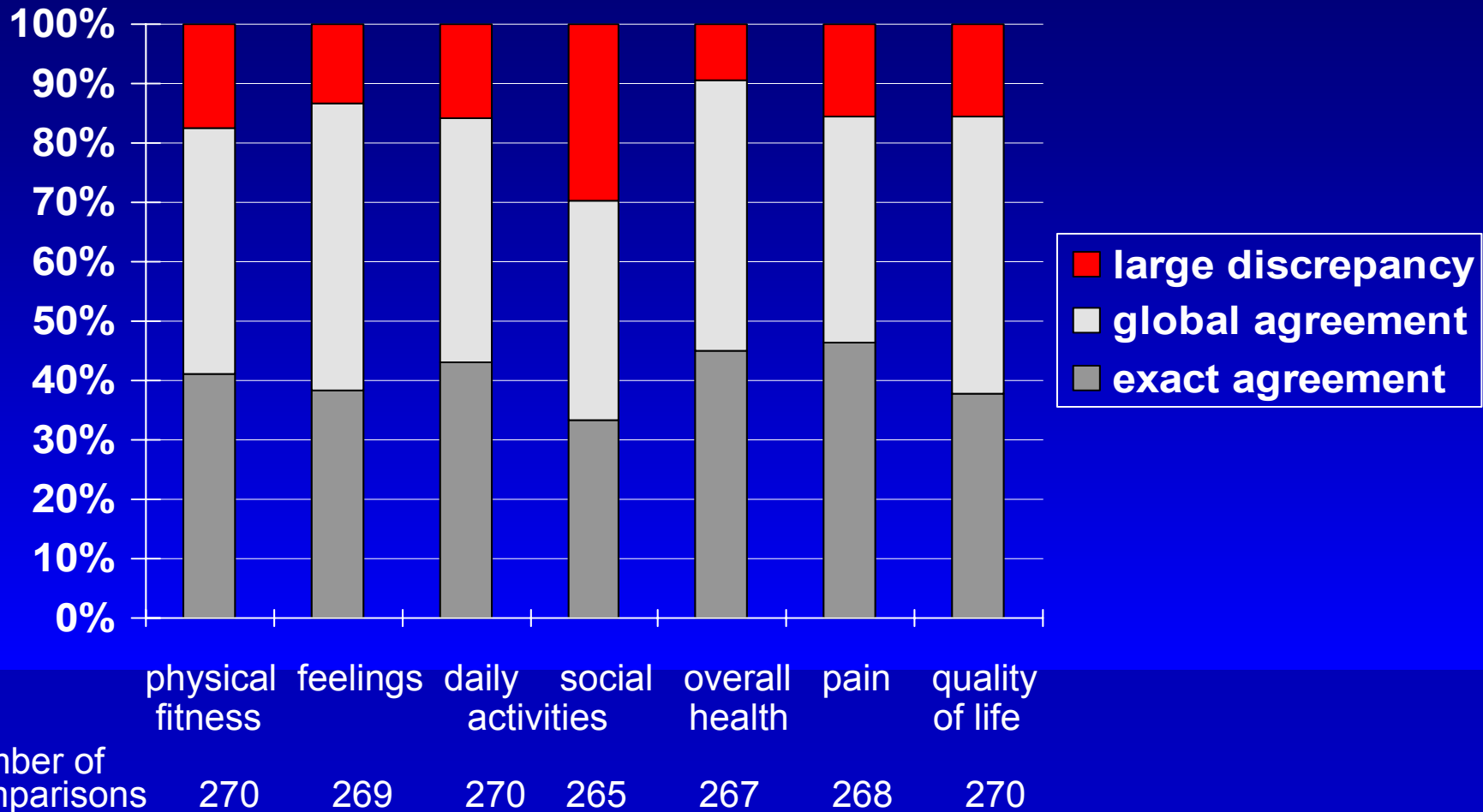
(partial) fiction

Self-report can be limited by:

- age (very young or old)
- cognitive impairment
- communication problems
- symptom distress
- physical disability
- emotional distress

Exclusion of highly relevant subgroup of patients can result in biased study outcomes

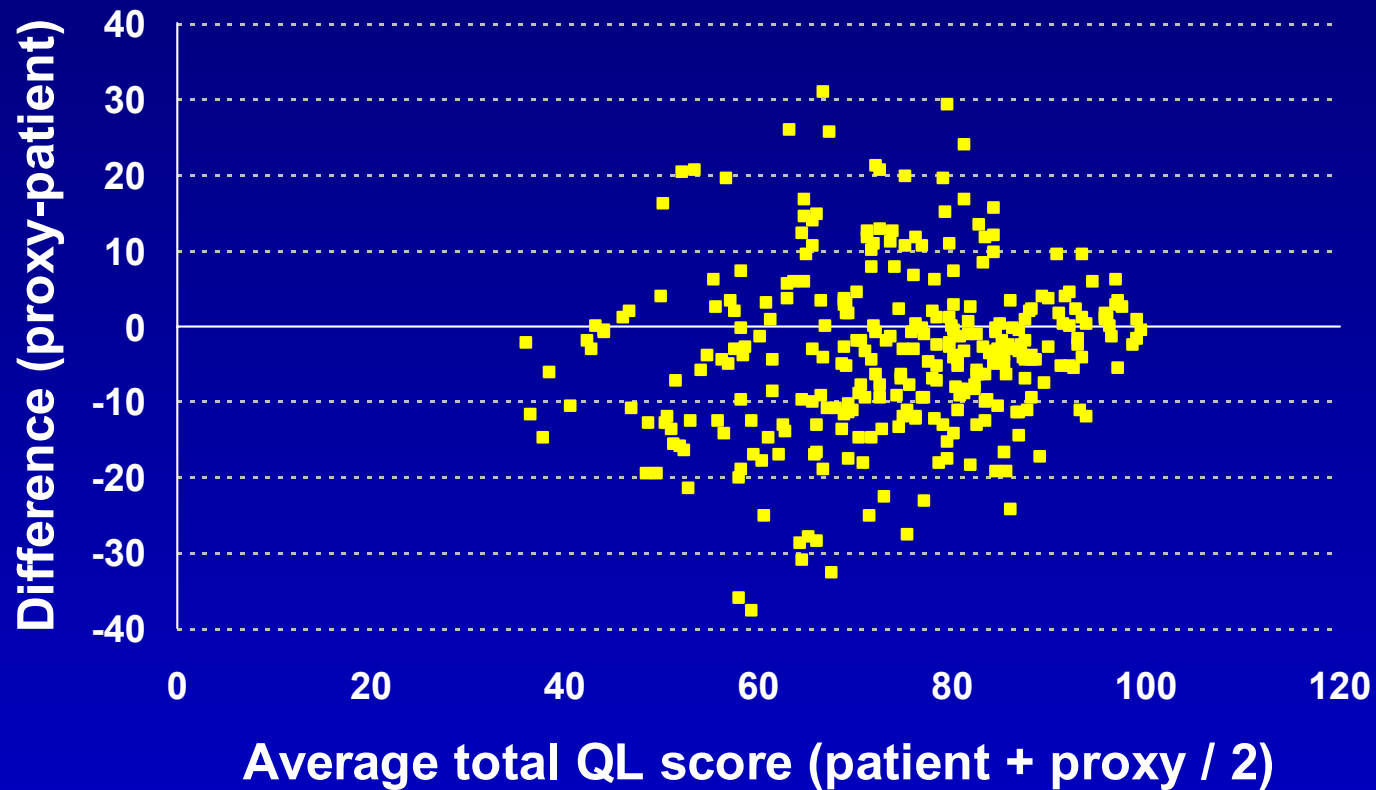# The role of health care providers and significant others in evaluating the QL of patients with chronic disease

- 23 studies published between 1991 - 2000
- Moderate/high patient – proxy agreement
- Proxies tended to rate patients as having more problems than did patients themselves
- Magnitude of differences was small (median standardized difference 0.20)

Proportion of agreement
by question/QL domain

Legend:
- large discrepancy (red)
- global agreement (light grey)
- exact agreement (grey)

	physical fitness	feelings	daily activities	social	overall health	pain	quality of life
number of comparisons	270	269	270	265	267	268	270

Bland-Altman plot for total QL score

fact or fiction?

Although there are many QL question-naires from which to choose, the dust is settling and a "best bet" can be identified based on a comparison of psychometric characteristics and performance.

fiction

Generic QL instruments

- Sickness Impact Profile (SIP)
- Nottingham Health Profile (NHP)
- Spitzer QL Index
- COOP/WONCA Charts
- MOS 36-Item Health Survey (SF-36)
- World Health Organization (WHOQoL)

Cancer-specific QL questionnaires

- Functional Living Index – Cancer (FLIC)

- Cancer Rehabilitation Evaluation System (CARES)

- Rotterdam Symptom Checklist (RSCL)

- EORTC QLQ-C30

- Functional Assessment of Cancer Therapy (FACT-G)

Key psychometric attributes of QL instruments

- measurement model
- reliability
- validity
- responsiveness
- interpretability
- cultural adaptability
- burden

Choice of QL instrument should be driven by:

- the research question(s) to be addressed
- the population under study
- the conceptual basis of candidate questionnaires
- the specific content and wording of candidate questionnaires

Negative affect items

SF-36   "Have you felt so down in the dumps that nothing could cheer you up?"

"Have you felt downhearted and blue?"

FACT-G   "I feel sad"

QLQ-C30   "Did you feel depressed?"

Future perspective items

SF-36   "I expect my health to get worse."

FACT-G   "I worry about dying."

QLQ-C30   --

fact or fiction?

Given the plethora of QL questionnaires currently available, there is little or no need for continued efforts at instrument development.

fiction

- Condition-specific questionnaires tend to be more sensitive to group differences and responsive to inter- and intra-individual changes over time

supplemental modules/scales

- combine "core" instrument with condition-specific modules/scales

  - EORTC "modules"
  - FACT subscales
  - NCIC symptom checklists

advantages of core + module approach to QL assessment

- facilitates comparison of results across studies

- provides sufficient flexibility to address questions specific to a given patient population or treatment

computer-adaptive testing

- item banking initiatives

- modern test (item response) theory

- QL measurement systems that are more flexible, efficient and precise

fact or fiction?

The major methodological challenges in QL analysis – missing data, multiple comparisons, and clinical interpretation of statistical results – have been resolved or are well on their way to being resolved.

Reasonably factual
(2 out of 3 ain't bad)

Missing data:
Items from questionnaires

- Relatively minor problem (less than 5%)

- For multi-item scales, missing responses can be estimated/replaced

- High level of missing values for a given item may signal problem of appropriateness or acceptability

Missing data:
Questionnaires

- **Missing at random (e.g., administrative failure)**
  - **Largely avoidable**

- **Systematic loss to follow-up due to illness or death ("informative censoring")**
  - **Often unavoidable (e.g., in advanced disease trials)**
  - **Complex problem with imperfect but workable solutions**
    - **Mixed effects ANOVA**
    - **Growth curve analysis**

Multiple comparisons

- Inherent problem with multidimensional QL measures (health profiles)
- Results in inflated p values
- Three primary solutions
  - Use summary scores, where available
  - Focus on a few "cardinal" (primary) outcomes
  - Apply statistical adjustments

Defining clinical vs. statistical significance in QL scores

- The post-test mean score of the intervention group was 74.3 vs. 67.5 in the control group ($p < .001$)

- The between-group difference in mean QL scores represented a moderate effect size (Cohen's d = 0.5).

- 35% of the intervention group experienced a QL benefit (> 10 point change) vs. 22% of the control group

fact or fiction?

Drug regulatory agencies and key clinical oncology groups are increasingly open to and supportive of the use of QL outcomes in clinical trials.

In theory, factual –

In practice, fiction?

U.S. FDA 1985

"…In the past, new anti-cancer drugs were approved solely on the basis of objective tumor response, but this is no longer the case.…Survival and quality of life are the key efficacy parameters."

Johnson and Temple, Cancer Trt Rep, 1985

U.S. FDA 1996

"The Oncologic Drugs Advisory Committee has recommended that beneficial effects on QoL and/or survival be the basis for approval of new anticancer drugs…"

Beitz, Gnecco & Justice, JNCI Monographs, 1996

Endpoints in U.S. F.D.A. approval of oncology drugs: 1990 - 2002

- Marketing approval given to 57 drugs via standard procedures
- Basis of approval:
  - Survival – 32%
  - Tumor response – 46%
  - Tumor-specific symptoms – 23%
  - Other – 16%
  - Quality of life – 0%

  - Johnson RJ et al. J Clin Oncol 2003; 21:1404-11

FDA drug approval based (in part) on symptom relief

Drug	Indication	Endpoint
Gemcitabine	Pancreatic cancer	Clinical benefit response (pain, PF, weight gain)
Paclitaxel	Karposi's sarcoma	Cosmesis
Liposomal daunorubicin	Karposi's sarcoma	Cosmesis
Methoxsalen	T-cell lymphona	Edema, scaling
Mitoxantrone	Metastatic prostate cancer	Pain
Topotecan	SCLC	Lung cancer symptoms

"Although QOL assessments have been submitted in oncology drug applications, this aspect of the clinical trials has generally not been well conducted. Problems have usually included unblinded assessment, large amounts of missing data, and poorly defined prospective analytic plans…"

Johnson RJ et al. J Clin Oncol 2003; 21:1404-11

Outcomes of cancer treatment for technology assessment and cancer treatment guidelines adopted by ASCO, July 24, 1995

- survival
- quality of life
- cancer response
- cost-effectiveness

"Survival…is the most important outcome in cancer treatment. Nevertheless, survival alone is not sufficient; the quality of survival and cost of maintaining or improving it must also be assessed."

ASCO, July 24, 1995

"In the case of metastatic disease, treatment can be recommended even without an improvement in survival, if it improves quality of life."

ASCO, July 24, 1995

HRQL measurement in randomized clinical trials in breast cancer

- 46 RCTs reviewed
  - 8 primary management
  - 7 adjuvant therapy
  - 20 metastatic disease
  - 11 symptom control/supportive care

Key conclusions:
RCTs in primary management

"…HRQOL measurement provided information that was useful in selecting optimal treatment when two medical treatments were demonstrated to have equivalent medical outcomes."

Key conclusions:
RCTs of adjuvant therapy

"In general, HRQOL effects were either absent, transient or associated with observed toxicity…HRQOL measures have had little impact on clinical decision making.

Key conclusions:
RCTs in metastatic disease

"Disappointingly, HRQOL outcomes in these studies have provided little additional information beyond that obtained from traditional medical outcomes…In none of the published studies…did HRQOL measurement provide information that had a clear effect on treatment recommendations."

Key conclusions:
RCTs in symptom control

" …measurement of HRQOL adds little if any benefit to traditional medical outcomes in these trials…focus on the specific symptoms being studied rather than selecting general or even cancer-specific HRQOL instruments."

fact or fiction?

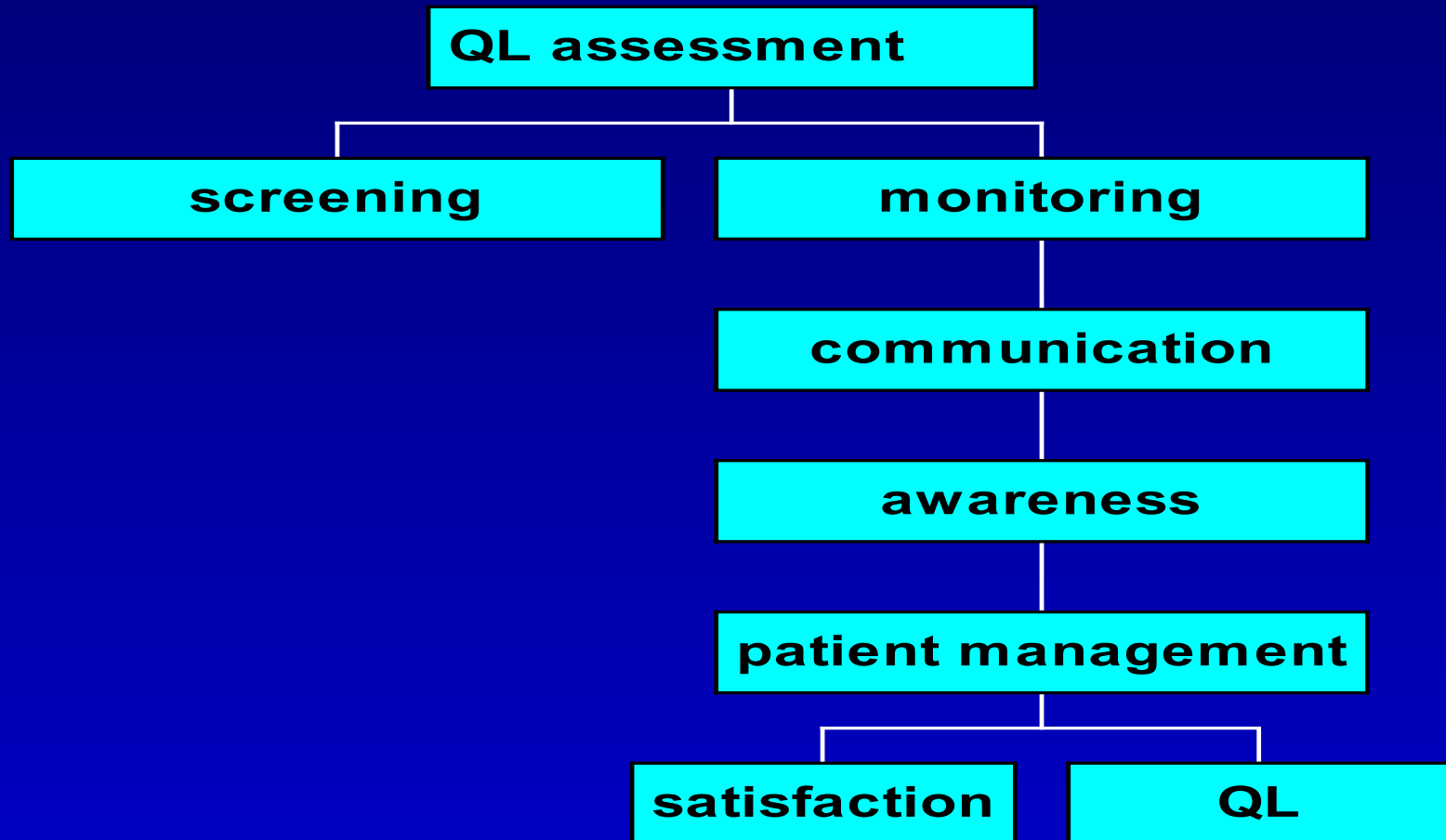QL assessment is ready for prime time as a tool in daily clinical practice.

"faction"

Brodman K. et al. The Cornell Medical Index: An adjunct to medical interview JAMA 1949; 140:531-34

- 195 item self-administered questionnaire on physical and psychological symptoms and medical history

- completed prior to office visit in 10-30 minutes; high compliance rates

- Elicited information not found in medical records

Modeling the use of QL assessment in clinical practice

```
            ┌─────────────────────┐
            │   QL assessment     │
            └─────────────────────┘
         ┌────────────┴────────────┐
┌──────────────────┐      ┌──────────────────┐
│    screening     │      │    monitoring    │
└──────────────────┘      └──────────────────┘
                                   │
                          ┌──────────────────┐
                          │  communication   │
                          └──────────────────┘
                                   │
                          ┌──────────────────┐
                          │    awareness     │
                          └──────────────────┘
                                   │
                          ┌──────────────────┐
                          │ patient management│
                          └──────────────────┘
                          ┌────────┴─────────┐
                  ┌──────────────┐    ┌──────────────┐
                  │ satisfaction │    │      QL      │
                  └──────────────┘    └──────────────┘
```

QL assessment in daily clinical practice: Feasibility

- Self-administered questionnaires can be completed quickly in office-based practice

- Computer-assisted (e.g., touchscreen) administration is acceptable and efficient

- No evidence that collection of standardized QL data interferes with normal clinic routine or lengthens average visit time

QL assessment in daily clinical practice:

16 randomized studies published 1987-2002

4 of which were in oncology setting:

(Taenzer et al. 2000; McLachlan et al. 2001; Detmar et al. 2002; Velikova et al. 2003)

- communication   +
- awareness   +
- patient management   +/-
- satisfaction   -
- health outcomes   +/-

Moving things forward (1)

- Make better use of existing conceptual models in:
  - shaping research questions

  - selecting appropriate measures and methodologies

  - guiding analysis strategies

Moving things forward (2)

- Don't reinvent the QL measurement wheel, but rather:
  - refine existing "traditional" QL measures

  - invest in the development of additional, condition-specific measures

  - contribute to the collective effort needed to develop "dynamic" (computer-adaptive) measurement systems

Moving things forward (3)

- Continue efforts to translate statistically significant QL findings into clinically meaningful terms

- Make substantial investment in a limited number of high profile clinical trials where QL is likely to yield added value

Moving things forward (4)

- Move forward with the application of QL measures in daily clinical practice, but don't make promises you can't keep

- Investigate (and attempt to strengthen) the "weak links" in the putative causal chain between routine QL assessment and improved patient QL over time