Quality of Life Measurement for NCIC CTG - Clinical Trials

Jolie Ringash

Co-Chair, QOL Committee

NCIC-CTG

The Princess Margaret Hospital, Toronto

Thank you

- Input: NCIC-CTG QOL Committee
- Slides:
  - Andrea Bezjak
  - Michael Brundage
  - David Osoba
  - Joe Pater

NCIC CTG
NCIC GEC

Learning Objectives

- Understand the meaning of “QOL”
- Learn basics of QOL measurement
- Interpret of QOL Results
- Learn the history of the CTG QOL Committee

NCIC CTG
NCIC GEC

What is QOL?

- “the goodness of life”
- patient’s perspective
- multi-dimensional
- includes hope/ hopefulness
- as related to health (not housing, income, environment, etc)

NCIC CTG
NCIC GEC

What is QOL?

- “Optimum levels of physical, role and social function, including relationships, and the perception of health, fitness, life satisfaction and well-being. It should also include some assessment of the patient’s level of satisfaction with treatment, outcome and health status and with future prospects.”

Bowling, 1995

Why is QOL an increasingly frequent outcome in clinical trials?

- Disease-centered outcomes (response rates, cause-specific survival) are not the only clinically relevant outcomes
- Toxicity has traditionally been assessed from the view-point of medical staff
- Patient-centered outcomes have become an important measure of the patient experience of their illness and treatment

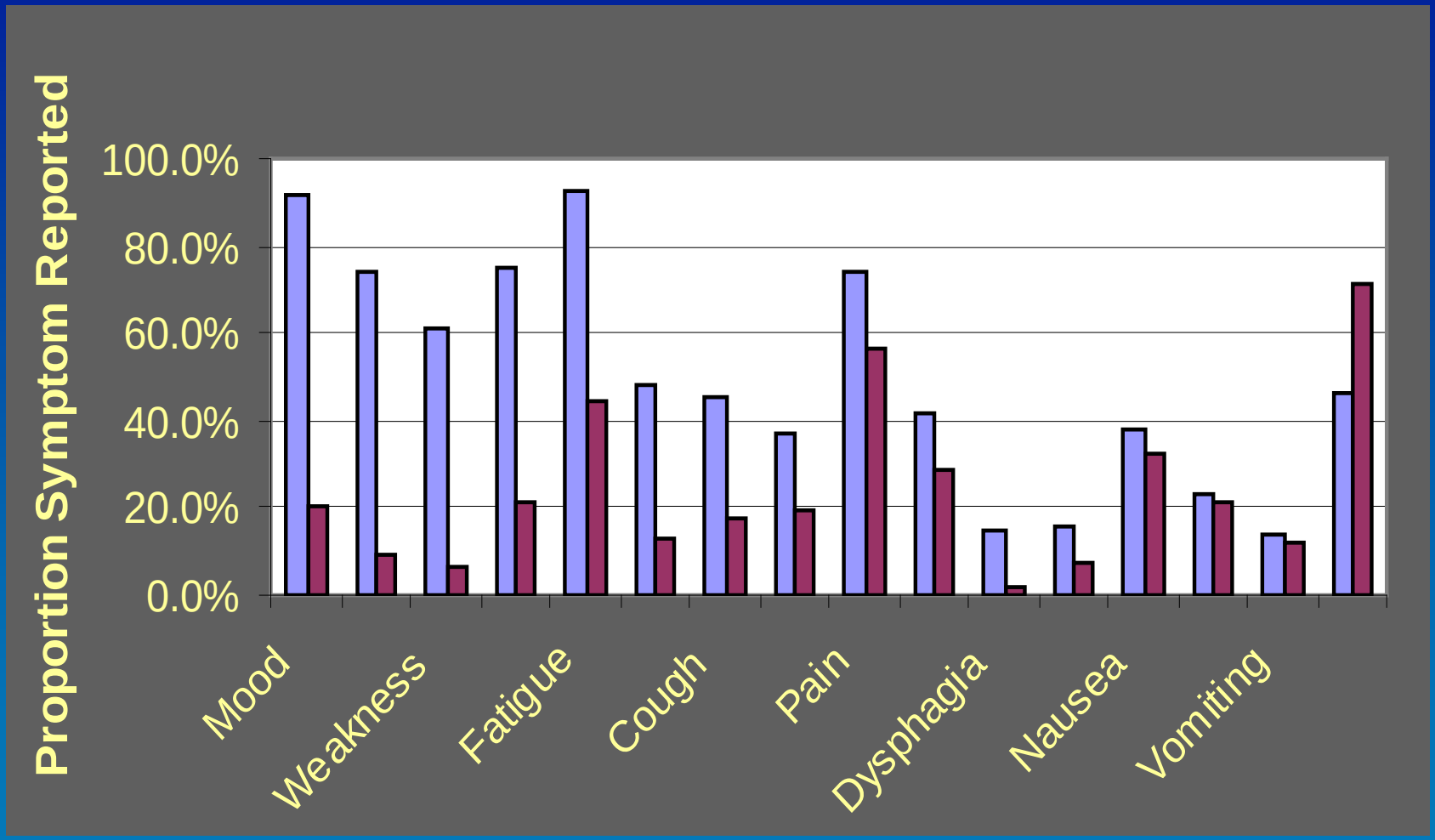
HRQL vs. Toxicity

Comparison of toxicity to HRQL data (advanced breast cancer):

- Agreement was found to be fair to slight
- (kappa 0.012 to 0.378)
- Patient's reported far more symptoms (by HRQL) than noted by toxicity scores.

These differences influenced the interpretation of trial results

HRQL vs. Toxicity



Savage et al, ASCO 2005

NCIC CTG
NCIC GEC

Why QOL is important

- Different treatments have similar survival
- Treatment improves survival but has severe side effects
- Treatment has no effect on survival but may improve QOL
- Cure is not possible
- Chronic diseases with high survival rates

Issues to Consider

- The 5 W's: Who, What, When, Where, Why, How
  - Who are the patients (cancer type, age, etc)
  - What are their concerns or issues
  - When & where will QOL be measured
  - Why measure QOL?
  - How? Self-completed, computerized, interview

Who and What?

- Consider cancer type, gender, age, level of education, stage, treatment, and point in disease course
- Review literature or interview patients about their issues and concerns
- Are there existing, validated instruments?
- Consider emotional, social, cognitive, role-fulfillment and spiritual issues as well as physical condition

When & Where?

- Cross-sectional vs. longitudinal design
- Do you want to describe a state, or measure change?
- Beware pitfalls of missing or untimely data
- Respondents may be more comfortable in their own homes, but the clinic may be more practical
- Timing with respect to doctors' visits

How? Administration

- Self-completion is the gold standard
- Interviews must be standardized
- Use of computers is promising
- Use of proxy information is difficult

NCIC CTG
NCIC GEC

How important is it to get all patients to complete QOL questionnaires?

- Extremely important!!
- The biggest problem with analyzing QOL information from clinical trials is missing data - are pts whose QOL data are missing different from pts supplying QOL data? Or is QOL data missing because pts are sicker that those providing info?
- Analysis can try to account for missing data but it is best trying to prevent missing data

Reliability

- Does the questionnaire produce reproducible results?
  - Internal consistency (similar items score similarly, eg. Cronbach's alpha)
  - Test-retest reliability (5-7 days) – ideally concordance (ICC) rather than correlation
  - Longer questionnaires are more reliable

Validity

- Does the questionnaire really measure QOL?
- Face and content validity
  - Do the questions make sense? Are they relevant? Is the administration and scoring sensible?
- Criterion validity
  - Compare to a “Gold Standard”

Construct Validity

- Also called: Concurrent validity, convergent validity, divergent validity
- Formulate and test hypotheses
  - Eg. CAROT score will be higher in younger patients and those with stage I toe cancer
  - Eg. CAROT score will correlate positively with EORTC QOL score and negatively with HAD (anxiety) score

Sensitivity (Responsiveness)

- Do scores accurately reflect change over time?
  - Usually measured in a group expected to change, eg. During toxic therapy, after cure
  - The more specific the questionnaire, the better its responsiveness

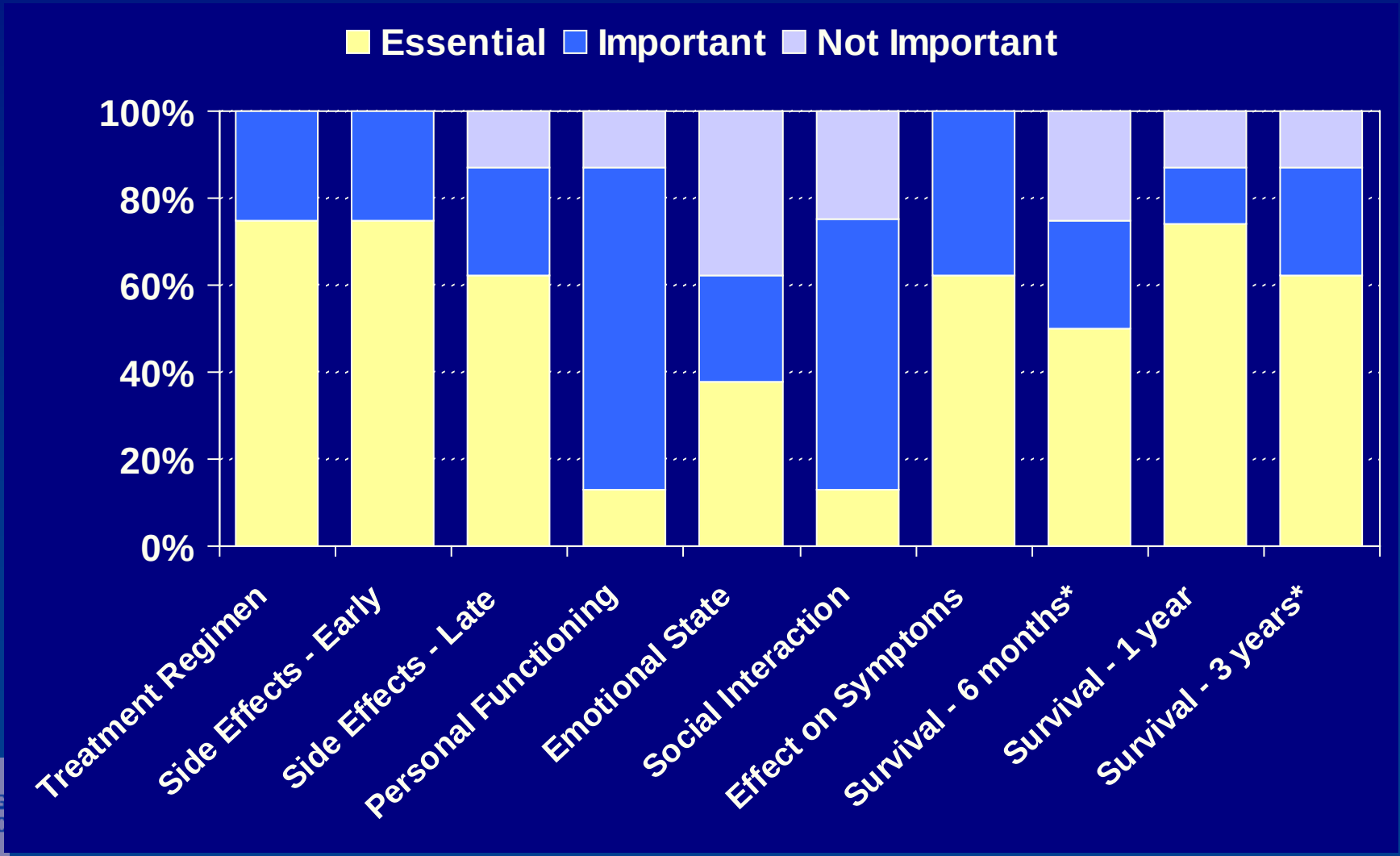
NCIC CTG
NCIC GEC

Patients' Information Needs

- Patients have a extensive list of information needs
  - collectively and individually
- Information may be needed for one or more of several purposes
  - Decision making, planning, preparing, understanding, and so on

NCIC CTG
NCIC GEC

Patients' Information Needs

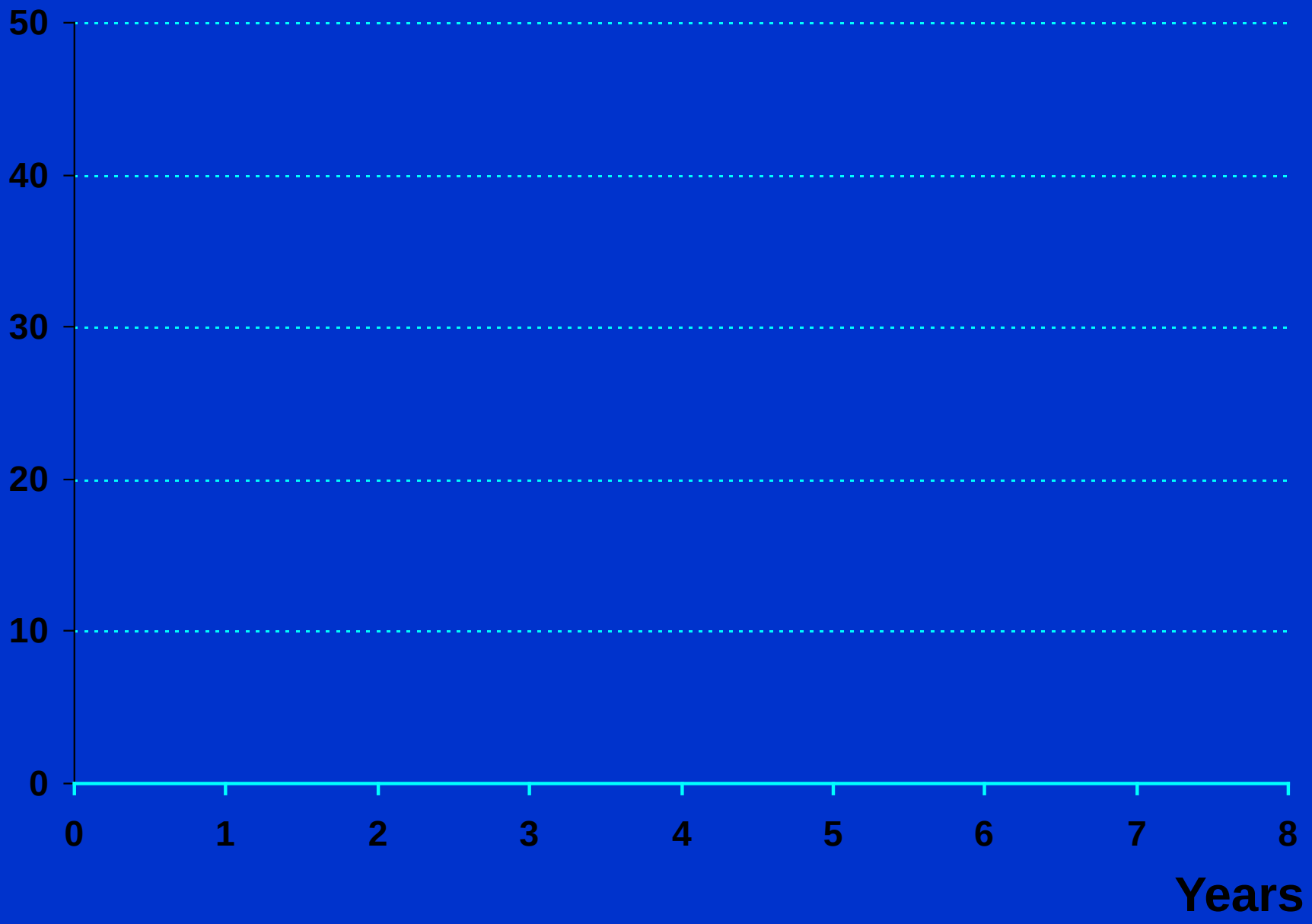


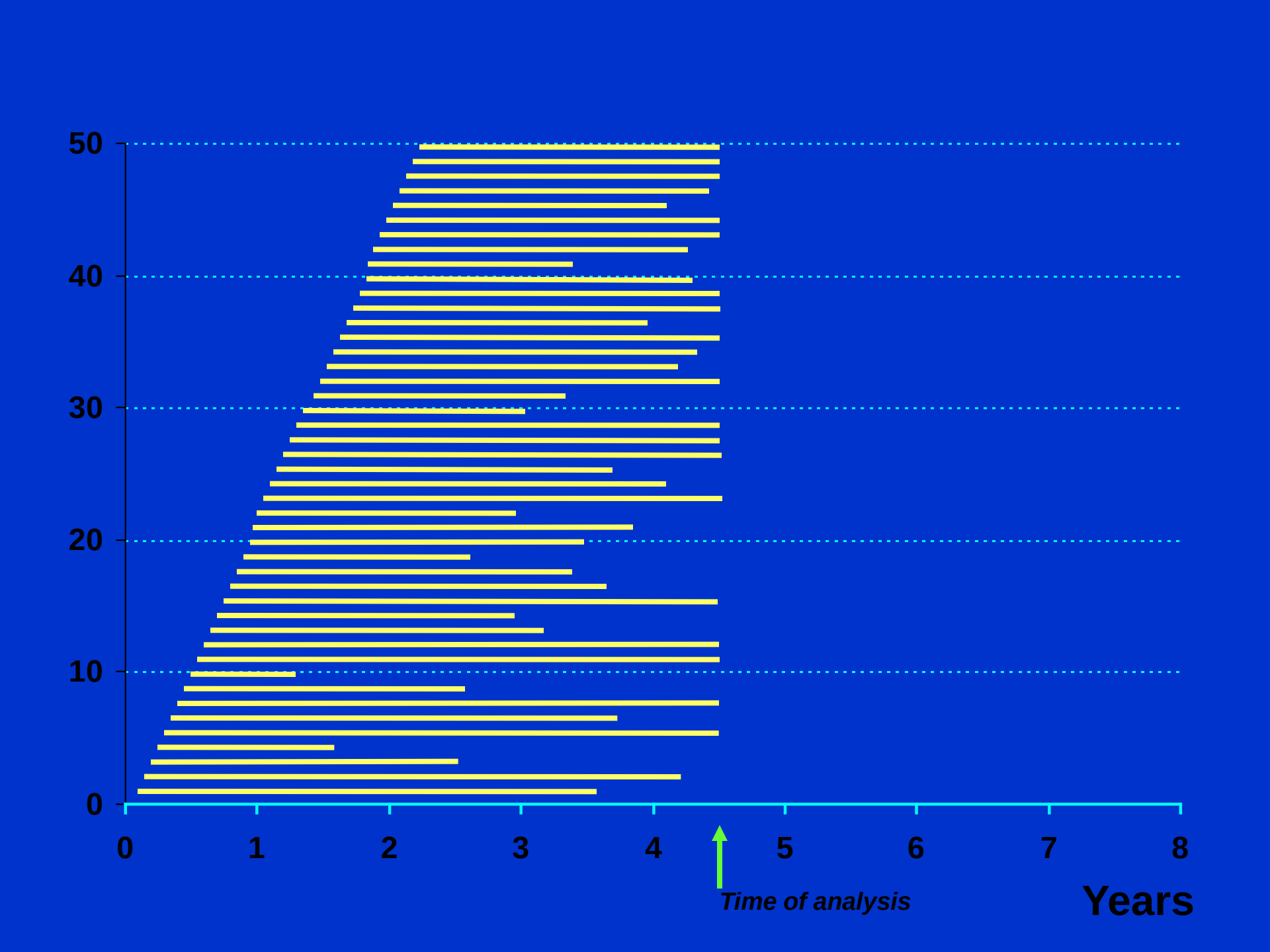
Interpretation of QOL Results: Example

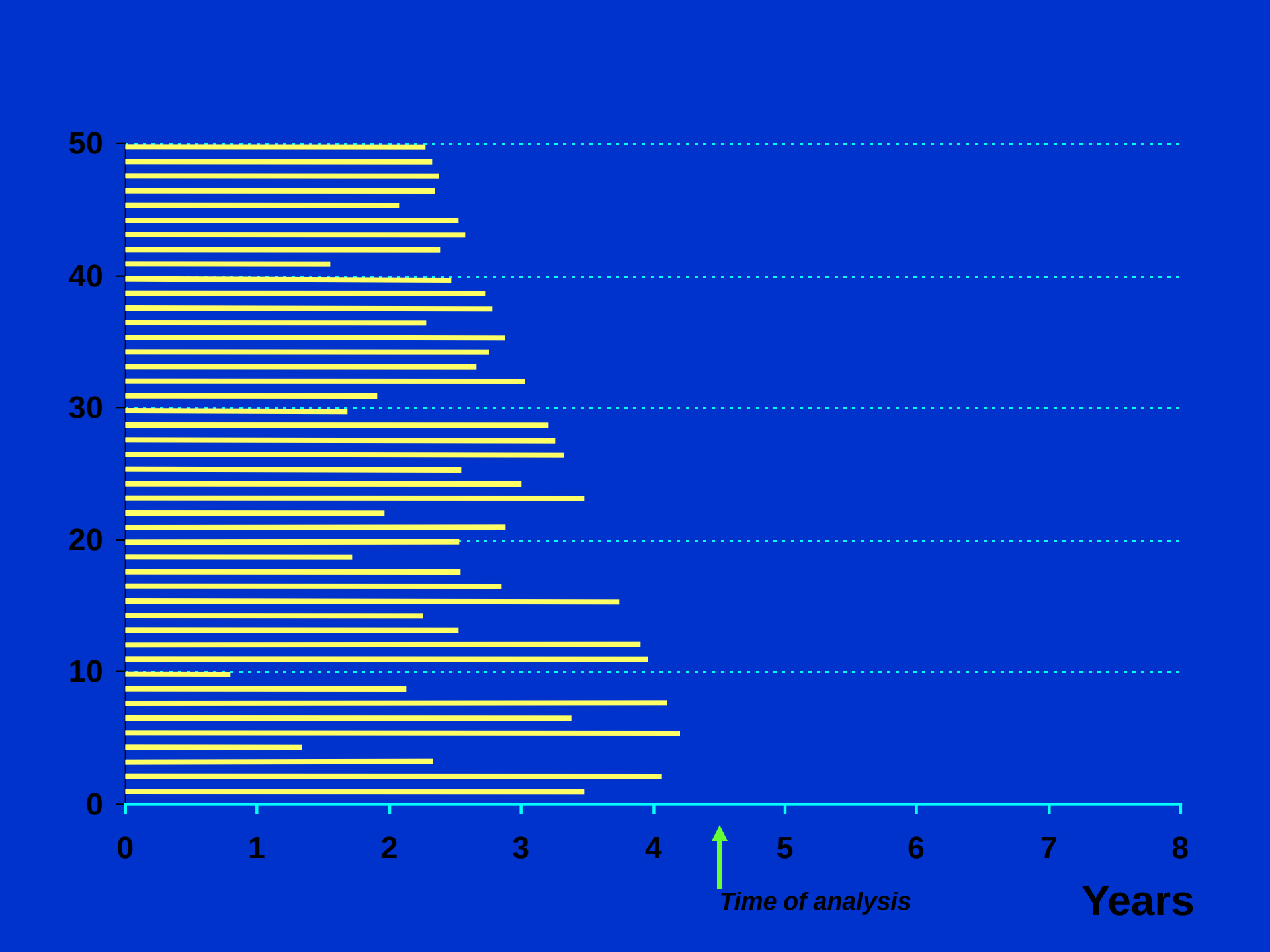
- A RCT claims to demonstrate that one treatment approach resulted in clinically superior QOL when compared with the other approach. How do you interpret this finding?
- Compare approach with a more familiar metric: survival

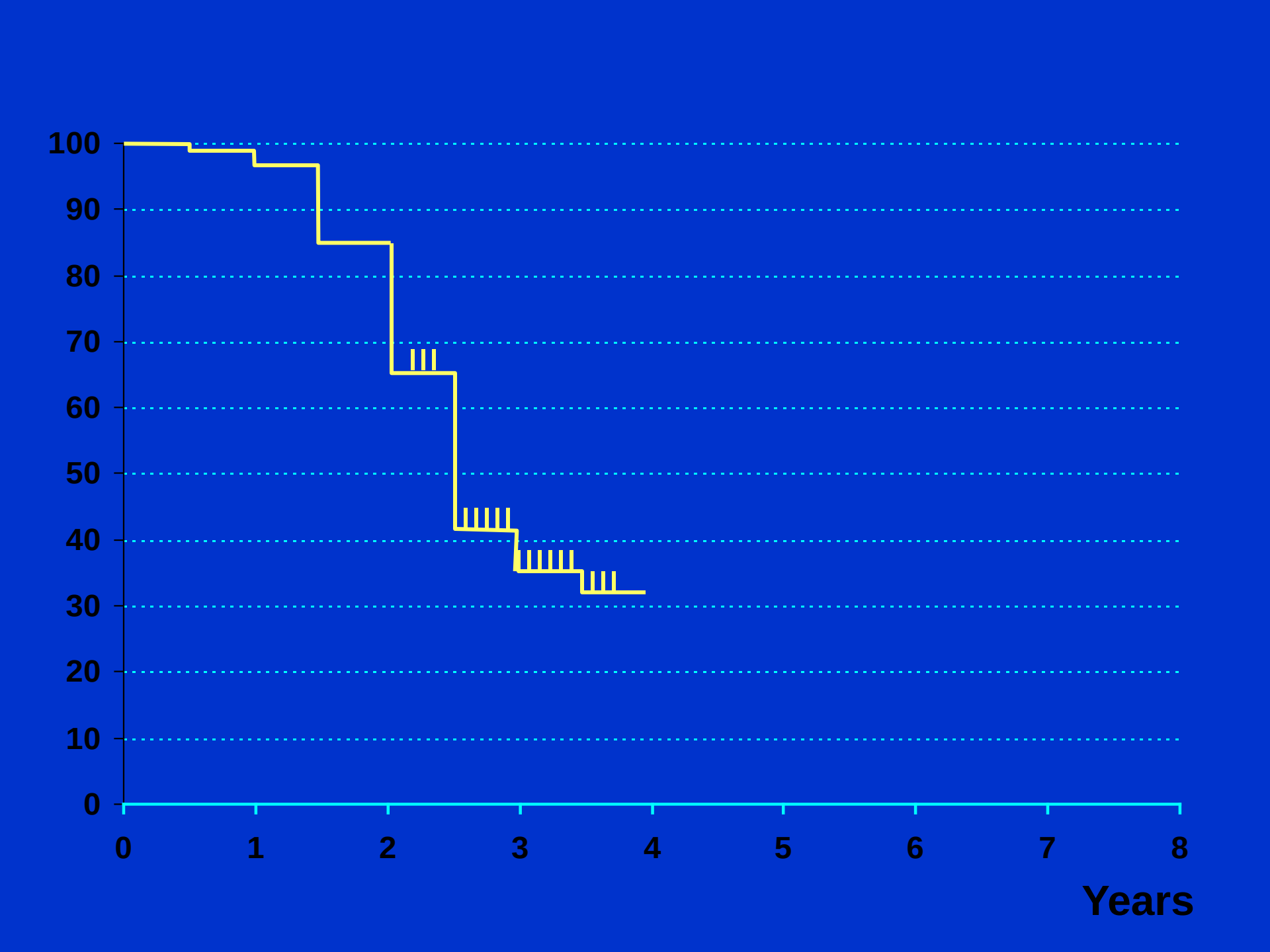
Example of an objective endpoint: Survival

- Patterns of survival on clinical trials are usually quite complex
- Clinically conveniently summarized by one or two statistics
  - E.g. Median survival, 3 year survival


50
40
30
20
10
0
0
1
2
3
4
5
6
7
8
Years

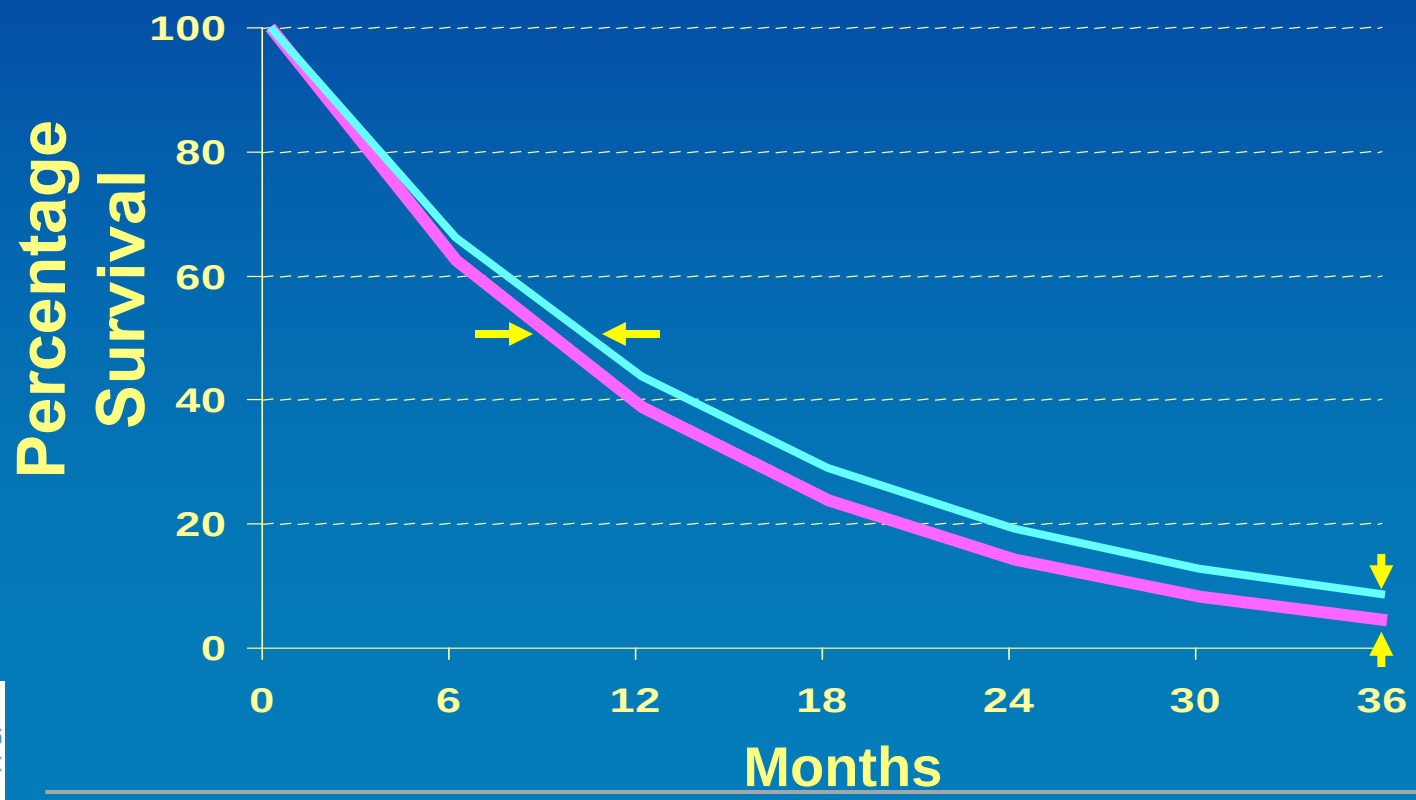

50
40
30
20
10
0
0
1
2
3
4
5
6
7
8
Time of analysis
Years


50
40
30
20
10
0
0
1
2
3
4
5
6
7
8
Time of analysis
Years


100
90
80
70
60
50
40
30
20
10
0
0
1
2
3
4
5
6
7
8
Years

Level 1 Evidence

Cambridge Meta-analysis, *BMJ* 1995



NCIC CTG
NCIC GEC

Example of a QOL “result”

- EORTC QLQ-C30+3 Instrument
- Domain: Global quality of life
- Patient questionnaire items:

How would you rate your overall health during the past week?

1 2 3 4 5 6 7

Very poor Excellent

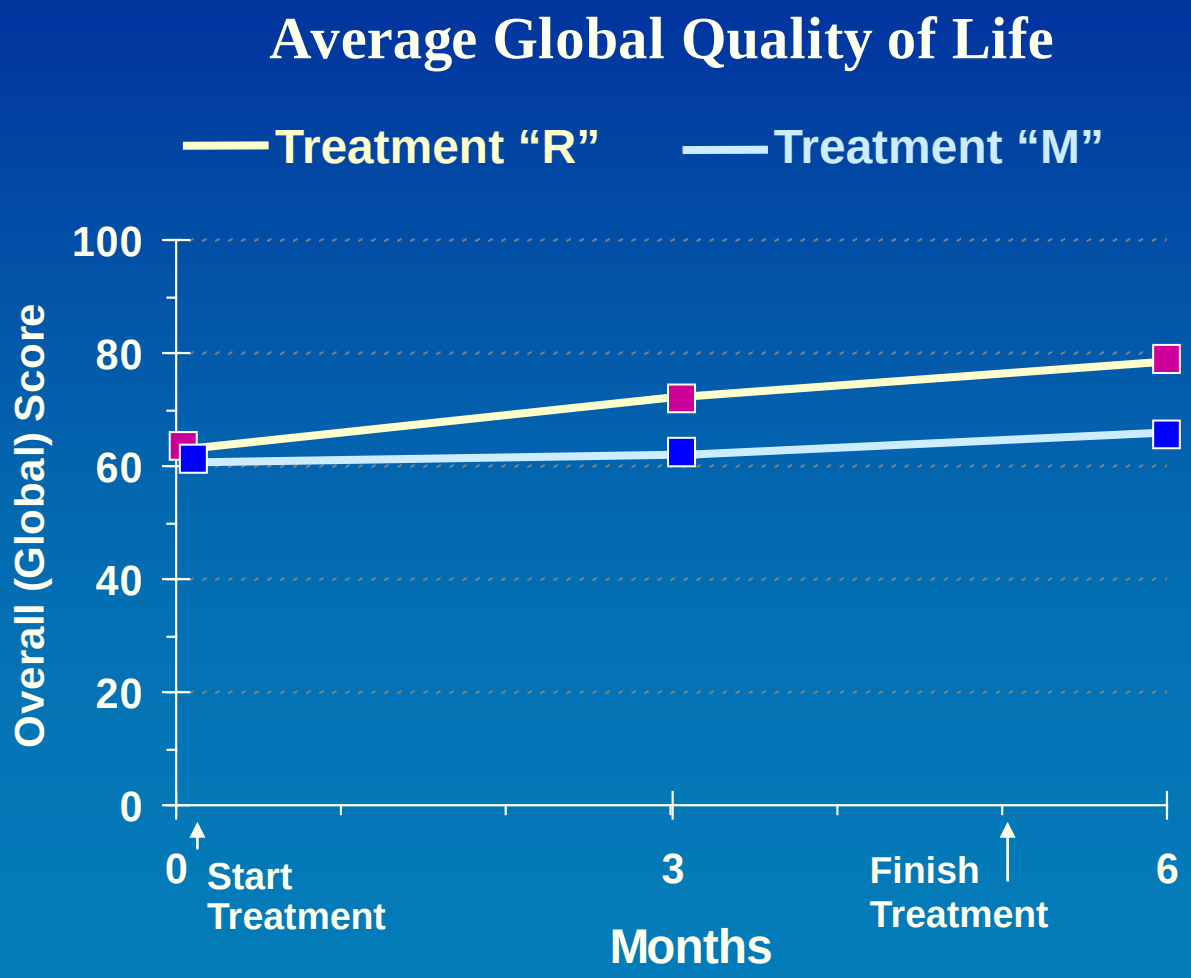
How would you rate your overall quality of life during the past week?

1 2 3 4 5 6 7

Very poor Excellent

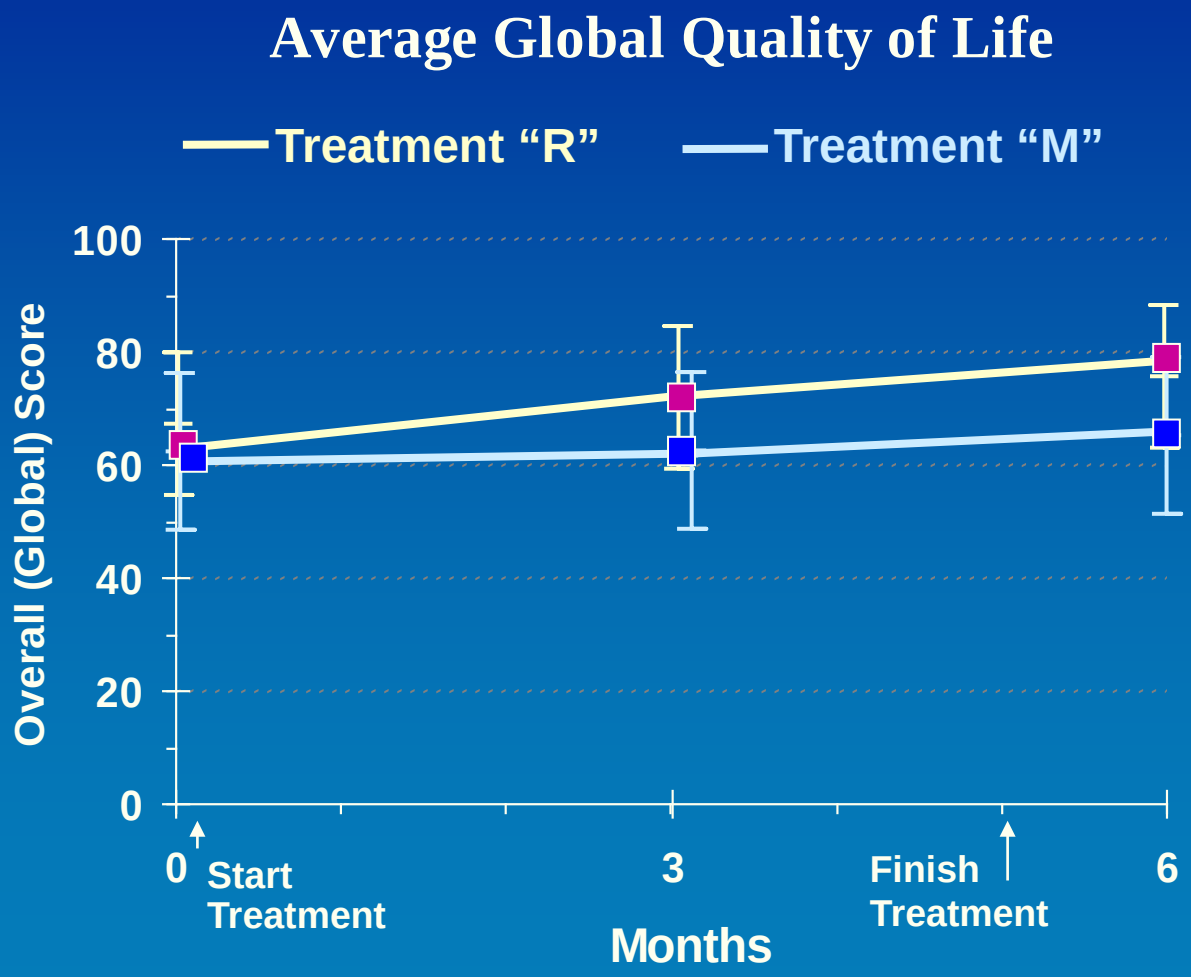
NCIC CTG
NCIC GEC

Example of a QOL “result”



NCIC CTG
NCIC GEC

Example of a QOL “result”



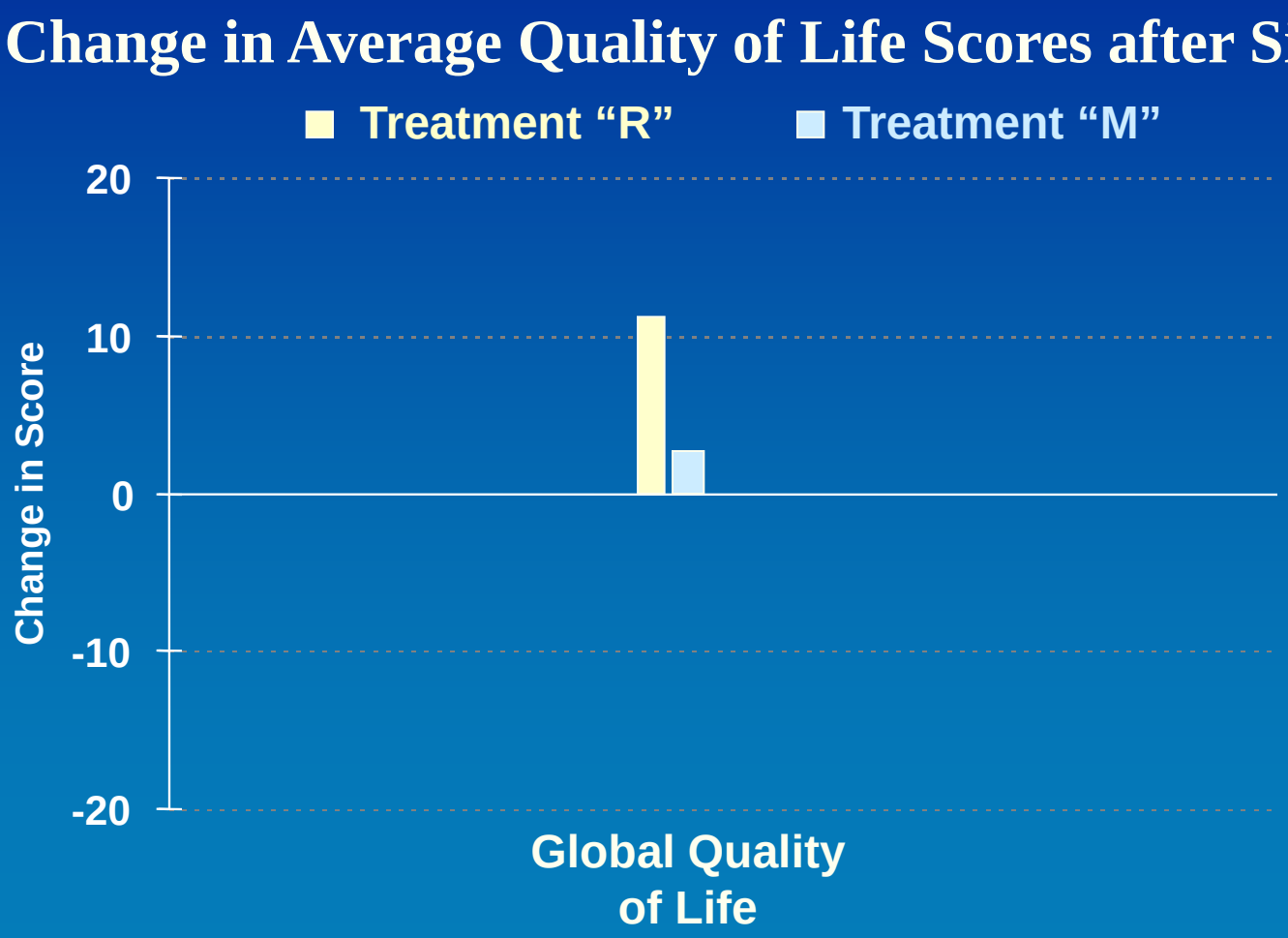
NCIC CTG
NCIC GEC

What “difference” is clinically significant?

- E.g.: Osoba et al, JCO 1998
  - Minimal change: 5-10 points
  - Moderate change: 10-20 points
  - Large change: >20 points

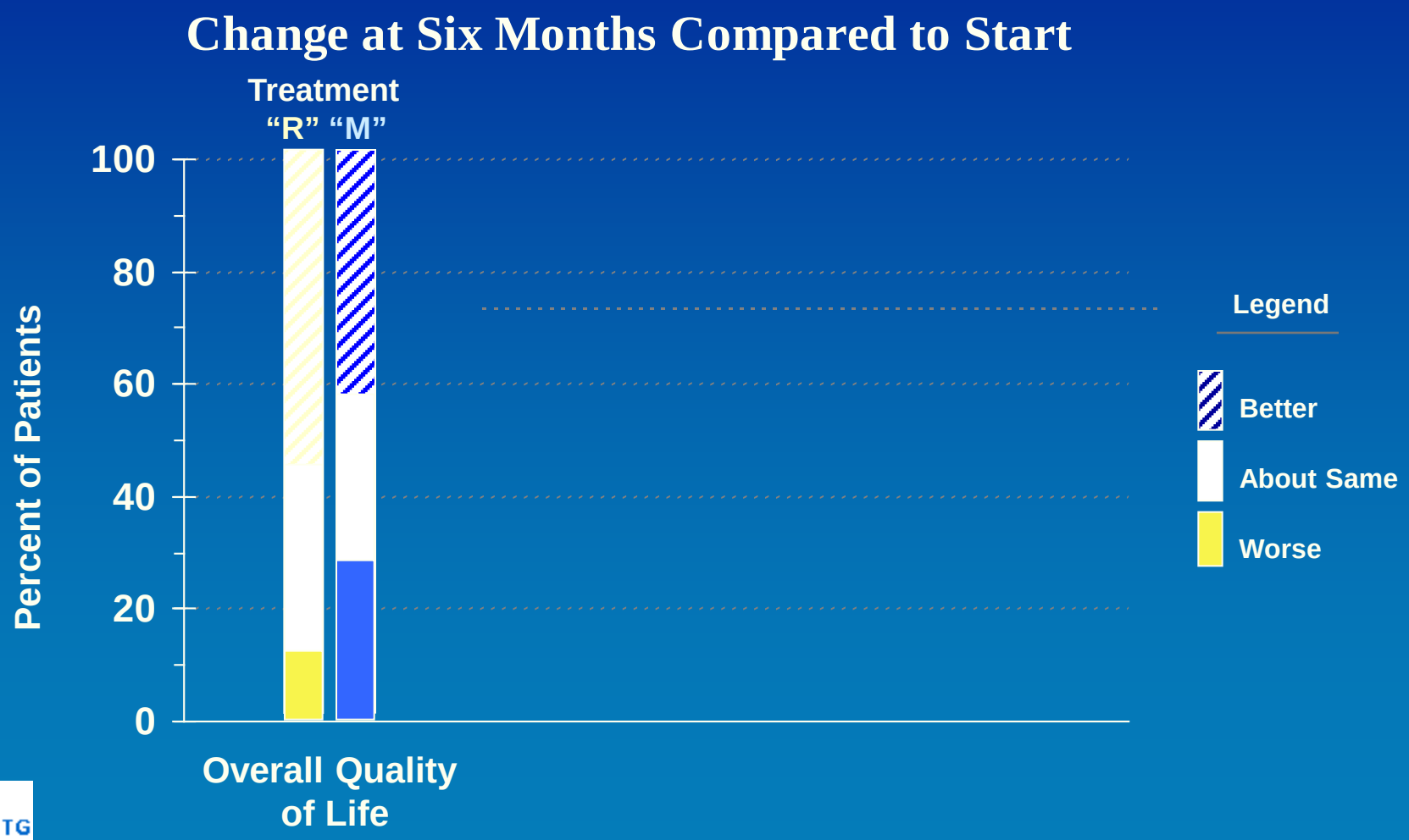
NCIC CTG
NCIC GEC

Same Data Presented Differently



NCIC CTG
NCIC GEC

Same Data Presented Differently



NCIC CTG
NCIC GEC

Some final thoughts…

Is there bias in the results?

Look for:

- Statement of a clear hypothesis for assessing HRQL
- An explanation of the choice of HRQL instruments
- A clear description of methodology
- Appropriate planning for handling and analyzing data

NCIC CTG
NCIC GEC

History of the NCIC CTG Quality of Life Committee

With thanks to Joe Pater and David Osoba

NCIC Clinical Trials Group - History

- 1979 - NCIC decides to establish a Clinical Trials Group
- 1980 - Joe Pater named Director of Group and headquarters moved to Kingston, Ontario
- 1982 - IND program established
- 1982 – BR.5 (1st QOL trial)
- 1981 - 2004 - Program expands through a series of site visits
- 2005 – Ralph Meyer appointed to succeed Joe Pater in 2007

NCIC CTG
NCIC GEC

BR.5

- 1982: two trials in advanced NSCLC appeared to show a survival advantage for chemotherapy
  - Cormier – MACC
  - Gralla - cisplatin/vindesine
- Best supportive care control arm
- Reviewer suggests should have a QOL endpoint
- “gold standard” instrument (the Sickness Impact Profile) along with a newly developed cancer instrument (FLIC)

BR.5 QOL

- Shortly after the trial started, centres were asked to participate in the QOL component of the trial
  - They were given the option to use both instruments, only FLIC or not participate
- Almost all centres agreed to participate and most chose to use both instruments
- Study completed 1986
- Survival benefit of chemotherapy

Audience Feedback

- What was the compliance rate with QOL questionnaire completion on BR.5?
  - >90%
  - 50-75%
  - 25-49%
  - <25%

NCIC CTG
NCIC GEC

Audience Feedback

- What was the compliance rate with QOL questionnaire completion on BR.5?
  - >90%
  - 50-75%
  - 25-49%
  - <25%

  Answer <25%

NCIC CTG
NCIC GEC

After BR.5

- The low compliance with QOL collection in BR.5 was due to many factors, including the fact the one of the key central office personnel had to leave due to illness in her family
- It was evident, though, the adequate QOL data collection would not just happen
- a “scientific session” was held at the 1986 spring meeting

Growth of QOL Committee

- Began as a Working Group in 1986
  - Symposium 1986 – Prof. Frits van Dam
  - Symposium 1987 – Dr. Neil Aaronson
  - Spring meeting 1988 – Dr. Jerome Yates
- Named as QOL Subcommittee in 1987
- Full Standing Committee - 1989 – present
  - Interim Chair, then Chair – D. Osoba ’86-’95
  - Chair – Andrea Bezjak ’95- 2006
  - Co-Chair – Michael Brundage ’03 - present
    - Jolie Ringash 2006-present

NCIC CTG
NCIC GEC

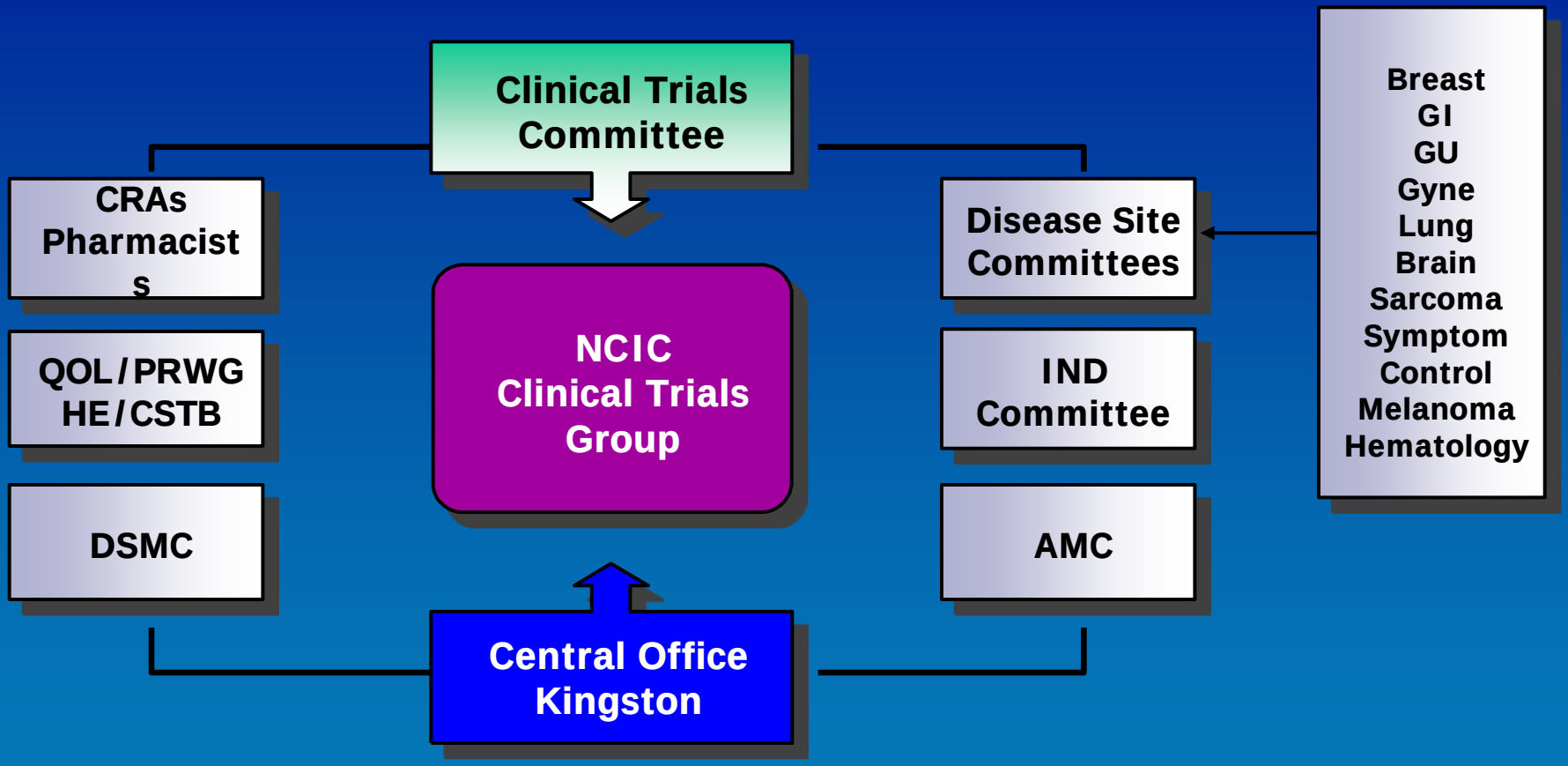
Developments Within QOL Committee

- Chose EORTC QLQ-C30
- Developed a policy re: QOL assessment in 1988; adopted in 1989
- “There should be a statement about the anticipated impact on QOL with every proposed phase III clinical trial and whether or not QOL measures will be incorporated in the protocol.”

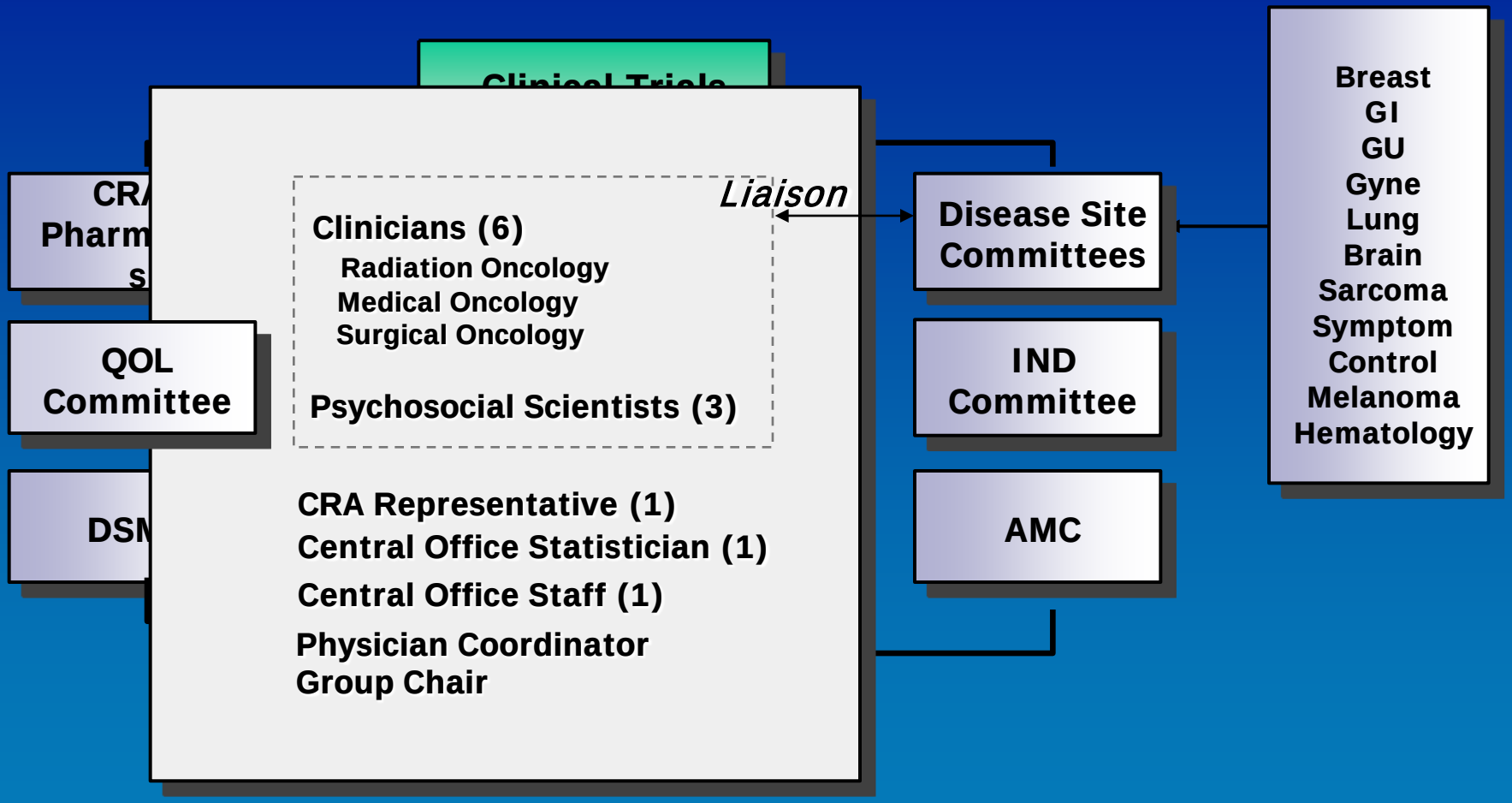
Developments Within QOL Committee

- Named members of the QOL Committee to act as Disease Site liaisons – usually as members of Disease Site Committees
- Maintained a liaison with the EORTC Quality of Life Study Group since 1987
- Maintained contact with several cooperative CTGs and NCI in USA
- First trial, ME.7 – an adjuvant trial of levamisole vs gamma interferon in malignant melanoma - November, 1988

Current Structure



Current Structure



NCIC CTG
NCIC GEC

QOL questionnaire	Number of Studies
EORTC QLQ-C30	35
SF-36	6
McMaster BCQ	1
FACT	6
SWOG Distress scale	1
Spitzer QOL index	1
Lung Cancer Symptom Scale (LCSS)	1
Brain Tumor QOL questionnaire	1
Menopause QOL questionnaire	1
PROSQUALY	1
Toronto Extremity Salvage Score	1
McMaster Head and Neck XRT questionnaire	1
NCCTG Symptom Distress Scale	1

Summary

- QOL refers to overall well-being, as reported by the patient
- There is a science of measurement which applies to QOL
- Interpretation of results is important
- NCIC-CTG has been a leader in QOL research

NCIC CTG
NCIC GEC

Dr. Jolie Ringash, Co-Chair, NCIC-CTG QOL Committee

Quality of Life Measurement for NCIC-CTG Clinical Trials

Traditionally, the outcome of cancer care was assessed in terms of survival and/or tumour response. As early as 1948, Karnofsky recognized that other outcomes were important to patients. In his study, "subjective improvement was indicated by the patient's feeling of well-being, his increased appetite and strength, and the relief of specific complaints…",[1] by applying a performance status scale which is still in use. In the intervening 60 years, his initial concept of patient well-being has been expanded into our modern conception of quality of life (QOL). QOL is now recognized as an important outcome of cancer care.

Definition

Broadly speaking, QOL is a measure of an individual's overall personal well-being. Three aspects critical to the concept are subjectivity (only the individual truly knows his or her own internal state), multi-dimensionality, and sociocultural context.

QOL and "Health-related" QOL

Overall QOL is impacted by issues such as income and adequacy of housing, which cannot typically be influenced by the health-care system. In the context of health care, QOL measures are often used to measure the effect of disease, illness, and treatment on the patient and family. For this purpose, issues which are not expected to change based on these effects become measurement "noise", and reduce the ability of questionnaires to detect actual changes. For this reason, the more limited concept of "health-related" QOL is usually applied. The WHO has defined it as: "an individual's perception of their position in life in the context of the culture and value systems in which they live and in relation to their goals, standards and concerns. It is a broad ranging concept affected in a complex way by the person's physical health, psychosocial state, level of independence, social relationships, and

their relationships to salient features of their environment."[2] When the term "quality of life" is used in the context of health care, it is usually health-related QOL which is meant.

Domains and Multi-dimensionality

Human beings are complex; the overall human experience reflects many underlying functions and roles. Under the stress of illness, that experience is influenced as well by specific symptoms. Such complexity may be addressed by two very different methods. The first method attempts to explicitly address the many dimensions of experience by constructing specific "domains" within a questionnaire, such as measures of particular symptoms, as well as cognitive, emotional, social, spiritual, role and physical functioning. This approach results in long questionnaires with multiple items organized into separate sub-scales relating to each domain. The alternative method is to rely upon the respondent's ability to internally integrate his or her experience, and to report overall QOL as a single item index. One example would be the use of visual analogue scales, such as the "feeling thermometer", originally developed in 1964 by the United States National Election Services to allow voters to rate their feelings toward political candidates, but more recently adapted as a health utility instrument.[3, 4] Some instruments use a mixture of both methods; for example, "overall QOL" may be included as a single item, along with more specific domains. Typically, multi-item instruments are more reliable and more sensitive to change over time than single items, however they require more time to complete.

Measurement of QOL: Basic Methodology

QOL instruments measure a subjective concept, but their measurement properties are based on sound scientific principles. Psychometrics, the science of indirect measurement through questionnaires and other related instruments, evolved in educational and psychology research over the course of the 20th century. It has been applied to health-related questionnaires and

PROs for over 20 years.[5] Instruments chosen for use in clinical research should adhere to the principles outlined below.

Item Generation should incorporate information about the issues of importance to patients from literature review, health professional expertise, and direct input from patients similar to the instrument's target population. Questions should be written at an appropriate educational level; grade 6 is often recommended. [6] Items should be formatted in a standard way, including both positively- and negatively-worded items, and avoiding jargon, skip formats and double-barreled questions. Utilization in other languages and cultural groups requires a formal process of cultural adaptation, including forward- and back-translation, pilot and field testing in the new language/culture.[7]

Item Reduction is often required to produce a questionnaire of practical length, but which remains sufficiently sensitive to change over time for *evaluative* (longitudinal) use. Direct testing in patients is typically carried out to identify the items most frequently endorsed by patients, and ranked as being of the greatest importance. Statistical methods may also be used to identify items which are most informative.[8]

Questionnaire Design includes principles of readability and clarity. Questionnaires should include a large proportion of white space, with font size and type which is easy to read. Special requirements for the target group need to be considered (eg. the visually impaired, young children, low-literacy populations, etc.).

Indices and Profiles

Controversy exists regarding the relative preferability of *indices* or *profiles* for QOL measurement. Different individuals may apply personal weights to aspects of their quality of life, so summation of scores over multiple domains, as is done for indices, may impose the developer's values inappropriately on the patient. Exploration of individual, patient-assigned weighting has proven cumbersome and is rarely used. Other instruments present scores

separately for each domain (profiles), without summation. Popular questionnaires of both types are currently in use.

Reliability refers to the reproducibility of scores. It may be assessed by repeated administration of the instrument to a population with stable QOL (test-retest reliability), or by correlation of items within a questionnaire (internal consistency). Higher levels of reliability coefficients are conventionally required for *evaluative* use (to measure change in individuals over time) than in *discriminative* use (to measure difference between groups of patients); typically, 0.7 and 0.8 respectively for internal consistency.[9, 10]

Validity refers to the ability of a questionnaire score to reflect the actual concept of interest. It is important that a "QOL questionnaire" is actually related to the patient's overall well-being during a defined period (eg. one week), and not his or her momentary comfort or passing mood. Questionnaire validation lacks a gold standard, so validity is defined by hypothesis testing with respect to convergence or divergence from other findings (concurrent validity). For example, QOL scores might be expected to be better in patients with better performance status, and to improve over time in patients who were gaining weight post-treatment. A disease site-specific QOL questionnaire would also be expected to show a moderate correlation with other, more general, QOL or utility instruments. It is important that validation studies included patients similar to those for whom the instrument will be used; a questionnaire validated exclusively in surgically-treated patients may not exhibit the same measurement properties in chemoradiation patients.

Responsiveness is the sensitivity of the instrument to changes over time in an individual patient. Responsiveness is inversely correlated with instrument length and directly correlated with the specificity of items. A very detailed, disease site-specific QOL instrument would be highly responsive, whereas a short, general QOL instrument would be less

responsive, to change in a patient with a given cancer type. Prospective evaluation is required to determine instrument responsiveness.

Minimal Clinically Important Difference (MID) is defined as the smallest change in value on a measurement instrument, which, from the point of view of the patient, represents an important rather than trivial change. In practice, it has been estimated for groups by the use of the minimal detectable difference, that is, the smallest difference which is detectable by the average patient.[11] It is important to differentiate this clinical concept from statistically significant differences, which reflect only the likelihood of observing a given difference, not what it may or may not mean to a patient. Ideally, MID should be determined for every new instrument; however, several studies suggest that a change of 5-10% of instrument range may represent the MID for many instruments.[12, 13]

Interpretation of QOL Results

Each individual conceptualizes QOL in a personal way. Life experience, optimism or pessimism, and psychological state all contribute to the perception of QOL. Consequently, cross-sectional comparisons among individuals are subject to measurement "noise" which should be less problematic when patient scores are self-controlled, by calculating one individual's change in QOL over time in a longitudinal study. For this reason, if QOL is to be used as an outcome of a treatment in a clinical trial, prospective measurement at multiple timepoints is preferred. However, it is important to realize that the baseline administration usually occurs soon after a patient has received a cancer diagnosis, or has been found to have disease recurrence or progression. Thus, the "baseline QOL" does not reflect that person's QOL when healthy. QOL scores that return to baseline over a period of time cannot be interpreted as indicating a resolution of tumour- and treatment-related effects; in many cases, the patient may, in fact, have exchanged tumour-related impairments for different problems induced by treatment.

Response Shift

An additional important consideration in the interpretation of longitudinal QOL data relates to response shift, or changing internal standards.[14] Over time, an individual confronted with critical illness may modify his or her values, or standards of measurement, and may also reconceptualize QOL entirely. Response shift may play a role in some initially unexpected findings, such as the fact that patients with serious illness will routinely rate their own QOL as better than the ratings applied to them by surrogates (eg. family members or health care professionals). Response shift may be viewed as a beneficial adaptive process, however it also introduces an additional source of measurement error. Methods of quantifying response shift exist, but are labour-intensive. One approach to descriptive studies is to compare QOL results with population norms drawn from healthy individuals.[15] Once again, the randomized trial design is favoured for studies with QOL outcomes, since it is hoped that unmeasured covariates such as response shift should be balanced between the arms by chance.

Compliance and Missing Data

Results of any study must be assessed for two types of validity: internal validity (does the study measure what it says it does?) and external validity (generalizability). In QOL studies, compliance with planned questionnaires and missing data can threaten both types of validity. Patients self-select study participation, which influences external validity (ie., study results are applicable only to the type of patients who agreed to participate). Once enrolled in the study, participants determine whether or not they complete requested evaluations. Certain questions or even pages of a given questionnaire may not be completed, or the entire questionnaire may have been missed, either because the patient did not attend a scheduled appointment, or because he or she attended but did not complete the QOL instrument. Missed questionnaires threaten both types of validity, since reported results do not really reflect the experience of ALL patients in the study. Specifically, it has been shown that healthier patients are more

likely to comply with QOL assessments.[16] While statistical methods exist to attempt to correct for missing data, they require the assumption that data is missing at random, which is known to be unlikely in QOL studies. Consequently, every effort should be made to maximize compliance in QOL studies. Strategies to do so include adequate resources, education and feedback for those administering the questionnaires, real-time monitoring of compliance, and back-up methods of administering questionnaires if an error is detected within an acceptable time window.[17]

Mean Changes versus Response Analyses

Longitudinal studies may report mean change in an overall group, however this can overestimate longer-term QOL due to "survivor effect": data from all patients will be included at baseline, but only patients who survive and continue to comply with assessments are included in follow-up. In comparison of two trial arms, it is even possible that the QOL may appear to be better in the arm with fewer survivors, since a more toxic treatment may selectively eliminate those with poorer QOL. One alternative is to pre-specify the QOL hypothesis and MID, and analyze QOL response. Each participant is categorized according to "improved", "stable", or "worsened" QOL, and arms are compared for proportion of patients with a QOL benefit.[18] This approach also allows calculation of a number needed to treat (NNT) statistic.[19]

Knowledge Translation

The concept of knowledge translation refers to the gap between evidence and practice.[20] Awareness, agreement, adoption and adherence have been proposed as the necessary steps required before clinicians will use new knowledge. A prerequisite of both awareness and agreement is that information must be presented in a manner which is interpretable and usable. This has been a challenge for QOL data.[21] Two User's Guides have been published to assist the clinician with evaluating and interpreting QOL results.[22, 23] In addition, two papers

have provided lists of study details which should be included in publications of QOL results.[24, 25] However, a recent review showed that in recent publications of oncology clinical trials, recommended information items were included 10-70% of cases; a trend to improvement of most data points was seen over time.[26] Additional research is needed to help bridge the current gap between QOL researchers and oncologists in the clinical setting.

History of the NCIC-CTG QOL Committee

(with thanks to Drs. Andrea Bezjak, Michael Brundage, David Osoba and Joseph Pater)

In 1982, just 2 years after Dr. Pater was named as inaugural Director of Canada's first national cancer clinical trials group, it was decided to measure a QOL endpoint in a trial of metastatic NSCLC partients. The study, BR.5, compared chemotherapy to best supportive care and used a generic QOL instrument, the Sickness Impact Profile (SIP) as well as a cancer-specific QOL scale, the (FLIC), as secondary outcomes. The study was completed in 1986 and showed a definitive survival benefit to chemotherapy [27]. However, the QOL assessment was not successful – the compliance was < 25%! Several problems were identified: 1) the QOL outcome was initiated after the overall study had already begun, 2) QOL measurement was optional, 3) participating centres were given the choice of using either or both of the questionnaires, 4) limited education/training about how to administer the QOL instruments was available, and 5) a key central office member who had championed the QOL assessment had to leave mid-way through the study due to a family issue. It was recognized that successful measurement of QOL would require greater infrastructure and expertise.

In 1986, a QOL Working Group was established. Visiting experts assisted the group in 1986 (Dr. Frits van Dam), 1987 (Dr. Neil Aaronson), and 1988 (Dr. Jerome Yates). In 1987, a QOL Sub-committee was established with Dr. David Osoba as Chair, which became a full standing committee in 1989. In that same year, the CTG established the following policy:

"There should be a statement about the anticipated impact on QOL with every proposed phase III clinical trial and whether or not QOL measures will be incorporated in the protocol."

Another important innovation was establishing QOL liaisons, members of the QOL committee with measurement expertise who were appointed to each disease site group, to assist with development of QOL endpoints for their proposed trials. The first successful trial with a QOL outcome was a melanoma trial, ME.7, which opening in 1988 [28]. Dr. Andrea Bezjak was Chair from 1995-2006, with Dr. Michael Brundage as Co-Chair from 2003. Current Co-Chairs are Dr. Brundage and Dr. Jolie Ringash; Dr. Osoba remains an active member of the committee.

Through the years, the NCIC-CTG QOL Committee has maintained close ties to the EORTC QOL Committee and Group. Additionally, our members have participated in initiatives with the NCI (US), several cooperative groups, and in selected industry trials. Over the years, influential papers have been published outlining our general approach to QOL compliance [17], an our analysis approach[18]. A workshop defining the "added value" of QOL assessment for clinical trials was held with Dr. Aaronson of the EORTC in 2008, resulting in a taxonomy of added value which is currently being prepared for publication. Our trials group has been recognized internationally for our success in the measurement and reporting of QOL.

References:

1. Karnofsky DA, Abelmann WH, Craver LF, et al. The use of nitrogen mustards in the palliative treatment of carcinoma. Cancer. 1948;1(4):634-656.

2. Rehabilitation after cardiovascular diseases, with special emphasis on developing countries. Report of a WHO Expert Committee. World Health Organ Tech Rep Ser. 1993;831:1-122.

3. National Election Studies, 1964: Pre/Post Election Study. In: Political Behavior Program tSRCotIoSR, ed. Ann Arbor: University of Michigan, Center for Political Studies; 1999.

4. Bennet KJ, Torrance G. Measuring health state preferences and utilities: rating scale, time trade-off, and standard gamble techniquesSpilker B,ed:Quality of life and pharmacoeconomics in clinical trials. Philadelphia: Lippincott-Raven; 1996

5. Feinstein AR. Clinimetrics. New Haven CT: Yale University Press; 1987

6. Woodard CA, Chambers LW. Guide to Questionnaire Construction and Question Writing. In: Canadian Public Health Association, ed. 1st edition, ed. Ottawa; 1980.

7. Cull A, Sprangers M, Bjordal K, et al. on behalf of the EORTC Quality of Life Group. EORTC Quality of Life Group Translation Procedure. Brussels: EORTC; 2002.

8. Reeve B, Hays R, Bjorner J, et al. Psychometric evaluation and calibration of health-related quality of life item banks: plans for the Patient-Reported Outcomes Measurement Information System (PROMIS). Medical care. 2007;45(5 Suppl 1):S22-31.

9. Nunnally JC. Psychometric Theory. New York McGraw Hill; 1978

10. Weiner EA, Stewart BJ. Assessing individuals: Psychological and educational tests and measurements. Boston (MA): Little Brown; 1984

11. Wright JG. The minimal important difference: who's to say what is important? J Clin Epidemiol. Nov 1996;49(11):1221-1222.

12. Ringash J, Jolie|O'Sullivan,B,Brian|Bezjak,A,Andrea|Redelmeier,DA,Donald A. Interpreting clinically significant changes in patient-reported outcomes. Cancer; 2007; 110:196-202.

13. Norman GR, Sloan JA, Wyrwich KW. Interpretation of changes in health-related quality of life: the remarkable universality of half a standard deviation. Med Care. May 2003;41(5):582-592.

14. Sprangers MAG, Schwartz CE. Integrating response shift into health-related quality of life research: a theoretical model Vol 48: Social Science & Medicine; 1999

15. Holzner B, Kemmler G, Cella D, et al. Normative data for functional assessment of cancer therapy--general scale and its use for the interpretation of quality of life scores in cancer survivors. Acta oncologica. 2004;43(2):153-160.

16. Guren MG, Dueland S, Skovlund E, et al. Quality of life during radiotherapy for rectal cancer. European journal of cancer. 2003;39(5):587-594.

17. Osoba D, Zee B. Completion rates in health-related quality-of-life assessment: approach of the National Cancer Institute of Canada Clinical Trials Group. Stat Med. 1998;17(5-7):603-612.

18. Osoba D, Bezjak A, Brundage M, et al. Analysis and interpretation of health-related quality-of-life data from clinical trials: basic approach of The National Cancer Institute of Canada Clinical Trials Group. European journal of cancer. 2005;41(2):280-287.

19. Laupacis A, Sackett DL, Roberts RS. An assessment of clinically useful measures of the consequences of treatment. N Engl J Med. 1988;318(26):1728-1733.

20. Davis D, Evans M, Jadad A, et al. The case for knowledge translation: shortening the journey from evidence to effect. BMJ. British medical journal (Clinical research ed.). 2003;327(7405):33-35.

21. Bezjak A, Ng P, Skeel R, et al. Oncologists' use of quality of life information: results of a survey of Eastern Cooperative Oncology Group physicians. Qual Life Res. 2001;10(1):1-13.

22. Guyatt GH, Naylor CD, Juniper E, et al. Users' guides to the medical literature. XII. How to use articles about health-related quality of life. Evidence-Based Medicine Working Group. JAMA. 1997;277(15):1232-1237.

23. Sprangers MA, Moinpour CM, Moynihan TJ, et al. Assessing meaningful change in quality of life over time: a users' guide for clinicians. Mayo Clin Proc. 2002;77(6):561-571.

24. Lee CW, Chi KN. The standard of reporting of health-related quality of life in clinical cancer trials. J Clin Epidemiol. 2000;53(5):451-458.

25. Efficace F, Bottomley A, van Andel G. Health related quality of life in prostate carcinoma patients: a systematic review of randomized controlled trials. Cancer. 2003;97(2):377-388.
26. Brundage M, Bass B, Davidson J, et al. Patterns of reporting HRQL outcomes in randomized clinical trials of cancer therapies: implications for clinicians. International Society for Quality of Life Research meeting abstracts. The QLR Journal. 2007;A-31.
27. Rapp E, Pater JL, Willan A, et al. Chemotherapy can prolong survival in patients with advanced non-small-cell lung cancer--report of a Canadian multicenter randomized trial. Journal of clinical oncology. 1988;6(4):633-641.
28. Quirt IC, Shelley WE, Pater JL, et al. Improved survival in patients with poor-prognosis malignant melanoma treated with adjuvant levamisole: a phase III study by the National Cancer Institute of Canada Clinical Trials Group. Journal of clinical oncology. 1991;9(5):729-735.