The reason why I'm unwilling to use QOL assessment for cancer clinical trials

Haruhiko FUKUDA
JCOG Data Center
National Cancer Center Hospital
Tokyo, Japan

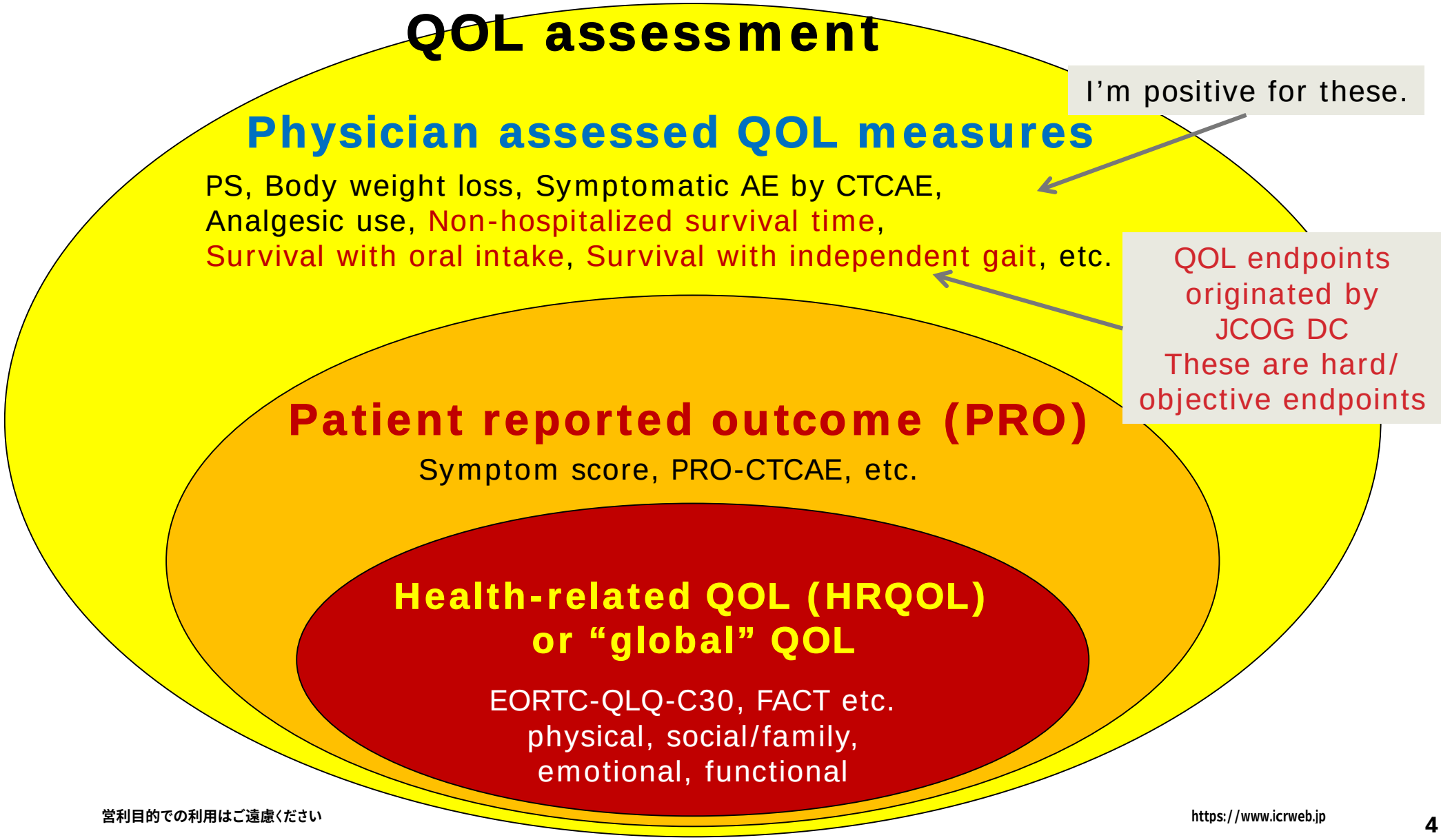
Preface / Disclaimer

- My talk today is based on my personal opinion, is NOT a vision of JCOG Data Center.
- My criticism on HRQOL/PRO is for those in open-label oncology trials, especially in confirmatory phase III trials to decide standard treatment, NOT for entire PRO nor PRO in masked(blinded) trials and clinical practice.
- Please do not confuse ‘to care about patients’ QOL as a physician in clinical practice and clinical trials’ and ‘to study using QOL questionnaire’. They are often confused, but are completely different things, personally I think.

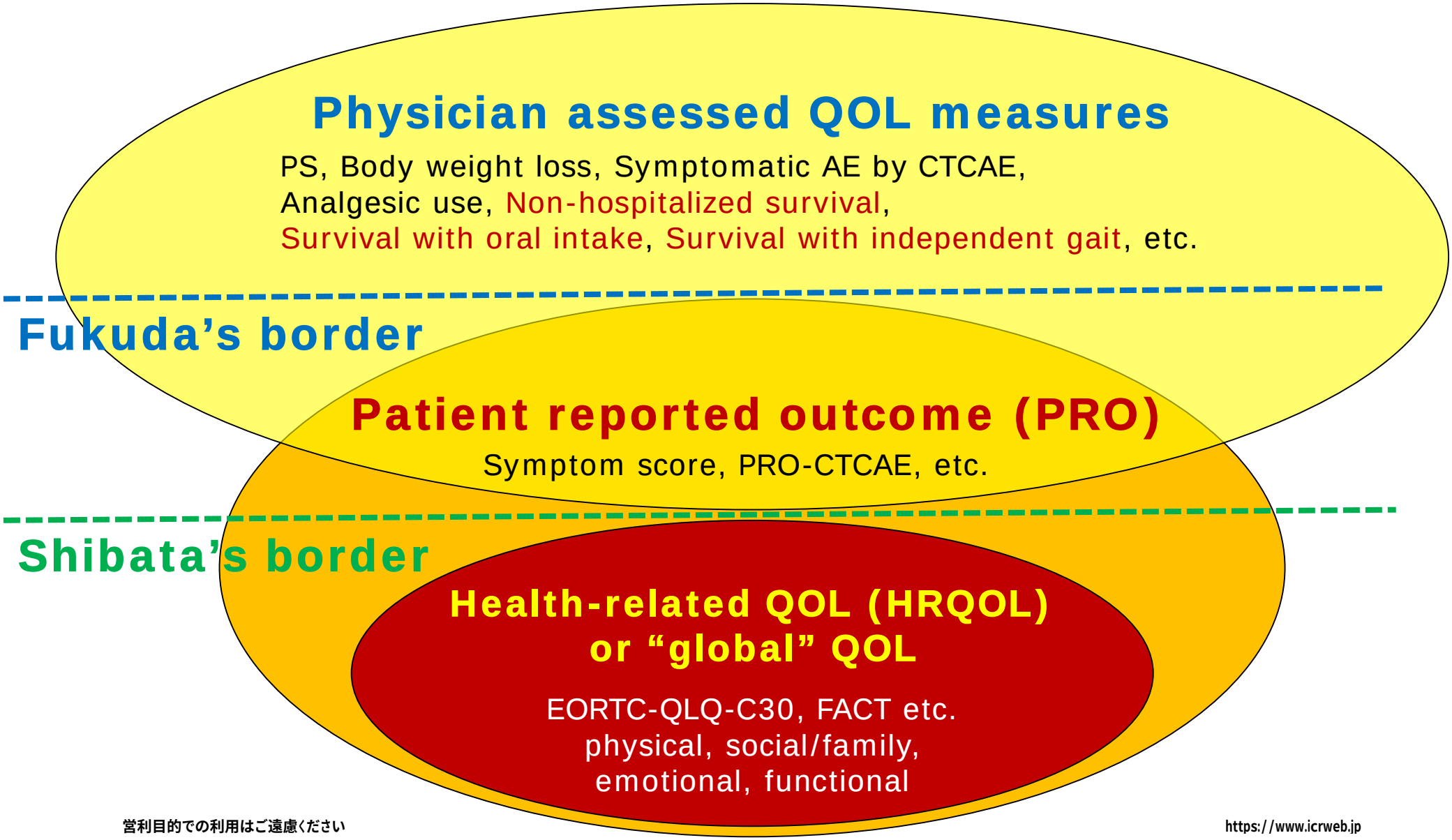
Outline

- Category of QOL assessment
- Clinical meaning
- Information bias
- Physician's underestimation
- Usefulness in choice of treatment
- Respondent burden
- Other's opinion
  - ASCO Value in Cancer Care Task Force
  - FDA
- Conclusions

Category of QOL assessment: Fukuda's view



Category of QOL assessment: Shibata's view



What is measured? – clinical meaning

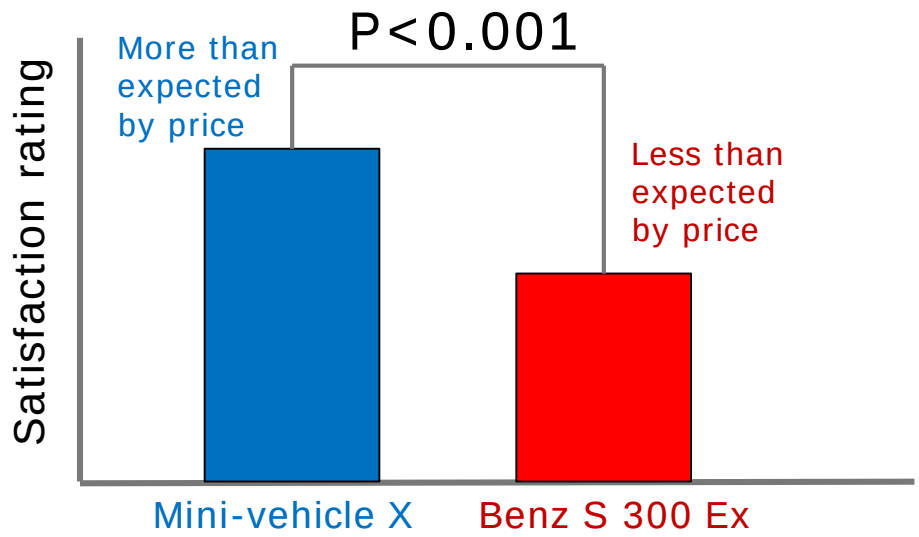
- 5-year overall survival: 30% vs. 40%
  - 30/40 patients out of 100 are surviving over 5 years after receiving treatment A/B
  - Treatment B is 10% more effective in terms of survival
- Treatment related death: 5% vs. 10%
  - 5/10 patients out of 100 died due to treatment A/B
  - Treatment B is more life-threatening
- Grade 3 sensory neuropathy: 10% vs. 20%
  - 10/20 patients out of 100 suffered numbness disturbing activity of daily living
  - Treatment B is more unfavorable for daily living
- Mean QOL score: 60/100 vs. 70/100
  - WHAT is better in treatment B?

Explainable and understandable for everyone

Clinical meaning of 10 points of QOL score is never explainable. Unexplainable information is useful for decision making?

Rightly measured? – information bias

- Imagine… To decide which the more convenient car is
  - Test drive on Mini-vehicle X: ‘Price is 8,000 $’
  - Test drive on Benz S 300 H Exclusive: ‘Price is 120,000 $’



Mean satisfaction rating was significantly higher in Mini-vehicle X.

Therefore, Mini-vehicle X is more convenient car than Benz S 300 H

Do you agree?

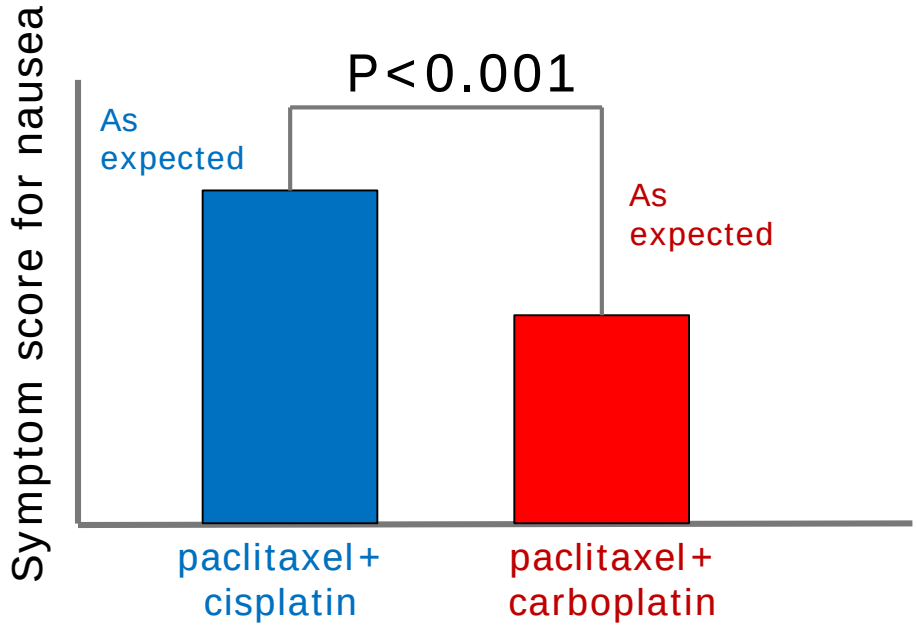
The answer of the guest is ‘very convenient **considering the price**’

Such phenomenon is called ‘information bias’

Prior information inevitably affects how a person feels.
In this scenario, ‘price’ is source of information bias.
Satisfaction rating by potential user must be wrong endpoint to answer to question ‘which is the more convenient car?’.

Rightly measured? – information bias

- Imagine… To decide which the better treatment regimen is
  - paclitaxel + cisplatin: ‘will cause strong nausea’ in IC form
  - paclitaxel + carboplatin: ‘will cause little nausea’ in IC form



Mean symptom score for nausea was significantly better in carboplatin

Therefore, paclitaxel + carboplatin is less toxic better treatment.

Do you agree?

The answer of the patients should be affected by prior information.
Observed difference is NOT always expected to reflect true difference.

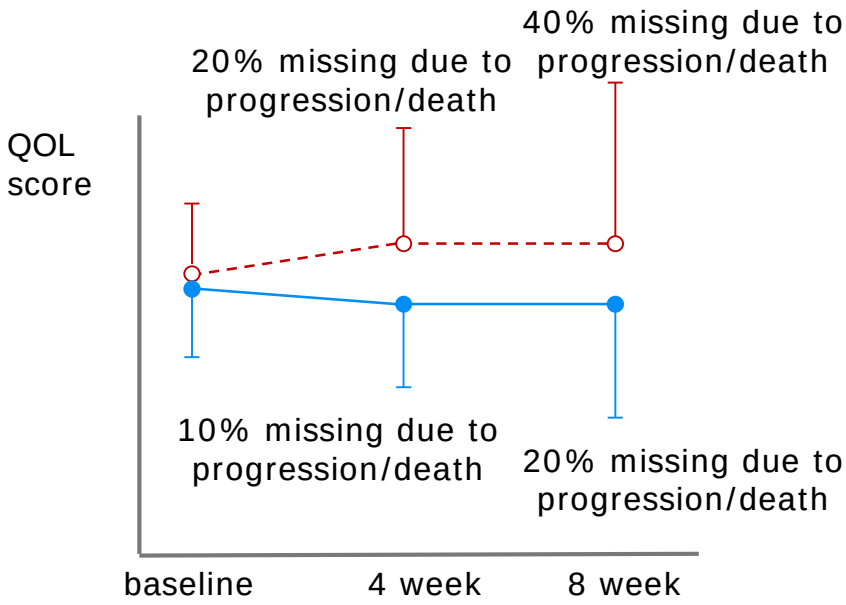
Prior information inevitably affects how a person feels.
In this scenario, ‘strong nausea’ is source of information bias.
Symptom score for nausea must be wrong endpoint to answer to question ‘which is more gentle treatment?’.

Remember that information bias is never eliminated by statistical analysis

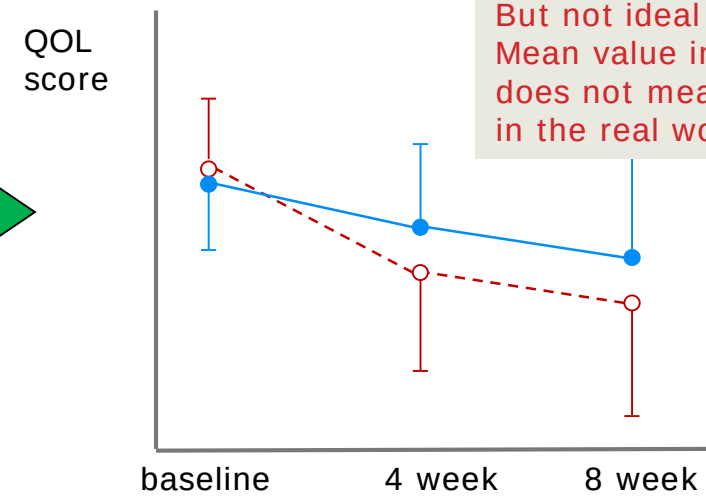
Rightly measured? – informative censoring

- Missing data is well-known source of bias in QOL

If true patients' QOL is worse in RED



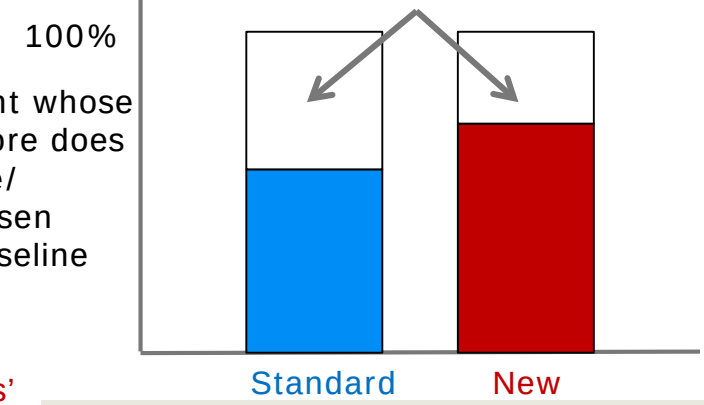
Worst value (zero) is imputed to missing



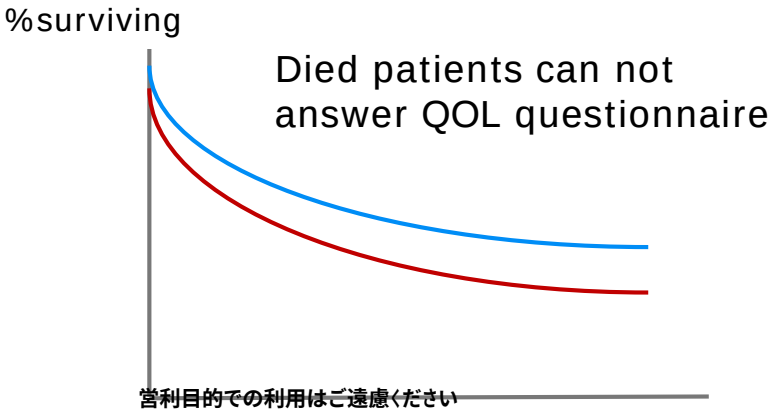
But not ideal way.
Mean value including zero does not mean something in the real world

ASCO recommends...

Patients who could not answer questionnaire are included here, not excluded from analysis.
So, this is unbiased method



"responder analysis"
by Dueck & Kunze
In 'Oncology Clinical Trials'



JCOG DC also recommends this method

Really matters? – physician's underestimation –

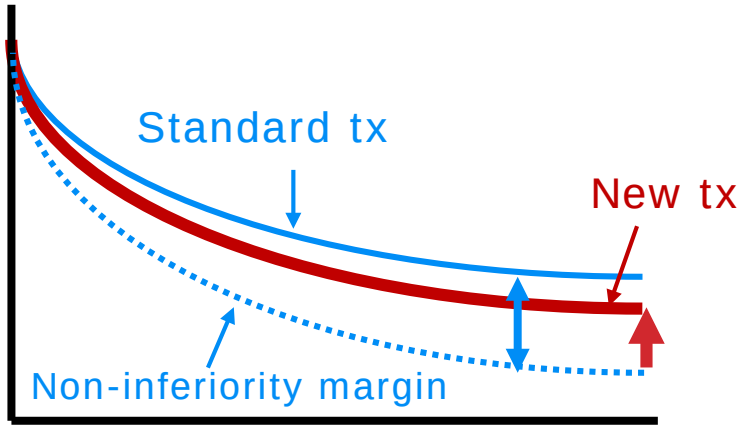
- There are many reports that physician's assessment tends to underestimate patients' true suffering
  - Yes, I agree.
  - Physicians tend to underestimate patient's hurt
  - Of course, it is not good thing

- But… It is expected that:
  - Physician who underestimates the hurt of patients treated with treatment A also tends to underestimate in patients treated with treatment B
  - Bias between the treatment arms may be smaller than patient-reported outcome with prior information
  - To underestimate equally between the treatment arms is less critical than to estimate differently between the arms

Critical things for choice of better treatment is biased assessment, not underestimation itself

Useful for choice of better treatment?

Non-inferiority trial

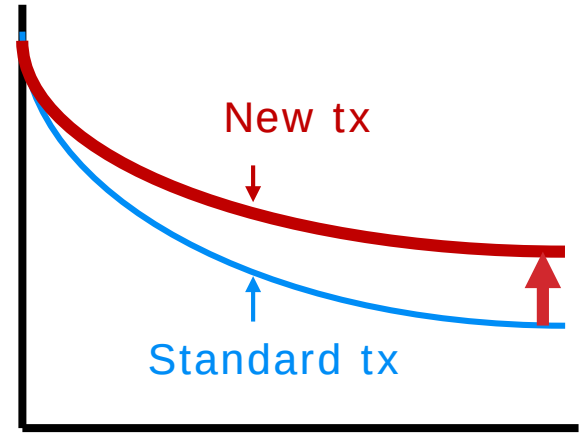


Survival non-inferior in New
Toxicity equivalent
QOL better in New

New is chosen

QOL results may influence treatment choice a little, but not indispensable

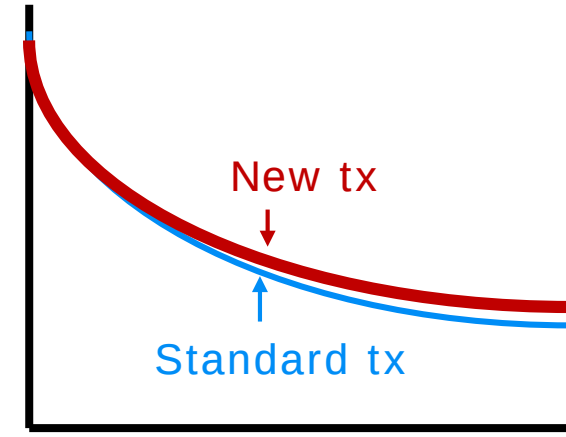
Superiority trial



Survival better in New
New more toxic
QOL worse in New

New is chosen

QOL results does not change the choice



Survival not differ
New more toxic
QOL better in New

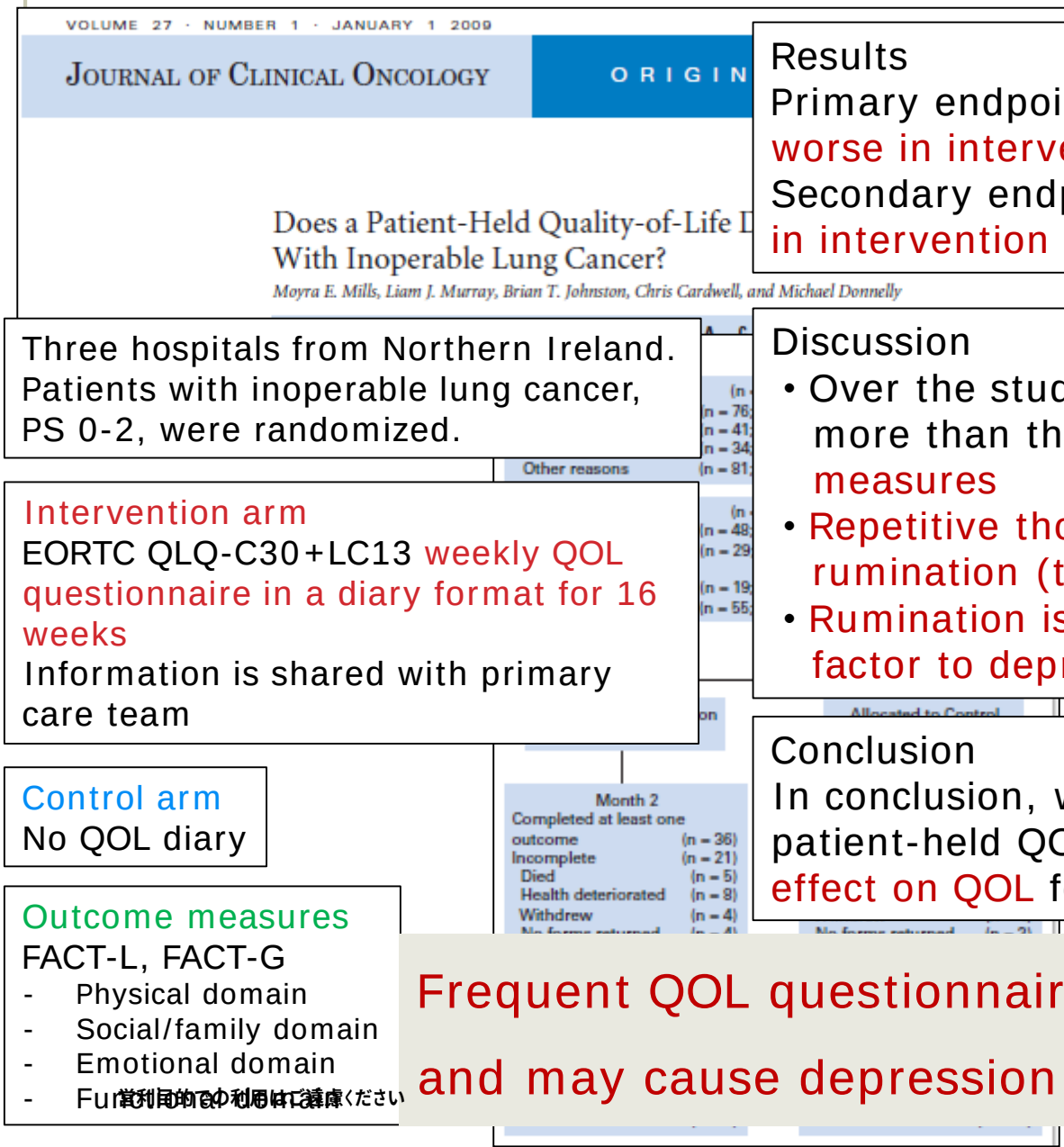
Standard is chosen

QOL results does not change the choice

QOL does not affect decision making in superiority trials

Respondent burden

Mills ME, et al.
J Clin Oncol 27:70-77. 2009


VOLUME 27 · NUMBER 1 · JANUARY 1 2009

JOURNAL OF CLINICAL ONCOLOGY ORIGIN

Does a Patient-Held Quality-of-Life D
With Inoperable Lung Cancer?

Moyra E. Mills, Liam J. Murray, Brian T. Johnston, Chris Cardwell, and Michael Donnelly

Other reasons (n = 81)

Month 2
Completed at least one outcome (n = 36)
Incomplete (n = 21)
Died (n = 5)
Health deteriorated (n = 8)
Withdrew (n = 4)

Allocated to Control

Three hospitals from Northern Ireland.
Patients with inoperable lung cancer, PS 0-2, were randomized.

Intervention arm
EORTC QLQ-C30+LC13 weekly QOL questionnaire in a diary format for 16 weeks
Information is shared with primary care team

Control arm
No QOL diary

Outcome measures
FACT-L, FACT-G
- Physical domain
- Social/family domain
- Emotional domain
- Functional domain

Results
Primary endpoint, TOI (Trial Outcome Index), was worse in intervention arm, but not significant (p=.07).
Secondary endpoint, FACT-L, were significantly worse in intervention arm (p=.03)

Discussion
- Over the study period, the diary group deteriorated more than the standard care group in all QOL measures
- Repetitive thoughts may lead to worry and rumination (thinking deeply)
- Rumination is strongly linked as a contributory factor to depression in cancer patients

Conclusion
In conclusion, weekly completion of a structured, patient-held QOL record may have a small negative effect on QOL for patients with inoperable lung cancer.

Frequent QOL questionnaires may deteriorate patient QOL, and may cause depression --- Caution!

Not only me, who is negative for QOL

- ASCO '*Value in Cancer Care Task Force*' declared NOT using QOL/PRO for risk-benefit assessment.

> - We did not find quality-of-life data or patient-reported outcomes to be end points reported in clinical trials with enough consistency or reliability to be informative in our assessment of clinical benefit.
> - Thus, we relied on a comparison of high-grade, acute toxicity, including rates of treatment-related death, to assess the negative physical effects of treatment that detract from overall health benefit.

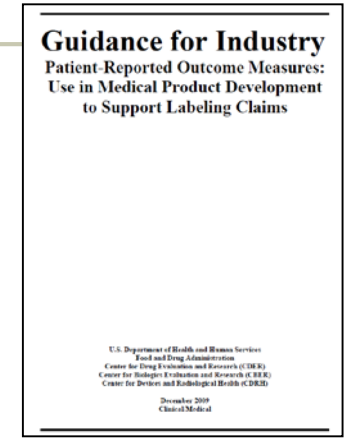
JOURNAL OF CLINICAL ONCOLOGY | ASCO SPECIAL ARTICLE

American Society of Clinical Oncology Statement: A Conceptual Framework to Assess the Value of Cancer Treatment Options

Lowell E. Schnipper, Nancy E. Davidson, Dana S. Wollins, Courtney Tyne, Douglas W. Blayney, Diane Blum, Adam P. Dicker, Patricia A. Ganz, J. Russell Hoverman, Robert Langdon, Gary H. Lyman, Neal J. Meropol, Therese Mulvey, Lee Newcomer, Jeffrey Peppercorn, Blase Polite, Derek Raghavan, Gregory Rossi, Leonard Saltz, Deborah Schrag, Thomas J. Smith, Peter P. Yu, Clifford A. Hudis, and Richard L. Schilsky

Schnipper LE, et al.
J Clin Oncol 33. 2015

How about FDA?


Guidance for Industry
Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims

- FDA Guidance for PRO 2009
  - Often misunderstood as “FDA recommends PRO”
    - No statement of “recommend PRO” in this guidance
  - PRO “can be used” for labeling claim if following this guidance – this is a certain ‘restriction’
    - Generally, findings measured by a well-defined and reliable PRO instrument in appropriately designed investigations can be used to support a claim in medical product labeling
  - FDA is reluctant for PRO used in open-label trials
    - Open-label clinical trials are rarely adequate to support labeling claims based on PRO instruments.

Remember that cancer trials are mostly open-labeled…!

Challenges in Assessing **Efficacy** with PRO Measures in Cancer Clinical Trials

- Many patients enrolled on cancer trials are asymptomatic with good performance status
  - Time to deterioration endpoints typically utilized
  - Enriching for symptomatic patients to measure symptom improvement/palliation should also be considered

- Trials supporting regulatory approval more often single arm or open-label in contemporary drug development
  - Degree of open-label bias is not well understood *I fully agree*
  - Research is needed to characterize the magnitude of potential overestimation of treatment benefit

FDA is cautious about information bias in open-label trials

Safety / Tolerability =
PRO Measurement Opportunity

- Symptomatic adverse events are best assessed by patients
- Safety and Tolerability- important in all phases of development
- PRO measures can offer different but complementary data to current clinician reported safety data
- PRO measures can be systematically and longitudinally obtained including a baseline measure

Complement to physician-assessed data, not substitute

Summary

- Thoughtful incorporation of patients into cancer clinical trials and drug development is becoming a priority
- Assessing safety and tolerability with PRO measures can have utility across the drug development life cycle
- Lack of flexibility with existing static PRO tools is problematic when assessing symptomatic adverse events
- PRO-CTCAE is a promising tool developed specifically to assess symptomatic adverse events and can involve patients directly in the assessment of safety and tolerability

FDA does NOT accept existing PRO tools, accepts only PRO-CTCAE

My conclusion – aggressive version

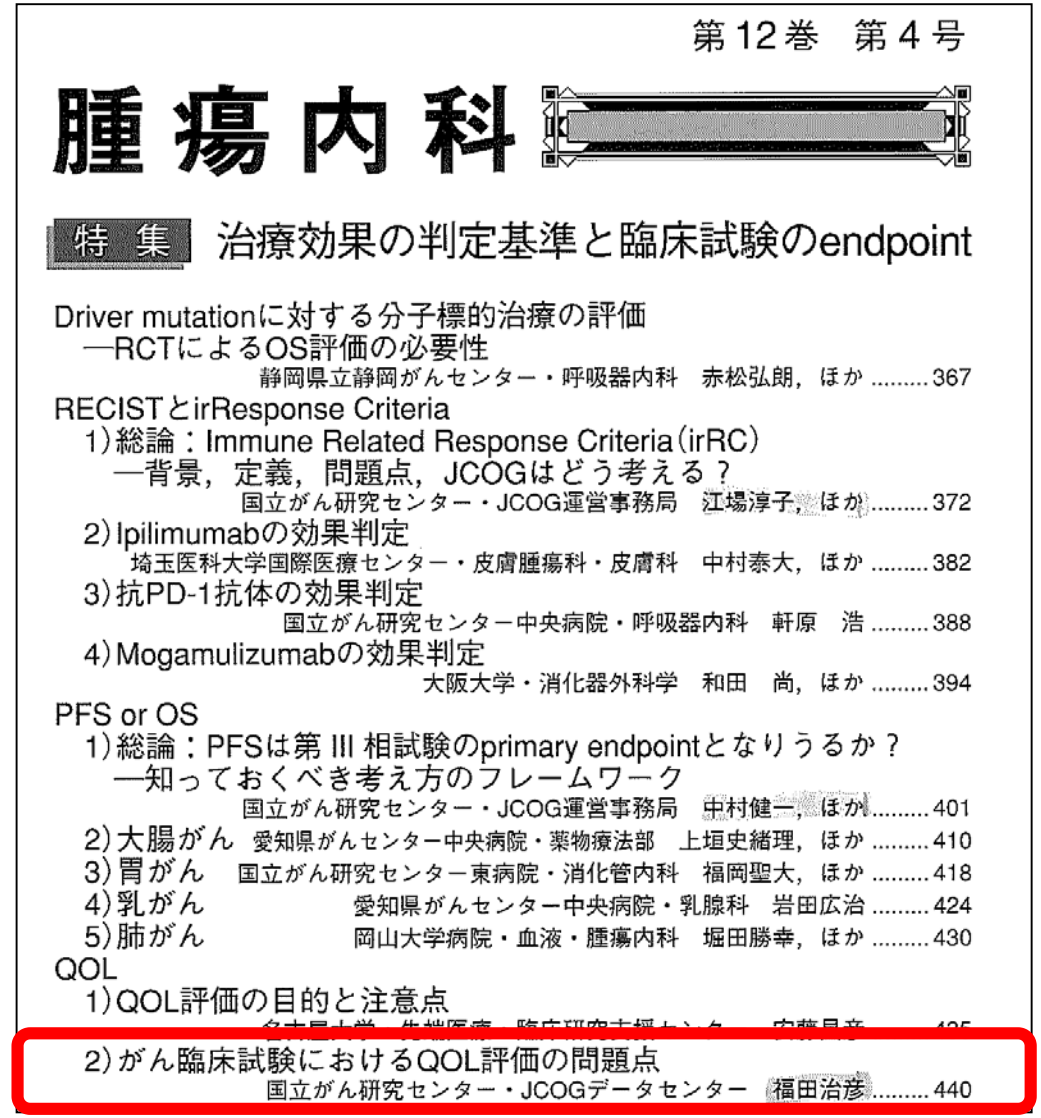
- HRQOL/PRO measures is NOT recommended to be used in open-label oncology trials
  - Patient burden and the risk of misinterpretation of trial results exceed the value of information gained from QOL questionnaires/PRO
  - Use in masked trials, eg. emetic drug, can be valuable and recommended
    - But, the more preferable is only one simple question to the patients; 'Which drug do you prefer?' after off-treatment in cross-over trial
- Only supplemental use of PRO-CTCAE for subjective toxicities is acceptable

I personally conclude that HRQOL/PRO is too complicated answer to simple question.

My conclusion – modest version

- Information bias should be carefully considered in open-label trials
  - Information given in the informed consent form and its possible impact on QOL scores should be considered in interpretation of trial results and drawing conclusion
- Be aware of the problem of missing data
  - Not too long term
  - Comparing %improved or %non-worsened is recommended according to ASCO guidance – responder analysis
- Respondent burden should be carefully considered
  - Minimize questionnaire sheets/items, only essential ones
  - Be conscious of risk causing depression in the patients
- Use in superiority trials is discouraged
  - Rarely useful in decision making for treatment choice in superiority trials

Further reading



If interested, please read my article in 'SYUYO NAIKA', 2013. Sorry, written in Japanese...

Thank you
for your kind attention