

“Important difference” for interpreting health-related quality of life outcome measures: important to whom?

Jackson S. Y. Wu

Received: 28 October 2010 / Accepted: 5 December 2011 / Published online: 21 December 2011
© Springer-Verlag 2011

In a hospital-based cancer clinic, a typical visit of an advanced cancer patient might start with this exchange: “Mrs. J, how are you doing?” “Well, I have been better...” “What’s bothering you the most these days?” “I feel nauseated all the time...I can’t do anything any more.” The clinician then evaluates the acuity and severity of the symptom, but also considers the whereabouts of the patient on her disease trajectory, whether her current combination therapy should be delayed, changed or stopped; whether opioid should be rotated, or steroid added, or new antiemetic, etc. Very quickly a complicated plan of management emerges, but one important thought the patient wishes to tell someone is lost in the shuffle, “I don’t like being so dependent on others; I’m a burden to my family...I have no energy for anything or anyone.”

This scenario demonstrates the multidimensional nature of health-related quality of life [1], for which the European Organisation for Research and Treatment of Cancer 30-item core quality of life questionnaire (EORTC QLQ-C30) [2] was developed to provide meaningful and interpretable patient-reported outcomes. The scale is an evaluative index [3] used within the context of clinical trials as aggregate measures of change in health status over time. Similar to other clinical measures, it follows that some cut point of minimal change has to be defined in order to identify a successful outcome (i.e., treatment response) that is relevant, important, and worth noticing (and worth paying) for *individual* patients.

But converting a scale measure into a dichotomized outcome poses considerable challenges, such as misclassification and the need to discern instrument-specific and context-specific variability. Of various methodological issues discussed in behavioral sciences and clinical epidemiology literature concerning cut points ([4], pp. 115–119), the reader should be cognizant of three issues when interpreting important changes in health-related quality of life measures:

1. *From whose perspective is the change deemed “important”?* Since one cannot assume “important” to carry the same meaning across societal, health policy, institutional, health care providers’, and patients’ perspectives, Schunnenman and Gyatt provided a definition of *minimal important difference* (MID) grounded on a patient-centered approach; that is, “the smallest difference in score in the outcome of interest that informed patients or informed proxies perceive as important, either beneficial or harmful...” [5]. They removed the focus on “clinical” (i.e., clinician-based) interpretations, thereby dropping the “C” from the MCID (minimal *clinical* important difference). Similarly in the recent IMMPACT recommendations, while acknowledging the relevance of clinical anchors, the authors emphasize the need for further research on patient-defined important difference ([6], p. 115).
2. *The criteria for defining an important change in individuals cannot be directly applied to the change between groups* ([6], p. 110). The cut point value for individuals who are deemed “improved,” based on some patient- or clinician-derived anchor, is not designed or expected to carry over as the minimal desirable difference between two comparison groups (e.g., treatment vs. placebo). Furthermore, the relevance of the anchor

J. S. Y. Wu (✉)
Department of Oncology, Tom Baker Cancer Centre,
Alberta Health Services Cancer Care, University of Calgary,
1331–29th St NW,
Calgary, AB T2N 4N2, Canada
e-mail: jackson.wu@albertahealthservices.ca

(s) should be considered with respect to the disease condition and natural history, and a given anchor may not be suitable for all quality of life subscales measured [7].

3. *Measurement of change reflects true change plus error* ([8], p. 1209). To evaluate change in individuals, the amount of change must be more than measurement error [9]. Essentially, the reliability (precision) of an instrument is data-driven (context specific) and will influence its interpretation. Although the standard error of measurement (SEM) is related to test reliability ([4], p. 117), the reliability change index (RCI, i.e., the change score divided by the SEM of the change) is suggested as a measure to describe the minimal threshold for meaningful change; for example, a minimally detectable change, as a function of the instrument for the population in question, may be defined as RCI of 1 (i.e., change score=SEM, or about 67% confidence level that the change score represents true change) [8]. More elaborate methods to compute RCI according to individual level and group level data are proposed by Hageman and Arrindell [10].

With this background in mind, we can turn our attention to the present study by Maringwa et al., who analyzed the QLQ-C30 data from two published non-small cell lung cancer studies with the aim to provide MIDs for various quality of life domains [11]. The first issue is the use of clinical anchors, including WHO performance status and weight changes, as criteria of important change. Unlike the study of breast and lung (small cell) cancer patients reported by Osoba et al. [12], the present study does not have patient ratings of “subjective significance” (subjective change in seven categories, from “very much worse” to “very much better”). Since performance status and weight changes are clinical parameters, the results in the present study are essentially MCIDs [5]. It is noteworthy that the correlations between changes in quality of life scores and changes in clinical anchors are no higher than 0.2 (Table 2), substantially lower than the 0.30–0.35 correlation recommended by Revicki et al. as a threshold of relevant correlation between anchors and patient-reported outcomes for establishing MID [7]. The MCID estimates in Table 5 are between-group differences (magnitude of mean change between improved and no change, between no change and deteriorated, as listed in the fourth and fifth data columns of Tables 3 and 4). These between-group differences represent the numerators of effect size estimates for responders and nonresponders.

Respecting limitations of available study data (the two randomized trials were probably not designed with an explicit objective of validating QLQ-C30 MID estimates), the MCIDs do provide a glimpse of the patterns of differences that

could arise across health-related quality of life domains. The magnitude of change *seems smaller* for within-group improvement than that for within-group worsening in some domains, and the magnitude of change *seems larger* for between-group improvement than that for between-group worsening for most domains. The magnitude of important change may indeed be different between improving and worsening status ([6], p.110), but without knowledge of measurement error in this dataset, one should refrain from reading too much into subtle differences.

With respect to measurement error, assuming the SEMs in Table 5 are derived from the between-group MCID estimates in the same table, most of the change estimates do not appear to surpass their respective SEMs (i.e., $RCI \leq 1$), limiting one's confidence in the results as representative of true change rather than error. However, unpublished data may yet support present study interpretation; alternatively, a more elaborate approach to group level data [10] may help confirm apparent between-group differences.

As for Mrs. J, the patient who presented for a routine visit, she decided to “take a break” from her chemotherapy. A month later, she reported better energy, less nausea, and better mood. She had altered her treatment course to meet her quality of life needs. Perhaps it is with this goal in mind that early palliative care for patients with metastatic non-small cell lung cancer provided survival improvement over standard oncology care [13]. Conceivably, if baseline quality of life scores could demonstrate independent prognostic value towards survival [14], treatment strategies that prioritize quality of life as a goal of therapy in advanced cancers might very well yield better overall outcome. And for quality of life, or status of “performance”, the better rater is more likely the patient who takes 10–20 min to fill out a questionnaire of 30 items of how he/she has been feeling and doing for the last week, rather than a clinician who has only a few minutes to ask if she is up and about for more than 50% of the day. The study by Maringwa et al. demonstrates some of the discriminative and evaluative properties of the EORTC QLQ-C30 measures of health-related quality of life, even though the conclusions are constrained by methodological limitations. Further accumulation of evidence towards patient-derived MIDs will help standardize health-related quality of life outcome measures and ease interpretation of results across cancer clinical trials.

Conflict of interest None declared.

References

1. Kassa S, Loge JH (2003) Quality of life in palliative care: principles and practice. *Pall Med* 17:11–20

2. Aaronson NK, Ahmedzai S, Bergman B et al (1993) The European Organisation for Research and Treatment of Cancer QLQ-C30: a quality-of-life instrument for use in international clinical trials in oncology. *J Natl Cancer Inst* 85:365–376
3. Kirshner B, Guyatt GH (1985) A methodological framework for assessing health indices. *J Chron Dis* 38(1):27–36
4. Streiner DL, Norman GR (2003) *Health measurement scales: a practice guide to their development and use*, 3rd edn. Oxford University Press, Oxford
5. Schunemann HJ, Guyatt GH (2005) Commentary—goodbye M(C) ID! Hello MID, where do you come from? *Health Serv Res* 40:593–596
6. Dworkin RH, Turk DC, Wyrwich W, Beaton D, Cleeland CS, Farrar JT et al (2008) Interpreting the clinical importance of treatment outcomes in chronic pain clinical trials: IMMPACT recommendations. *J Pain* 9(2):105–121
7. Revicki D, Hays RD, Cella D, Sloan J (2008) Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. *J Clin Epi* 61:102–109
8. Beaton DE, Bombardier C, Katz JN, Wright JG (2001) A taxonomy of responsiveness. *J Clin Epi* 54:1204–1217
9. Jacobson NS, Truax P (1991) Clinical significance: a statistical approach to defining meaningful change in psychotherapy research. *J Consult Clin Psychol* 59:12–19
10. Hageman WJJ, Arrindell WA (1999) Establishing clinically significant change: increment of precision and the distinction between individual and group level of analysis. *Behav Res Ther* 37:1169–1193
11. Maringwa JT, Quinten C, King M, Ringash J, Osoba D, Coens C et al (2011) Minimal important differences for interpreting health-related quality of life scores from the EORTC QLQ-C30 in lung cancer patients participating in randomized controlled trials. *Support Care Cancer* 19(11):1753–1760
12. Osoba D, Rodrigues G, Myles J, Zee B, Pater J (1998) Interpreting the significance of changes in health-related quality-of-life scores. *J Clin Oncol* 16(1):139–144
13. Temel JS, Greer JA, Muzikansky A, Gallagher ER, Admane S, Jackson VA et al (2010) Early palliative care for patients with metastatic non-small cell lung cancer. *N Engl J Med* 363:733–742
14. Quinten C, Coens C, Mauer M, Comte S, Sprangers MAG, Cleeland C et al (2009) Baseline quality of life as a prognostic indicator of survival: a meta-analysis of individual patient data from EORTC clinical trials. *Lancet Oncol* 10:865–871