Review

Quality of life assessment and reporting in randomized controlled trials: a study of literature published from Japan

Mariko Naito*¹, Takeo Nakayama¹ and Shunichi Fukuhara²

Address: ¹Department of Health Informatics, Kyoto University School of Public Health, Konoe-cho, Yoshida, Sakyo-ku, Kyoto, Kyoto 606-8501, Japan and ²Department of Epidemiology and Health Care Research, Kyoto University School of Public Health, Konoe-cho, Yoshida, Sakyo-ku, Kyoto, Kyoto 606-8501, Japan

Email: Mariko Naito* - naito@pbh.med.kyoto-u.ac.jp; Takeo Nakayama - nakayama@pbh.med.kyoto-u.ac.jp; Shunichi Fukuhara - fukuhara@pbh.med.kyoto-u.ac.jp

* Corresponding author

Abstract

Background: Standardization of quality of life (QOL) assessment and reporting in clinical trials is an imperative issue. While English-speaking countries have led this movement in standardization, there persists to be a limited amount of information from non-English-speaking including Japan. In this study, we bibliographically analyze the reporting of randomized controlled trials (RCT) conducted in Japan that used a QOL instrument.


Results: Percentages of QOL reporting in RCTs have increased between 1970–2003 both worldwide (0% for 1970–1974 to 4.4% for 2000–2003) and in Japan (0% to 1.8% for the identical periods). We found and evaluated 46 RCT reports published from Japan (32 in English, 14 in Japanese). The most commonly studied clinical condition was cancer (26, 56.5%) and the most common intervention was drug therapy (29, 63.0%). QOL was used as the primary endpoint in 10 studies (21.7%). Authors used established QOL instruments in 12 studies (26.1%), developed original instruments in 8 studies (17.5%) and assessed the symptoms or performance status in 10 studies (21.7%). Authors conceptually defined QOL in only 6 studies (13.0%). Neither response rate nor number of respondents for questionnaire surveys was specified in 16 studies (34.8%); furthermore, 11 studies (23.9%) did not describe respondents’ attributes.

Conclusions: Findings on relative frequency suggested that Japanese authors of RCT reports have less interest in QOL instruments than other international researchers in Western Europe and North America. Examination of RCT reports published from Japan revealed that there were several points to be improved in reporting QOL instruments. This study highlights the need to define QOL measures specific to clinical specialty and to examine methodology for assessing and reporting QOL.

Background

During the past two decades, ideas of health have transformed from ones that focus on illness to ones that consider patient well-being and quality of life (QOL) [1,2]. Currently, QOL instruments that integrate clinical and economic indices serve as a key to understanding...
treatment outcomes [2,3]. As Spilker defines, "Quality of life" in clinical medicine represents the functional effect of an illness and its consequent therapy upon a patient as perceived by the patient [4]. Although randomized controlled trials (RCT) generally have the objective of evaluating intervention effect (i.e., drug, modes of care), significant results are not recognized unless a suitable measure of outcome is evaluated [5-7]. Accordingly, the use of QOL assessment as an outcome measure in RCTs continues to gain attention.

While the CONSORT statement [8] establishes methods of RCT reporting, this guideline does not address problems with QOL instruments, standards of cultural validity and psychometric validity. Sanders et al. [9] studied QOL assessment in clinical trials, indicating a high frequency of incomplete responses and a poor quality of reporting. Findings described a need for specified standards of QOL assessment and reporting. In 2002, international standards for QOL assessment and reporting in clinical trials were proposed [10]. But the fact that QOL is culturally influenced [4] reminds us that standardization based on information having originated from English-speaking countries may be impetuous. Accordingly, an examination of QOL-assessment methodology in RCTs collected from nations worldwide could provide valuable insight into the ramifications of standardization.

A PubMed-based bibliographical search found approximately 4000 reports on clinical trials published in Japanese between 1987 and 2001; this ranks second to German-published reports in number of non-English articles [11]. When limited to RCTs, the number of articles published between 1995 and 1999 from Japanese institutions ranks top amongst Asian nations and within the top ten internationally [12]. An understanding of how QOL is assessed in Japanese clinical trials, particularly in RCTs, will contribute to the ongoing international discussion.

In this literature study, we calculate the number of RCT reports that refer to QOL and compare Japanese publication trends with international ones. We then examine all reports published from Japan as of 1970. Based on our findings, we point to several imperative problems concerning the assessment and reporting of QOL that require further examination.

Methods
We performed a literature search of reports published between 1970–2003 using PubMed. We first located articles reporting RCTs by using "randomized controlled trial" as a PubMed "publication type s [pt]" tag – this was also used as a denominator to gain relative frequency. Second, we identified RCTs that referred to QOL by using the free text term "quality of life". For each year, we calculated the proportion of reports mentioning QOL among all RCTs.

Third, we looked for RCTs published from Japan that referred to QOL as follows: "randomized controlled trial" [pt] AND "Quality of life" [text] AND (Japan [address] OR Japanese [language]). Results were not only examined by frequency of reporting, but were also assessed by hand in more detail. Reviews, editorials, meeting abstracts, letters and publications without abstracts were excluded from the latter analysis.

Items of evaluation were based on previous reports [9,13,14] and included: (1) type of subjects' condition; (2) type of intervention tested; (3) type of QOL measures used and whether they were validated; (4) whether QOL was defined conceptually; (5) reasons why QOL measurement was introduced into the trial; (6) whether QOL was categorized as the primary or secondary endpoint and description of (7) the response rate and, lastly, (8) respondents of QOL measurements. Two of the authors independently evaluated all articles for eligibility and resolved disagreements by consensus. In cases when information on QOL-instrument validation was not mentioned in the article, we verified what type of validation study was conducted by checking references.

Results
Counting the number of articles with "publication type" and a free text term, frequency of referring to QOL was found to increase over time worldwide – from 0 (0%) between 1970–1974 to 1930 (4.4%) between 2000–2003. This trend was similar among reports published from Japan with 0 (0%) in 1970–1974 to 27 (1.8%) in 2000–2003 (Figure 1).
65 RCT reports referring to QOL, published from Japan during 1970–2003, were found. We excluded 19 of these for not being RCTs (n = 4), not having an abstract (n = 1), not reporting on QOL (n = 12) or for being a review article (n = 2). The remaining 46 reports (32 were written in English and 14 were written in Japanese) met our criteria and were evaluated (Table 1 [see additional file 1]). Table 2 and Table 3 show the distributions of conditions and interventions examined and the language of each article. The most commonly studied condition was oncology (15 English reports; 11 Japanese reports; 26 in total, 56.5%), followed by cardiovascular disease (5 English reports; 0 Japanese reports; 5 in total (10.9%)) and urologic disorders (4 English reports; 2 Japanese reports; 6 in total, 13.0%), with other diseases occurring in less than 10% of reports (Table 2). The largest number of interventions was with drugs (18 English reports; 11 Japanese reports; 29 in total, 63.0%) followed by modes of care (9 English reports; 3 Japanese reports; 12 in total (26.1%)); other interventions were studied in less than 5% of reports (Table 3).

QOL was the primary endpoint in 10 reports (7 English reports; 3 Japanese reports; 10 in total, 21.7%). Only one report used a generic health-related QOL instrument. Authors defined QOL in 15 reports (9 English reports; 6 Japanese reports; 15 in total (32.6%)) (Table 4). Of these, authors of 9 reports (5 English reports; 4 Japanese reports; 9 in total, 19.6%) described their own definitions of QOL. After 1999, there were no RCT reports in Japanese that had investigators' definition of QOL.

A rising trend was recognized for reports describing why they performed a QOL assessment (Table 5) – prior to 1993, 2/7 reports (1 English report; 1 Japanese report; 2 in total (28.6%)); between 1994–1998, 6/14 reports (2 English reports; 4 Japanese reports; 6 in total (42.9%));
between 1999–2003, 11/25 reports (11 English reports; 0 Japanese reports; 11 in total (44.0%)).

Table 6 shows the methods for assessing QOL in each respective report. Authors used established QOL instruments in 12 reports (11 English reports; 1 Japanese report; 12 in total (26.1%)) and modifications of established instruments in six reports (2 English report; 4 Japanese report; 6 in total (13.0%)). Seven reports (5 English reports; 2 Japanese reports; 7 in total (15.2%)) used non-established QOL instruments but quoted references that described their methodology. Eight reports (3 English reports; 5 Japanese reports; 8 in total (17.5%)) developed original instruments. Except for three QOL scales in which each was used in two reports, a different scale was used in all reports.

Of the remaining articles, 10 reports (9 English reports; 1 Japanese report; 10 in total (21.7%)) assessed only symptoms or performance status, and three reports (2 English report; 1 Japanese report; 3 in total (6.5%)) used methods that were unclear. 30 reports (21 English reports; 9 Japanese reports; 30 in total, 65.2%) specified response rate or number of respondents. Lastly, 35 reports described respondents (26 English reports; 9 Japanese reports; 35 in total (76.1%)).

**Discussion**

Health status, functional status, and QOL are three concepts often used interchangeably to refer to the same domain of "health." [15] Health-related QOL continues to be used as a measure of outcome in clinical trials [16]. In the present study, we performed a literature search using PubMed in aims of comparing QOL assessment frequencies between international and Japanese RCT reports. Our study, however, has the following limitations. A text-based search was unable to locate articles that do not contain the term, "quality of life," in the abstract, title or as a Medical Subject Headings (MeSH). Also, PubMed does not hold all reports of RCTs published in Japanese [17]. Nonetheless, given that there currently exists no Japanese database of RCT-related articles in Japanese, PubMed serves as the best descriptive database for our study.

While the number of RCT reports referring to QOL published from Japan drastically increased after 1982, this does not necessarily mean that a similar shift exists in relative frequency of total RCTs. The relative frequency of Japanese RCTs rose in parallel with a rising global trend during the 1980s. In the following decade, however, this trend began to diverge from the international one. Although Japanese interest in QOL measures rose in the 1990s [18,19], doubt persisted among clinicians and...
researchers concerning the validity of subjective measures using interviews and questionnaires [20]. At that time, moreover, little research had focused on the methodologies of QOL assessment. This study's findings, we surmise, reflect this history of QOL in Japan.

Among the 46 reports examined, a total of 10 studies used QOL as a primary endpoint. Previous systematic reviews have criticized researchers for not using a suitable health-related QOL instrument as an outcome measure when studying the effect of intervention [21-25]. Nevertheless, it is still unclear as to whether this reflects a lack an applicable scale or the misuse of available scales [26].

27 of the 46 reports did not describe their reasons for using QOL as an outcome measure. It remains unclear whether this derives from a lack of description concerning outcome measures or insufficient discussion on the need for QOL assessment. It has also been suggested that sponsors and researchers may measure health-related QOL in aims of raising the image of a clinical trial or of the drug under investigation [27]. This study was unable to address this issue because all reports lacked a description of sponsorship and/or conflict of interests. Further studies are needed to investigate and delineate the respective objectives and significance of QOL assessment within each field.

Gill et al. [13] conducted a systematic review of 75 reports published between 1987 and 1991 containing the term QOL in the title and reported that only 15% of reports defined QOL. Our findings show that 32.6% of reports define QOL. Editorial procedures and limited word count, however, may possibly be associated with why QOL measurement selection and its definition were not mentioned. 60.9% of reports used neither an established QOL instrument nor an instrument without modification. This lack of consistency indicates the need for studies to describe their definition of QOL – at least when using their own instrument. Moreover, 50% of studies that used an established instrument did not provide a definition of QOL. This was probably due to the assumption that QOL assessment is equivalent to an establishment scale. This point to the need for defining QOL and for better description as to why an established scale was selected.

In the present study, all 15 (32.6%) reports that defined QOL were published after 1993. Furthermore, only four of the 10 reports that used QOL as their primary endpoint actually provided a definition. A total of nine reports (19.6%) that measured symptoms and/or reported the assessment of performance status (PS) in terms of QOL did not clearly define QOL. Only one report published in 1993 used the measure of PS in defining QOL assessment. While the percentage of reports that assessed symptoms and PS in terms of QOL was 36% (5 out of 14 reports) before the mid-1990s, this percentage dropped to 15.6% (5 out of 32 reports) in later years. In light of this, we surmise that, since 1997, QOL instruments that use a symptom index have slowly been replaced with subjective QOL instruments. That is, QOL instruments that measure subjective QOL have gained acceptance and the concept of QOL has shifted from being objective to subjective.

Clinical researchers can use QOL instruments to measure treatment efficacy, yet questions remain as to how QOL instruments should be selected, used, and how findings should be interpreted [28]. We based our current review on the methods discussed in the reports by Gill et al [13], and the Scientific Advisory Committee of the Medical Outcomes Trust [14]. This review also took the following items into consideration when examining instrument selection: whether the instrument was disease-specific or comprehensive, whether or not established instruments were tested for validity, whether or not modification of an established instrument was tested for validity, whether or not instruments were tested for validity despite the citation of references, and whether or not instruments were original (developed by the investigators themselves).

### Table 6: Type of evaluated measure used in 46 randomized controlled trials reporting on QOL

<table>
<thead>
<tr>
<th>Measure</th>
<th>English Reports</th>
<th>Japanese Reports</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Established QOL instruments</td>
<td>11 (34.4)</td>
<td>1 (7.1)</td>
<td>12 (26.1)</td>
</tr>
<tr>
<td>Modifications of established QOL instruments</td>
<td>2 (6.3)</td>
<td>4 (28.6)</td>
<td>6 (13.0)</td>
</tr>
<tr>
<td>Non-established QOL instruments (references quoted)</td>
<td>5 (15.6)</td>
<td>2 (14.3)</td>
<td>7 (15.2)</td>
</tr>
<tr>
<td>Original QOL instrument (non-established)</td>
<td>3 (9.4)</td>
<td>5 (35.7)</td>
<td>8 (17.5)</td>
</tr>
<tr>
<td>Symptoms or performance status only</td>
<td>9 (28.1)</td>
<td>1 (7.1)</td>
<td>10 (21.7)</td>
</tr>
<tr>
<td>Unclear measure</td>
<td>2 (6.3)</td>
<td>1 (7.1)</td>
<td>3 (6.5)</td>
</tr>
<tr>
<td>Total</td>
<td>32 (100.0)</td>
<td>14 (100.0)</td>
<td>46 (100.0)</td>
</tr>
</tbody>
</table>
Most of reports of this review used different disease-specific scales, and this may be due to the wide range of diseases studied in such a limited sample of RCT reports.

A serious problem apparent in several studies is the use of original instruments that have not been tested for validity or reliability. Garrett [26] has cautioned against this "easy development of new instruments." Previous studies have also indicated the need to consider issues of cultural phenomenon, spirituality [29] and level of medical-service satisfaction in the context of QOL assessment. QOL, being a multifactorial concept, should not be assessed with only one generic instrument [28] or with only one disease-specific instrument, but rather with multiple instruments.

Sanders et al. [9] found that respondents are not clearly described in approximately 30% of RCT reports that use QOL. Our findings intimate the same. The reason may be a lack of interest in the issue of respondents’ criteria and/or a mere insufficiency in reporting procedures. Slevin et al. [30] suggests that QOL assessment may have more validity when conducted by the patient him/herself than when conducted by medical providers. We found 18.8% (6 out of 32 articles) of English reports and 35.7% (5 out of 14 articles) of Japanese reports did not describe respondents. Moher et al. [31] further exemplifies that RCT articles written in English provide a clearer description than similar reports written in a non-English language. Our findings show that 34.4% (11 out of 32 articles) of English written reports as opposed to 35.7% (5 out of 14 articles) of Japanese-written reports showed neither response rate nor number of respondents. These insufficiencies in description highlight a need for better reporting standards. Needless to say, a study's quality and validity are most easily assessed by information described in a report [32-34].

Conclusions
Japanese interest in QOL assessment used in RCTs may remain at low levels when compared internationally. Our findings suggest insufficient discussion on when to use QOL as an outcome measure, how to select an appropriate scale and how to accurately use that scale. Further discussion is needed from the standpoints of study objective and psychometrics on the use of QOL assessment as an outcome measure in RCTs and on assessment methodology.

Before doing so, however, it is essential to establish a clear definition of QOL within each field for each particular disease. Contemporary research concerning QOL assessment in RCTs has demonstrated the need for international minimal standards including, but not limited to, scale selection, minimal standards of psychometric validity, and agreed standards of cultural validity [10]. We are confident that by collecting data from nations worldwide, a basis for further discussion on an international framework for QOL will foster. By participating in this discussion, researchers will subsequently gain the tools necessary to improve their nation's domestic handling of QOL assessment.

Authors’ contributions
All authors were responsible for planning, conducting and reporting this work and approved the final manuscript.

Additional material

Additional File 1
Table 1 Study of 46 reports. The contents of 46 reports which were assessed in detail
Click here for file
http://www.biomedcentral.com/content/supplementary/1477-7525-2-31-S1.doc

References
11. Nakayama T, Yamasaki S: Percentages of reports of clinical trials, written in seven non-English languages, that have structured abstracts. General Medicine in press.
33. Guyatt GH, Sackett DL, Cook DJ, for the Evidence-Based Medicine Working Group: User’s guides to the medical literature, II. How to use an article about therapy and prevention, A: are the results of the study valid? JAMA 1993, 270:2598-601.