Investigation of Evidence Sources for Health-Related Quality of Life in Cost-Utility Analysis of Pharmaceuticals in Japan

Mie Kasai Azuma, MS1,2, Shunya Ikeda, MD, DMSc1,*
1Graduate School of Pharmacy, International University of Health and Welfare, Ohtawara, Japan; 2Eisai Co., Ltd. Government Relations, Tokyo, Japan

A B S T R A C T

Objective: To provide an overview of how current utility values were obtained in the cost-utility analysis (CUA) for pharmaceuticals in Japan on the basis of methodological guidelines developed in England and Wales, Australia, Canada, France, and Japan by conducting a systematic review of the published literature. Methods: We searched and reviewed CUAs conducted for pharmaceuticals in Japan, reporting the results as cost per quality-adjusted life-year (QALY). The databases we used were PubMed, EconLit, Centre for Reviews and Dissemination, and the Japan Medical Abstracts Society. The search terms were “QALY” and “Japan” or “cost utility” and “Japan” in the PubMed database, Centre for Reviews and Dissemination, and EconLit. In the search on the Japan Medical Abstracts Society database, we used the term “QALY.” Results: In total, 41 CUA articles met the selection criteria and the most common method of obtaining utility values was derived from the published literature (31 CUAs, 168 utility values). Five CUAs were elicited by directly asking the participants regarding their own health state, and four CUAs used “mapping” techniques in which utility values were linked to clinical results. The most commonly used instrument was the EuroQol five-dimensional questionnaire followed by the time-trade-off. A few CUAs mentioned how they selected the literature for the utility values, and some utility values were combined across different sources, using different methods, and obtained from different locations. Conclusions: Practical methodological guidelines need to be developed to provide standardized methods of presenting the procedure of using utility values from the literature. Although transferability of utility values across jurisdictions has not been discussed fully, this topic should be covered in methodological guidelines and recommend best practices for evaluations. Keywords: cost-effectiveness, cost-utility, health-related quality of life, QALY.

Introduction

The concept of cost-effectiveness has been used as a decision-making criterion for the allocation of scarce resources for health care in several countries, such as in the United Kingdom, Australia, and Canada. Cost-utility analysis (CUA) is one of the major methods used for determining cost-effectiveness, and it has been used in these countries for more than a decade. Last year, France composed its own set of cost-effectiveness guidelines for use in the pricing of pharmaceuticals. In April 2012, Japanese government’s consulting body called the Central Social Insurance Medical Council (Chu-i-kyo) set common ground for the discussion of the possible introduction of the cost-effectiveness concept for health technologies, including pharmaceuticals. In parallel, a methodological guideline was proposed by an academic group in early 2013 [1]. After 2 years of deliberations, Chu-i-kyo now focuses more attention on drugs as cost-effectiveness evaluation targets, either the newly listed ones or those with a certain period passed since listing. In addition, from April 2014, pharmaceutical companies are being requested to submit economic evaluations for designated products, as a trial.

To conduct CUAs for pharmaceuticals, a substantial amount of data is required. This data not only includes clinical data but also data that are not usually taken from clinical trials, such as long-time transition. One piece of such necessary data is the utility values used when calculating quality-adjusted life-years (QALYs). It was not common to take utility data in clinical trials in Japan; therefore, CUAs often used utility values that were derived from the literature, sometimes derived from studies conducted at different locations, with different instruments. The transferability of utility values, however, has not been fully discussed. Although some analysts have argued that utility scores can be used across jurisdictions [2], others insisted that differences in ethic and cultural backgrounds might affect perceptions of health [3,4]. Badia et al. [4] suggested that efforts should be made to obtain local health state indices, wherever possible, and further cross-national comparisons of preference values should be conducted.

Conflict of interest: The authors have indicated that they have no conflicts of interest with regard to the content of this article.

* Address correspondence to: Shunya Ikeda, Graduate School of Pharmacy, International University of Health and Welfare, Ohtawara, Japan; Eisai Co., Ltd. Government Relations, Tokyo, Japan.

E-mail: shunya@iuhw.ac.jp.

2212-1099/$36.00 – see front matter Copyright © 2014, International Society for Pharmacoeconomics and Outcomes Research (ISPOR). Published by Elsevier Inc.

http://dx.doi.org/10.1016/j.vhri.2014.04.011
Another issue surrounding the utility values is that different utility measurement methods or instruments yield different utility values. This could occur in direct methods, such as the standard gamble (SG), the time trade-off (TTO), and the rating scale (RS), as well as in indirect methods, such as the Euroqol Five-Dimension (EQ-5D) the Health Utilities Index (HUI), and so on. Among the direct methods, for example, it was reported that for the same patient, the scores derived from the SG were generally greater than their scores on the TTO [5–7]. A Japanese study by Noto et al. [8] reported that the elicited utility scores from students and medical staff under hypothetical stroke rehabilitation revealed greater scores on the RS than on the TTO. Similarly, with reference to indirect methods, Neumann et al. [9] reported that scores of patients with dementia on the HUI 2 were greater than their scores on the HUI 3, for all levels of severity. Furthermore, systematic reviews of 28 empirical studies that compared direct and indirect methods of estimating utilities revealed that direct methods (SG and TTO, in this case) resulted in higher scores than indirect methods (the EQ-5D and HUI 3, in this case) [10].

To review the current practices related to these two issues, we examine what was addressed in the national methodological guidelines for economic evaluation across different countries because these guidelines represent the best practices in economic evaluation. Thus, we examined the guidelines provided by the following organizations: the National Institute for Health and Care Excellence (NICE) in the United Kingdom (England and Wales) [11], the Pharmaceutical Benefits Advisory Committee in Australia [12], the Canadian Agency for Drugs and Technologies in Health in Canada [13], the Haute Autorité de Santé (HAS) in France [14], and the proposed guideline provided by an academic group in Japan [1] that was funded by the Japanese Ministry of Health and Welfare.

Methods to Derive Utility Values as Prescribed in the National Guidelines of Five Countries

Although most of the guidelines prioritized the utility values derived from the studies conducted in their own country, the guidelines from all five countries allowed for the elicitation of utility values from the literature subject to the fulfillment of certain conditions. For instance, in France, the utility values and the life-years were recommended to be extracted from French empirical data. If French data were not available, the use of foreign data was allowed if the methodological quality of the study was good (although “good” was not defined) and its external validity was justified. The NICE guideline requires the utility values to be measured directly from the patients using the EQ-5D. If such information is not available in data from relevant clinical trials, it recommends the use of EQ-5D data from the literature, provided it is identified using a systematic and transparent method and there is clear explanation of the justification for choosing a particular data set. The Australian guideline recognized that utility estimates might sometimes be derived from the literature. It specifies, however, that the details of the methods used to identify the studies and to elicit the utility estimate should be presented clearly, so that validity should be assessed. According to the Canadian guideline, if utility values (preference scores) were not measured in a prospective study, they could be sourced from the literature if they are appropriate for the population of interest. In the proposed Japanese guideline, it is recommended that utility values be, in principle, elicited from the public. If Japanese studies, however, are not of optimal quality, the use of study results from other country settings is allowed.

Situations in which more than one plausible set of utility data for a health state was available from the literature, the guidelines from the five countries differed on their recommendations. The NICE and HAS guidelines recommended that scores from a single source should be used and sensitivity analyses should be conducted to show the effect of the alternative utility values. The Australian guidelines provide a warning that combining utility weights across different sources, for different health states, makes it difficult to interpret the results, particularly across those using different methods. We also noted that this issue was not addressed by the Canadian and Japanese guidelines.

Measuring Utility Values as Prescribed in the National Guidelines of Five Countries

Recognizing the variation in the utility values depending on the instrument, the NICE guideline applied strict criteria by specifying that only the EQ-5D measurement scale should be used for utility measurements. We also found that the recommendations in the guidelines may be influenced by the availability of the instruments in the country. For instance, the French authority recommended the use of “validated preference-based scores available in France”; currently, only the EQ-5D and HUI 3 meet this criterion. The proposed Japanese guideline recommended the use of index-type instruments with a newly developed Japanese scoring algorithm, which is currently met by the EQ-5D. The Canadian guidelines are relatively flexible and encouraged using indirect methods, such as HUI, the EQ-5D, the Short Form 6 Dimension (derived from the MOS 36-Item Short-Form Health Survey), and the 15D because they are easy to obtain, compare, and interpret. The guidelines, however, recommended that analysts select in advance the most appropriate instrument for the condition, one that best suits the study question and that prevents bias. The Australian guidelines are realistic in this regard and do not specify any particular instrument because such instruments were not always used in all the trials. The guidelines generally preferred, however, the indirect methods of multiattribute utility instruments. Although the five guidelines varied in terms of the instruments used to measure the utility values, all guidelines recommended at least the EQ-5D.

Considering these issues surrounding utility values, the objective of this study was to provide an overview of how current utility values were obtained in CUAAs for pharmaceuticals in Japan on the basis of recommendations in these five methodological guidelines. If the values were elicited from the literature, we examined the original sources (original study or not/study location/instrument used), the criteria and methods used to choose those selected original sources, and if more than one plausible set of utility data was available, methods of dealing with or selecting the data. We accomplished these objectives by conducting a systematic review of the published literature on CUAAs for pharmaceuticals in Japan.

Methods

Literature Review

We searched the published literature on the CUAAs conducted in Japan that reported the results as cost per QALY. The following databases were used for this search: PubMed, Centre for Reviews and Dissemination (CRD) in the University of York, EconLit, and the Japan Medical Abstracts Society (Ichushi). We included all articles that fit the search criteria, irrespective of the language of publication. The search terms used for the PubMed, EconLit, and CRD databases were “QALY” and “Japan,” or “cost utility” and “Japan.” For the Ichushi database, we used “QALY” as the search term. We conducted the search on February 20, 2013, and the search period was not specified. We also searched reference lists of the included studies.
Inclusion/Exclusion Criteria

Because we were interested in the process of evaluating the cost-effectiveness of individual drugs, we only included those studies that targeted drugs. The comparators, the items compared with the target drug, however, could be any health care technologies, such as therapeutic and/or diagnostic techniques, and medical devices, as well as drugs.

Reviews, commentaries, or educational articles were excluded. We excluded one study that collected only utility values, without compiling the cost per QALY. With reference to CUAs that targeted drugs, we also excluded the studies that compared a pharmaceutical agent in different regulatory situations, because current discussions in the Chu-i-kyo focus on drugs and explore their cost-effectiveness assessment for using such information to determine either the reimbursement status or the list price of some pharmaceuticals. Therefore, comparisons such as prescription drugs versus those in the over-the-counter setting, or branded versus generic versions of the same agents, were excluded. Studies comparing the same agents across a number of different regimens were also excluded for the same reason. We also excluded pharmaceutical agents that are not covered by social health insurance in Japan, such as prophylactic agents for healthy people (e.g., vaccines), because these were irrelevant to the current interest.

Items of the Review

When the published CUA studies were extracted on the basis of inclusion and exclusion criteria, the following research questions were addressed:

1. How did the CUAs obtain their utility values (empirical study/literature/others)? If the values were elicited from the literature, how did the author deal with or select those data?

2. What kind of studies were the original sources (original study or not/study location/instrument used)?

3. How were those original sources selected?

4. If more than one plausible set of utility data was available, how did the author deal with or select those data?

Results

Overall, 124 articles were obtained from the PubMed database, and 90 articles were excluded on the basis of exclusion criteria. Similarly, 161 articles were obtained from the Ichushi database, and 141 articles were excluded according to the exclusion criteria. In addition, two Japanese articles, Tanno et al. [15,16], were excluded because these two articles were Japanese versions of their previous English article Tanno et al. [17]. Finally, after excluding 11 duplicate articles, 41 CUA articles [17-57] that met the selection criteria were chosen (Fig. 1). No additional articles were found in the EconLit and CRD databases.

Methods of Obtaining Utility Values in the 41 CUA Articles (From Empirical Studies/Literature/Others)

We found that the most common method of obtaining utility values in the 41 CUA articles [17–57] was deriving them from published reports (31 CUAs). The remaining five CUAs elicited them by directly asking participants to evaluate their own health state, four CUAs used a “mapping” technique in which utility values were linked to clinical results, and one CUA assigned arbitrary values (Table 1).

If the values were elicited from the literature, what kind of study were the original sources (original study not/study location/instrument used)?

Fig. 1 – Disposition of publication identified for this systematic literature review. CUA, cost-utility analysis.
of the type of quoted studies from which the utility values were including guidelines [34,36,49,50,52], other economic evaluations more, 54 utility values were derived from either review articles, from original studies that measured the utility values from authors deal with or select those data? in Japan.

Methods of Selecting and Dealing with Original Sources

With reference to the study locations, out of the 107 utility values derived from original studies, 49 utility values were from studies conducted in Japan, 51 utility values were from studies conducted in other countries, and 7 utility values were derived by combining results from studies conducted in Japan and other countries (Table 3).

Regarding the instruments measuring utility values, the most commonly used instrument was the EQ-5D, and 39 utility values, in the original studies, were measured using this tool. Furthermore, the TTO was used to measure 15 of the utility values derived in the sample. The remaining instruments used for measuring the utility values were HUI 3 (7 utility values), mapping (7), visual analogue scale (VAS) (4), HUI 2 (3), the Quality of Well-Being Scale (3), RS (3), SG (3), MOS 36-Item Short-Form Health Survey (converted to Short Form 6 Dimension) (2). Twenty-one utility values were derived by combining different sources in the original studies. As presented in Table 2, we were unable to evaluate how the utility values were measured for 54 utility values, because those values were not derived from the original studies. As mentioned before, these values were derived from sources such as review articles, other economic evaluations, or a combination of expert opinion and results from other studies. Including the seven utility values that were arbitrarily assigned numbers, all instruments measuring the utility values in the 31 CUAs are illustrated in Table 4.

Methods of Selecting and Dealing with Original Sources

Only a few CUAs explained their methods of selecting the studies for the utility values. Some studies quoted reasons such as the absence of such data in Japan.

If more than one set of utility data was available, how did the author deal with or select those data?

As illustrated in Table 4, six CUAs fit this situation. The 21 utility values in those studies were obtained by combining the utility weights across original studies from different sources. In two cases, two types of instruments were used in one original study (the EQ-5D and VAS [39], and TTO and SG [25]); however, the former article did not mention which results were used and why, while the latter article assigned approximate values by using both the results. In another case, the original study measured the health state using the EQ-5D, TTO, SG, and VAS. The author of the original study [58] determined the mean of the scores from the TTO and VAS (converted 0 to 1) as the most appropriate utility; thus, two CUAs [37,40] used the mean. In one other case, there was more than one data set of utility values for each utility value, but the ways of selection were not mentioned by the author [24,29].

Discussion

This study provided an overview of how utility values were being obtained in CUAs for pharmaceuticals in Japan, evaluated with reference to the national methodological guidelines from five countries. The guidelines that we used were NICE in the United Kingdom (England and Wales) [11], Pharmaceutical Benefits Advisory Committee in Australia [12], and Canadian Agency for Drugs and Technologies in Health in Canada [13], which have a long-term history in cost-effectiveness evaluations, and newly developed guidelines provided by HAS in France [14] and the proposed guideline provided by an academic group in Japan [1]. In Japan, economic evaluation has not yet been officially used for decision making of pricing or reimbursement policy; therefore, the published literature is not directly linked to the assessment. We have illustrated, however, the current Japanese health economic studies at a glance.

All five guidelines allowed for deriving utility values from the literature, subject to systematic and transparent methods of identifying data and/or ensuring the quality of the study. Among the 41 CUAs, however, very few studies met these conditions. Furthermore, most of the CUAs did not explain the source of the

<table>
<thead>
<tr>
<th>Type of utility source</th>
<th>Number of CUA articles</th>
<th>References</th>
<th>Number of utility values used in the articles</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elicited by directly asking participants</td>
<td>5</td>
<td>[19,27,42,55]</td>
<td>34</td>
</tr>
<tr>
<td>Mapped from clinical results</td>
<td>4</td>
<td>[17,26,32,46]</td>
<td>–</td>
</tr>
<tr>
<td>Assigned any arbitrary numbers</td>
<td>1</td>
<td>43</td>
<td>5</td>
</tr>
<tr>
<td>Total</td>
<td>41</td>
<td></td>
<td>207</td>
</tr>
</tbody>
</table>

CUA, cost-utility analysis.

* Duplicated.
utility values in six CUAs [24,25,29,37,39,40] had been derived by combining the utility weights across different original studies, and the methods used for the same varied across these studies. In addition, five CUAs [20,23,33,41,51] quoted utility values from other economic evaluations, which required additional effort to identify the original study that measured the utility values, which was sometimes difficult to accomplish because the original article was too old or the study may not have cited the original studies appropriately. Therefore, we suggest that guidelines need to be developed to recommend standardized methods of presenting and assessing the procedure of using utility values from the literature. Regarding review articles as the reference for utility values, we did not focus on them in the present study, because we classified all CUAs on the basis of whether the original study measured the utility values. We acknowledge, however, that it may be an ideal way to derive utility values because the process of identifying the data is usually explained. It cases in which it may not be easy to identify the quoted data in the articles, authors are recommended to provide more information when utility values are derived from review articles.

Finally, even if the details of each of the cited utility values were explained, and the processes to identify those values were systematic and transparent, transferability of the utility values remains a vital question that needs to be addressed. Can the utility values derived from different sources, with different instruments, be used in the same CUA? For instance, within one CUA [21], some utility values were derived from a Japanese study conducted by using the TTO, while other utility values were derived from the other study conducted on samples from other countries using the EQ-5D. As described in the Introduction, such ways of transferring utility values across jurisdictions have not been fully discussed, but this topic needs to be covered in the methodological guidelines and relevant best practices need to be provided for conducting such evaluations.

In the present study, we could not fully assess the methodologies of utility value measurements in CUAs in which the utility values were elicited by asking patients [19,27,42,45,55] and using mapping techniques [17,26,32,46]. In former CUAs, the most commonly used instruments were the EQ-5D [19,27,42], followed by the TTO [55], and a combination of the TTO and SG [45]. These studies used a sample that ranged from 6 patients to 763 patients [19,27,42,55] and used health care professionals as proxies [45]. In terms of using mapping techniques, the guidelines of the five countries are described as follows: France does not recommend the use of mapping techniques in case analysis in their current state of development. If data from the EQ-5D were not available, NICE guidelines allowed using mapping techniques to estimate the EQ-5D questionnaire scores from other values observed in clinical trials. In Canada, if health-related quality of life was not measured in a prospective study, it is recommended that “preference scores can be gathered retrospectively through a separate exercise and then mapped onto the outcomes of the efficacy-effectiveness trial.” Australian guidelines do not seem to encourage mapping because the procedures were not well established, and they warn that special attention is required when mapping is presented. Finally, in Japan, using mapping techniques to estimate utility values from patient-reported outcomes is accepted.

### Recommendations

Reflecting the preferences of the local population (in this case, the Japanese population) is ideal; however, in the real world, performing an empirical study to derive the utility values of the local area for every CUA would be a challenge. So how utility values should be extracted from the literature while ensuring the plausibility and validity of the studies? On the one hand, if utility values obtained from the literature do not significantly affect results in terms of cost-effectiveness, measuring utility as accurately as possible becomes less important, such that utility values can just represent health status “off-the-shelf.” If validated utility catalogs were prepared in appropriate disease areas, more researchers would be encouraged to perform CUAs. In addition, in some disease areas, changes in utility values may not have a significant effect on results; therefore, further research is needed to determine which disease areas are more affected by changes in utility values.

### Table 3 – Study locations of the original studies.

<table>
<thead>
<tr>
<th>Study locations</th>
<th>Number of utility values</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Japan</td>
<td>49</td>
<td>[18,21,24,25,29,37,40,47–49,54]</td>
</tr>
<tr>
<td>The other countries</td>
<td>51</td>
<td>[21,22,28,30,31,33–35,38,39,44,47,53,56,57]</td>
</tr>
<tr>
<td>Combining the utility data from several locations</td>
<td>7</td>
<td>[24,29]</td>
</tr>
<tr>
<td>Total</td>
<td>107</td>
<td></td>
</tr>
</tbody>
</table>

### Table 4 – Instruments to measure utility values in 31 CUAs.

<table>
<thead>
<tr>
<th>Type of utility values instrument</th>
<th>Number of utility values</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>EQ-5D</td>
<td>39</td>
<td>[18,21,22,24,28,29,31,37,40,49]</td>
</tr>
<tr>
<td>TTO</td>
<td>15</td>
<td>[21,22,33,44,48]</td>
</tr>
<tr>
<td>HUI3</td>
<td>7</td>
<td>[30,35,54]</td>
</tr>
<tr>
<td>Mapping</td>
<td>7</td>
<td>[47]</td>
</tr>
<tr>
<td>VAS</td>
<td>4</td>
<td>[48]</td>
</tr>
<tr>
<td>HUI2</td>
<td>3</td>
<td>[57]</td>
</tr>
<tr>
<td>QWB</td>
<td>3</td>
<td>[53]</td>
</tr>
<tr>
<td>RS</td>
<td>3</td>
<td>[47]</td>
</tr>
<tr>
<td>SG</td>
<td>3</td>
<td>[38,56]</td>
</tr>
<tr>
<td>SF-36 (SF-6D)</td>
<td>2</td>
<td>[24,29]</td>
</tr>
<tr>
<td>Combining utility weights across different sources</td>
<td>21</td>
<td>[24,25,29,37,39,40]</td>
</tr>
<tr>
<td>Unknown*</td>
<td>54</td>
<td>[20,23,33,34,36,41,49–52,56]</td>
</tr>
<tr>
<td>Assigned arbitrary numbers</td>
<td>7</td>
<td>[21,24,31]</td>
</tr>
</tbody>
</table>

Total 168

EQ-5D, EuroQol Five-Dimension; HUI, Health Utilities Index; QWB, Quality of Well-Being Scale; RS, rating scale; SF-6D, Short Form 6 Dimension (derived from SF-36); SF-36, MOS 36-Item Short-Form Health Survey; SG, standard gamble; TTO, time trade-off; VAS, visual analogue scale.

* Quoted from either review articles, other economic evaluations, or combination with expert opinion and results from other studies.
Conclusions
A systematic review of the published literature was conducted, and 41 CUAUs met the selection criteria. We found that more than 80% (168 out of 207) of the utility values were derived from published studies. Among the utility values derived from original sources, about half the values were obtained from studies conducted overseas and the EQ-5D was the most commonly used instrument for measuring utility in CUAUs. Only a few CUAUs mentioned how they selected a study for utility values, and 21 utility values were reportedly derived by combining different sources, using different methods, and obtained from different locations. The other utility values were elicited from economic evaluations, which required additional effort to identify the original study. Therefore, practical methodological guidelines need to be developed, which recommend standardized methods of presenting and assessing the procedure of using utility values from the literature. In addition, although the transferability of utility values across jurisdictions has not fully been discussed, this topic should be covered in methodological guidelines and relevant best practices should be recommended for conducting such evaluations.

Source of financial support: The first author Mie Azuma is an employee of a pharmaceutical company called Eisai Co., Ltd. However, no specific grant has been received from any funding agency in public, commercial, or not-for-profit sectors, as well as the affiliate for this study.

R E F E R E N C E S


