Introduction

Korea is one of many countries that use cost-effectiveness (CE) evidence for reimbursement and pricing decisions on pharmaceutical products. In 2007, the government-run National Health Insurance (NHI) system in Korea adopted a positive drug listing system to improve the rational and efficient use of drug therapy in the nation. Consequently, pharmaceutical companies are required to provide the NHI with CE evidence on newly approved drugs if they require listing in the NHI drug formulary. In 2006, the Health Insurance Review and Assessment Service (HIRA), a government agency charged with reimbursement decisions on medical services and pharmaceutical products, developed a CE-assessment guideline for drugs. The guideline was updated in 2011 to improve the scientific quality of CE evidence produced by drug companies and to minimize the gap between the ideal goal and the feasible reality of CE evidence generation.

However, concerns have been raised that even the updated HIRA guideline is subject to pitfalls in the conduct of valid CE analyses. One of the concerns focuses on the inclusion of appropriate comparators. Since CE is a relative concept in which the incremental costs of a new drug versus comparator treatment are compared with the incremental effects, the choice of a comparator heavily influences the incremental CE ratio. It has been argued that the selection criteria of comparator treatments suggested by CE guidelines are not well constructed to produce a valid CE assessment of drugs. The present study investigated types of potential problems in comparator selection in complying with the Korean CE guideline. Methods: We conducted focus group interviews (FGIs) with experts from drug companies charged with generating and/or submitting CE evidence of their products for reimbursement decisions. To investigate the frequency of the potential problems in the field, we conducted an online survey among those working in drug companies, using a questionnaire developed from the FGI results. Results: Among the six problems identified by the FGIs, the highest proportion of survey respondents experienced “difficulty in obtaining reliable market share data necessary to choose comparator treatment” (94%), followed by “drugs widely used for a long time selected as comparators because they were recognized as standard treatments” (88%) and “therapeutically nonequivalent drugs selected as comparators” (72%). Conclusion: The results of our investigation would contribute to improving the quality of CE guidelines and the true value assessment of pharmaceutical interventions.

Keywords: Cost-effectiveness analysis · Comparator · Focus group · Guideline · Qualitative research · Selection criteria.
ternatives among comparable drugs should be chosen as comparators. Here, comparable drugs refer to drugs that can be substituted for the submitted drug if the latter is included in the reimbursement system; being comparable does not necessarily imply an identical mechanism of action. Where there are no comparable drugs, other treatment methods such as surgical interventions can be selected as comparators. However, CE data of existing interventions are not usually available. Therefore, the guideline recommends an alternative approach: choose the most frequently used interventions as comparators.

However, there are arguments that the selection criteria of comparator drugs or treatments suggested by the HIRA guideline are not well constructed to produce a valid CE assessment. There is a concern that the use of inappropriate comparators could result in an incorrect value assessment of a new drug. This may lead to invalid reimbursement decisions. Similar issues have been addressed in other countries. For example, a report from the University College London School of Pharmacy has criticized the UK’s National Institute for Health and Care Excellence for using inappropriate comparators in appraisals of the CE of new drugs by using low-cost non-licensed medicines as comparators.

Although concerns about inappropriate comparator choice in current new-intervention CE assessment practices have been addressed in the literature and in ad hoc discussions, the different types of comparator-related problems have not been systematically investigated. We believe that identifying the practical gaps in comparator selection for CE assessment should be the first step to improve the CE guidelines and, consequently, the quality of CE evidence. Therefore, our study aimed to investigate the potential problems in appropriate comparator selection to conduct a valid CE analysis of drugs in Korea, from the perspective of drug industry which are the main stakeholders in preparing and submitting CE evidence to the government.

### Methods

#### Focus group interview

To identify potential problems in selecting appropriate comparators for CE assessment of new interventions following the HIRA guideline, we performed focus group interviews (FGIs) in September 2014 with health technology assessment (HTA) experts working in research-based drug companies in Korea. As recommended by the academic advisory committee of the Korea Research-based Pharmaceutical Industry Association (KRPIA), 10 experts were recruited for FGIs. Each of them worked in one of the 35 member companies of KRPIA and was charged with submitting CE evidence to HIRA for listing of new or repositioned drugs on the NHI formulary. The focus group methodology was selected for two reasons. First, FGIs have been reported as a useful approach to research topics involving exploration of un-researched areas, such as comparator selection for CE analysis. Second, the interactive nature and group dynamics of focus groups would facilitate discussion and induce participants to explore their ideas, beliefs, and values about the discussion topic.

Each participant was provided with the interview question in advance via e-mail and was asked to present his or her experience and opinions related to the question at a round-table discussion. The interview question was framed as follows: “Please share your experiences of difficulties or problems in selecting appropriate comparators in compliance with the HIRA guideline when performing CE analysis or when commissioned to research for the listing of your company’s products on the NHI reimbursement list.” To fully retrieve individual participants’ opinions, we limited the number of FGI participants to five. Participants were randomly divided into two groups and a separate FGI was carried out for each group. In each FGI session, we asked the same question and the interview lasted about two hours. A facilitator moderated the focus groups and two research assistants recorded the discussions. All focus group discussions were audio-recorded digitally and the participants’ comments were transcribed verbatim.

Data analysis was conducted using the following phases of thematic analysis proposed by Braun and Clarke: 1) data familiarization, 2) generation of initial codes, 3) searching for themes, 4) reviewing themes, and 5) defining and naming themes. Through the data familiarization process, in which two authors separately listened to recordings and read written field notes several times, each author identified potential themes. They discussed these themes and developed a preliminary codebook. The third author reviewed the preliminary codebook and provided feedback. The authors arrived at a consensus on the themes via in-person discussion, and then defined the chosen themes.

#### Survey

To investigate the frequency of the potential problems in the field, we conducted an online survey among those working in drug companies, using a questionnaire developed from the FGI results. In the questionnaire, we presented each of the potential problems identified by the FGIs and asked the respondents whether they had experienced it. In answering those questions, respondents were allowed to choose multiple responses. For each problem, we suggested a method to resolve the problem based on the FGI discussions and asked the respondents to specify their degree of agreement with the suggestion on a 5-point Likert-type scale, with a higher score implying stronger agreement. Two-stage pilot tests were carried out.

60 J Health Tech Assess 2017;5(1):59-65
to improve the wording and structure of the questionnaire, one with the FGI participants and the other with a random sample of five subjects from the target population for the survey. We defined the target population as all staff of the 35 member companies of KRPIA whose workplace duty is related to CE evidence generation and/or submission to HIRA. KRPIA contacted each of the 35 member companies to obtain a list of the target population for the survey. Consequently, 93 people were identified and we e-mailed them requesting their participation in the survey. A total of 50 people responded to the survey between November 11 and 26, 2014, yielding a response rate of 53.8%. We generated descriptive statistics such as mean, median, and frequency for each practical problem with comparator selection.

The study, including both FGI and an online-based survey, was approved by the Yonsei University Institutional Review Board (IRB No. 1040917-201408-SB-207-03).

Results

Results of focus group interviews

The qualitative analysis of FGIs identified six themes associated with selecting appropriate comparators to conduct valid economic evaluation of innovative drugs for reimbursement decisions. Although we asked participants of FGI to share their experience of difficulties or problems in selecting appropriate comparators in compliance with the HIRA guideline, some of the themes derived from FGI are more a matter of disagreement with the guideline itself. Potential problems in selecting appropriate comparators and disagreement with the guideline are not identical issues. However, we believe that they are related to each other and provide useful information to improve the HIRA guideline. Thus, we present all of the six themes derived from the FGIs as study results in the following section.

Drugs used widely for a long time selected as comparators

The HIRA guideline recommends choosing the most frequently used interventions as comparators. As a result of complying with such recommendation, drugs that had been in the market for a long time and consequently obtained the largest market share are often selected as comparators. For this case, many of the FGI participants argued that drugs widely used for a long time could not be a standard treatment or a comparator by default. If a comparator drug had been used for a long time with no next-generation drugs available to replace it, the comparator drug could possibly be obsolete in terms of technology. Moreover, the price of the comparator drug set in the past would be low compared with that of a new innovative drug. If a new innovative drug is compared with such an obsolete drug, the resulting ICER could be high because of the rock-bottom price of the old drug, no matter how superior the treatment effectiveness of the new drug is in comparison with the comparator drug. Therefore, the FGI participants argue that evaluating a new innovative drug by reference to obsolete drugs characterized by non-comparable technology and pricing systems is considered inappropriate and the HIRA guideline should provide reasonable solutions to resolve this problem.

Therapeutically nonequivalent drugs used as comparators

According to the HIRA guideline, we are supposed to select comparators from among alternative treatments having the same indication as the new drug of interest. However, the guideline does not mention whether comparators should have an equivalent therapeutic status as well as the same indication. A number of participants stated that only treatments with an equivalent therapeutic status are valid comparators. For example, one participant commented:

"Anticancer drugs are good examples of this case. Anticancer drugs are nowadays classified as first-, second-, or third-line therapy, each of which denotes a different therapeutic status. On the other hand, most anticancer drugs in the past were licensed simply as first-line or higher therapy because of the lack of robust clinical data or specific reimbursement criteria. New anticancer drugs developed in recent years are mostly targeted anti-agents and licensed as second-line therapy. Often, no pre-existing treatment is replaceable with the new drug since most of the old anticancer drugs were licensed as first-line therapy."

Best supportive care used as a comparator

According to the HIRA guideline, comparable drugs, referring to the drugs that can be substituted for the submitted drug if the latter is included in the reimbursement system, should be chosen as comparators. For cases where best supportive care (BSC) is the only treatment choice and newly approved drugs are intended to be used to treat patients receiving BSC, BSC becomes a comparator based on the HIRA guideline. BSC in clinical trials is often defined as a control condition and is provided to patients nearing the end of life for the purpose of palliative care. According to Van Cutsem et al., BSC in oncology is defined as "the best palliative care per investigator excluding antineoplastic agents." Jassem et al. also defined BSC as "treatment administered with the intent of maximizing quality of life without a specific antineoplastic regimen, which includes antibiotics, analgesics, antiemetics, pleurodesis, blood transfusions, nutritional support, and focal ex-beam radiation for control of pain, cough, dyspnea, or hemoptysis."

Some of the FGI participants argued that it would not be appropriate if BSC were selected as a comparator when no

http://www.kahta.or.kr 61
treatment alternative with pharmacologic function is available. They argued that in such a case “no treatment” should be used as the comparator. One FGI participant referred to a clinician’s comment that “BSC is a sort of salvage and cannot be viewed as a treatment.” A condition in which no treatment choice other than BSC is available implies that the new drug of interest is the first innovation and is worthy of reimbursement.

Difficulty in selecting standard treatment in the face of too many treatment regimens

When various regimens are used for a single indication, it is not easy to choose a comparator regimen. A participant commented: “There are so many treatment choices, especially for cardiovascular diseases. For some cases, beta-blockers are considered the right treatment choice. For other cases, diuretics are recommended for patients. When various treatment regimens are used in clinical practice, it is not simple to choose one.” Another participant shared her experience: “So many regimens were used in clinical practice. My company chose a standard treatment among various regimens. However, the HIRA recommended a different one as the standard treatment. Both parties used different rationales to define their standard treatments.”

Difficulty in obtaining reliable market share data needed for comparator selection

According to the HIRA, an alternative treatment with the highest market share could be selected as a comparator. Thus, figuring out the market share of each alternative drug is a prerequisite to identify an appropriate comparator. Most of the participants believed that NHI claim records are perfect data for this purpose. However, they complained of limited access to the data. Participants also mentioned that Intercontinental Marketing Services (IMS) Health data are often used as an alternative source to figure out market shares of comparator drugs. However, IMS data are limited in that they provide only sales information, but not drug utilization information.

Generic product prices used for off-patent products

When off-patent products were chosen as comparators, the price of the new drug was compared with generic product prices rather than the original products’ initial prices set during the patent period. The FGI participants argued that the value of the comparator drug could be more accurately reflected using the initial price of the original product, set when the comparator drug was licensed and introduced in the market. This implies that the ICER of the new drug, computed using the price of generic products, would be the result of comparison of the new drug with undervalued comparators. The following are some sample comments on this issue:

“Once the patent expires and generic products are introduced, the price of off-patent products is lowered to half of the initial price set during the patent period. In this case, no matter how innovative a new drug is, the cost effectiveness of the new drug cannot be guaranteed if it is compared with half-priced off-patent products. The value of a new drug should be assessed on the basis of the comparator’s initial price set when it was first registered on the NHI reimbursement list.”

Results of survey

To examine the extent to which the potential problems identified by the FGIs occurs in generating CE evidence, we conducted a survey among all staff of the 35 member companies of KRPIA whose workplace duty was related to CE evidence generation and/or submission to HIRA. The mean age of the survey respondents was 36.9 years. The proportion of female respondents (58%) was higher than that of their male counterparts (42%). The respondents worked in drug companies for 10.3 years, on average. They worked for 6.6 years in the market access division of drug companies. They had a mean experience of being involved with 4.1 cases of generation and/or submission of CE evidence (Table 1).

Table 1. Demographic and work characteristics of survey respondents

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>n (%)</th>
<th>Mean (SD)</th>
<th>Min.–Max.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td></td>
<td>36.9 (6.08)</td>
<td>27–54</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>21 (42)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>29 (58)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Education</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>College graduates</td>
<td>17 (34)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Master’s degree</td>
<td>24 (48)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Doctoral study</td>
<td>9 (18 )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. of years worked in drug companies</td>
<td></td>
<td>10.3 (5.4)</td>
<td>2–27</td>
</tr>
<tr>
<td>No. of years associated with generating or submitting CE evidence in drug companies</td>
<td>6.6 (5.0)</td>
<td>1–27</td>
<td></td>
</tr>
<tr>
<td>No. of cases of involvement with generating or submitting CE evidence of drugs</td>
<td>4.1 (2.7)</td>
<td>1–10</td>
<td></td>
</tr>
</tbody>
</table>

CE: cost-effectiveness, SD: standard deviation
Among the six themes identified by the FGIs, the highest proportion of respondents experienced "difficulty in obtaining reliable market share data needed to determine comparator drugs" (94%) (Table 2). Other problems with relatively high frequencies of experience included "drugs used widely for a long time as comparators" (88%), and "therapeutically non-equivalent drugs as comparators" (72%).

A mean score was computed for the degree of agreement on each suggestion to resolve the comparator selection problems, with a score of 1 for "strongly disagree," 2 for "disagree," 3 for "unsure," 4 for "agree," and 5 for "strongly agree." All suggestions to resolve the problems had a mean scale score higher than 4, indicating close to strong agreement. The highest mean score was shown for the suggestion to "secure accessibility to reliable market share data for appropriate comparator selection" (4.83), followed by the suggestions to "select only therapeutically equivalent drugs as comparators" (4.62), "use the price of the patent drug for the off-patent product as comparator" (4.52), and "select standard drugs recently recognized as comparators rather than those that have been used widely for a long time" (4.50) (Table 3).

### Discussion

To the best of our knowledge, this is the first empirical study that explores the potential problems and concerns associated with the selection of valid comparators to conduct CE analysis of newly approved drugs from the perspective of drug companies. Considering the lack of preliminary knowledge about this issue, we carried out a qualitative study using FGI methods to retrieve experts’ experience and suggestions on this issue. The FGIs elicited six issues or disagreement with the CE guideline regarding comparator selection. The inappropriate comparators that the focus group pointed out reflect non-comparable alternatives in terms of therapeutic competency or pricing system. The following are examples of treatment alternatives with non-equivalent therapeutic status selected as comparators for a new drug of interest: a new anticancer drug com-

---

**Table 2.** Frequency of potential problems or disagreement with the guideline faced by survey respondents in selecting appropriate comparators for cost-effectiveness analysis of drugs

<table>
<thead>
<tr>
<th>Types of problems</th>
<th>Experienced, n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Drug widely used for a long time selected as comparator for new drug because it was recognized as a standard treatment</td>
<td>44 (88) 6 (12)</td>
</tr>
<tr>
<td>Drug selected as comparator because it had the same indication as, but not the equivalent therapeutic status of, the new drug (e.g., secondary anticancer drug compared with primary drug)</td>
<td>36 (72) 14 (28)</td>
</tr>
<tr>
<td>Best supportive care selected as comparator with no appropriate treatment alternatives available</td>
<td>25 (50) 25 (50)</td>
</tr>
<tr>
<td>Difficulty in selecting standard treatment in the face of too many treatment regimens</td>
<td>33 (66) 17 (44)</td>
</tr>
<tr>
<td>Difficulty in obtaining reliable market share data needed to select a comparator</td>
<td>47 (94) 3 (6)</td>
</tr>
<tr>
<td>Price of new drug compared with price of generic product for off-patent product selected as comparator</td>
<td>35 (70) 15 (30)</td>
</tr>
</tbody>
</table>

Multiple responses were allowed.

---

**Table 3.** Agreement on suggestions to improve appropriateness of comparator selection for cost-effectiveness analysis of drugs

<table>
<thead>
<tr>
<th>Suggestions</th>
<th>Mean scale (SD)</th>
<th>Strongly disagree</th>
<th>Disagree</th>
<th>Unsure</th>
<th>Agree</th>
<th>Strongly agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>Select standard drugs recently recognized as comparators rather than those used widely for a long time (n=48)</td>
<td>4.50 (0.74)</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>18</td>
<td>28</td>
</tr>
<tr>
<td>Select only therapeutically equivalent drugs as comparators (e.g., a primary anticancer drug should not be compared with a secondary drug) (n=47)</td>
<td>4.62 (0.82)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>9</td>
<td>35</td>
</tr>
<tr>
<td>If no alternative treatment, but only best supportive care is available, define comparator as “no treatment” (n=47)</td>
<td>4.23 (0.91)</td>
<td>0</td>
<td>3</td>
<td>6</td>
<td>15</td>
<td>23</td>
</tr>
<tr>
<td>Secure access to reliable market share data to select appropriate comparators (n=48)</td>
<td>4.83 (0.43)</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>For off-patent comparator, use initial price of product, set during patent period (n=48)</td>
<td>4.52 (0.97)</td>
<td>2</td>
<td>0</td>
<td>4</td>
<td>7</td>
<td>35</td>
</tr>
</tbody>
</table>

Each item was measured on a 5-point Likert-type scale, with a higher score reflecting stronger agreement (1: strongly disagree, 2: disagree, 3: unsure, 4: agree, 5: strongly agree). SD: standard deviation.
pared with BSC, which does not have antineoplastic action; a new innovative drug compared with an old drug that has been widely used for a long time partly because of a lack of innovation; a new second-line therapy compared with an old first-line therapy.

Ambiguity or inconsistency in comparator selection criteria also seems to be an issue. With too many regimens to treat the same condition, a consensus between the HIRA and drug companies on the reference regimen is difficult to work out. This problem causes confusion and wasteful conflict, which delay the reimbursement decision and therefore the market entry for newly approved drugs.

Concerns were also raised on comparisons between innovative and existing drugs with the extremely low level of drug prices due to factors other than value-based pricing, such as statutory price reduction of pharmaceuticals or price-volume agreements. Calculating ICER for comparator drugs based on generic instead of patent product prices is an example. Another example of a low-priced comparator drug is an old medicine marketed for 20 to 30 years. Maintaining the same drug price for a long time without adjusting for the inflation rate of the overall economy is a routine practice in Korea. Thus, drugs marketed for a long time often have extremely low prices. Strong pressure on cost savings may make cheaper products the most appropriate comparator. This approach leads to undervaluation of pharmaceutical innovation, which works as a barrier to inclusion in the reimbursement list, impeding public access to better treatment.

Although the issues collected from the FGIs are informative toward improving the acceptability of the CE guideline, we should be cautious in interpreting the study results, since this study has been performed only from a drug-company perspective. The potential problems in comparator selection derived from the FGIs are more industry-experienced problems in conducting economic evaluations. Some of the problems presented in this study seem to be about how the choice of comparator may impact the results of economic evaluations, namely unfavorable results. Thus, for a balanced discussion about the issues presented in this study, we recommend that a future study include stakeholders other than industry representatives, such as the HIRA, which makes CE guideline and appraises the CE evidence, those who use CE evidence for treatment decisions (i.e., clinicians and patients), as well as health economists and others dealing with HTA.

In addition to the lack of representativeness of the stakeholders investigated, this study has several limitations. First, compared with other FGI-based investigations, the present study used fewer focus groups, two versus four or more groups in other studies.11,17-18 The small number of groups suggests that our FGI findings may not stimulate the full range of discussions on the topic. However, the 10 participants from the two focus groups represent about 10% of the target population, which consist of 97 subjects. Thus, we believe that FGI discussions undertaken in this study achieve a relatively good representation of industry views. Second, the findings from our investigation, which is based on local issues, are limited to Korea and may not be applicable to other countries. Since each country has its own HTA guideline and appraisal system for CE evidence, based on a consensus among interest groups in that society, we do not suppose that all of the practical difficulties in complying with the HTA guideline of Korea pointed out in the FGIs of our study will apply to other countries. Lastly, considering the nature of FGIs, which are based on the open discussion group dynamic, it is possible that discussions are not solely based on an individual participant’s view, but are influenced by peers. In particular, the FGI participants in our study are from the same work field and have known each other for many years. Thus, pre-existing relationships and dynamics among participants may play a role in the dominance of specific issues. Moreover, some participants may have adjusted their statements either to conform to a popular viewpoint or because of concerns about offending others.

Conclusion

In conclusion, use of an inappropriate comparator could result in an incorrect value assessment of a new drug. Reimbursement decisions based on invalid assessments of clinical and economic value of pharmaceuticals negatively affect patients’ access to new drugs and medical innovation. Although our investigation reflects local issues in South Korea and a one-sided view from the drug industry, we believe that the issues and potential problems identified in our study could be used as a good starting point to specify potential areas for improvement of CE evidence by other countries using CE evidence for reimbursement decisions on newly approved pharmaceutical interventions.

Acknowledgments

Financial support for this study was provided in part by a grant from Korea Research-based Pharmaceutical Industry Association (grant number 2014-11-0814). The funding source had no involvement in study design, the collection, analysis and interpretation of data, in the writing of the report, and in the decision to submit the article for publication.

REFERENCES

1) Bae EY, Lee EK. Pharmacoeconomic guidelines and their implementation in the positive list system in South Korea. Value Health 2009;12 Suppl 3:S36-S41.