Nominal group technique to select attributes for discrete choice experiments: an example for drug treatment choice in osteoporosis

Mickael Hiligsmann, Caroline van Durme, Piet Geusens, Benedict GC Dellaert, Carmen D Dirksen, Trudy van der Weijden, Jean-Yves Reginster, Annelies Boonen

Department of Health Services Research, School for Public Health and Primary Care (CAPHRI), Maastricht University, The Netherlands, Department of Internal Medicine, CAPHRI, Maastricht University, The Netherlands, Department of Clinical Epidemiology and Medical Technology Assessment, CAPHRI, Maastricht University, The Netherlands, Department of General Practice, CAPHRI, Maastricht University, The Netherlands, Department of Business Economics, Erasmus Rotterdam University, The Netherlands, Department of Public Health, Epidemiology and Health Economics, University of Liege, Belgium

Background: Attribute selection represents an important step in the development of discrete-choice experiments (DCEs), but is often poorly reported. In some situations, the number of attributes identified may exceed what one may find possible to pilot in a DCE. Hence, there is a need to gain insight into methods to select attributes in order to construct the final list of attributes. This study aims to test the feasibility of using the nominal group technique (NGT) to select attributes for DCEs.

Methods: Patient group discussions (4–8 participants) were convened to prioritize a list of 12 potentially important attributes for osteoporosis drug therapy. The NGT consisted of three steps: an individual ranking of the 12 attributes by importance from 1 to 12, a group discussion on each of the attributes, including a group review of the aggregate score of the initial rankings, and a second ranking task of the same attributes.

Results: Twenty-six osteoporotic patients participated in five NGT sessions. Most (80%) of the patients changed their ranking after the discussion. However, the average initial and final ranking did not differ markedly. In the final ranking, the most important medication attributes were effectiveness, side effects, and frequency and mode of administration. Some (15%) of the patients did not correctly rank from 1 to 12, and the order of attributes did play a role in the ranking.

Conclusion: The NGT is feasible for selecting attributes for DCEs. Although in the context of this study, the NGT session had little impact on prioritizing attributes, this approach is rigorous, transparent, and improves the face validity of DCEs. Additional research in other contexts (different decisional problems or different diseases) is needed to determine the added value of the NGT session, to assess the optimal ranking/rating method with control of ordering effects, and to compare the attributes selected with the different approaches.

Keywords: discrete choice experiment, nominal group technique, patient preference, medication attributes, osteoporosis

Introduction

Over the last decade, discrete choice experiments (DCEs) have been increasingly used to elicit preferences for health care. Identification and selection of attributes are fundamentally important to obtain valuable results but are often poorly reported. Methods used to identify attributes include literature review, discussion with experts, professional recommendations, surveys, indepth interviews, focus groups, and repertory grid techniques. This first stage would generate a list of potential attributes for inclusion. In some situations, the number of attributes identified may exceed what one may find possible to pilot in a DCE. When the number of attributes may need to be restricted, the International Society for Pharmacoeconomics and Outcomes Research...
Good Research Practices for Conjoint Analysis Task Force has suggested that rating and/or ranking exercises may be beneficial in assessing the importance of attributes in order to construct the final list of attributes to be included. Some of the earlier techniques can again be used (eg, focus groups and interviews), but with a different objective from that in the identification stage.

The nominal group technique (NGT) seems particularly suitable for the “second” stage in which attributes are selected from the list by ranking them. The NGT is a structured, multistep, facilitated group meeting technique used to elicit and prioritize responses to a specific question. This method has been shown to be feasible and reliable for prioritizing health and health care research/problems but has never been investigated to select attributes for DCEs.

Therefore, this study was designed to assess use of the NGT to prioritize attributes for inclusion in DCEs. The study context is preferences for osteoporosis medications among adult patients. With the recent introduction of new therapies, conducting a DCE would be useful to understand patient preferences for these treatments, especially in view of the poor adherence with drug treatment in osteoporosis, but a DCE requires a rigorous and transparent approach to select attributes, given that many potential attributes have been identified in surveys and prior DCEs. While the results provide insight into preferences regarding the attributes of medications for osteoporosis, the primary objective of this paper was to test the feasibility and usefulness of the NGT to select attributes for DCEs. A secondary objective was to assess the influence of attribute ranking order on the selection of attributes.

Materials and methods

Patients

Five patient group discussions (consisting of 4–8 participants per NGT session) were conducted in June 2011 in The Netherlands and in Belgium to prioritize patient preferences for medication attributes. Patients were recruited during outpatient clinics or by telephone. Participants were considered eligible for inclusion in the study if they were diagnosed with osteoporosis or had a recent fracture that required osteoporosis medication during at least a period of their osteoporosis history. They were selected to represent the full clinical spectrum of age, educational level, history of osteoporosis (primary or secondary), and osteoporosis medication (switched, stopped, experienced side effects). The ethics committee of Maastricht University Medical Center approved the study and all patients received an information letter before participating and provided written consent.

Identification of candidate attributes

Fourteen potentially important attributes of osteoporosis drug therapy were established from a literature review and discussion with experts. Two attributes, ie, treatment duration and drug interactions, were not included in the final list because these attributes did not meet the important conditions of attributes for DCEs, such as being capable of being traded and being policy relevant. Indeed, treatment duration does not differ between first-line osteoporosis therapies, so is not relevant to policy, and drug interactions are very rare for all drugs used to treat osteoporosis. The final list of 12 attributes (Table 1) was approved by the working group, which included project investigators (rheumatologists, DCE experts), advisors, and a patient. A different ordering to present and discuss the attributes was used in each of the groups. Attributes were randomly divided into five sets (attributes 1–3, 4–6, 7–8, 9–10, and 10–11), and each of the five discussion groups received a different ordering of these sets.

Nominal group process

The NGT process consisted of three steps. After being informed about the purpose of the study (“to determine the most important characteristics for drug therapy in osteoporosis from the perspective of the patients”) and being given a brief description of the NGT process and its attributes, each participant was asked to rank the list of attributes by importance from 1 (the most important) to 12 (the least important) on a worksheet. Patients also had the opportunity to include any missing attribute. In comparison with a traditional NGT, and because many attributes were already identified in the literature, we did not include a first stage of silent generation of ideas whereby participants are asked to write down all ideas (attributes in this context) about a question.

During a second step (discussion and sharing ideas), a group discussion on each of the attributes was held, and

Table 1 List of potentially important attributes

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Efficacy (effect) in reducing the risk of future fractures (decreasing by between 20%–75% of the risk of future fractures)</td>
</tr>
<tr>
<td>2.</td>
<td>Side-effects (mild and common; serious and rare)</td>
</tr>
<tr>
<td>3.</td>
<td>Biological mechanism of action (bone resorption or bone formation)</td>
</tr>
<tr>
<td>4.</td>
<td>Frequency of administration (daily, weekly, monthly, yearly, etc)</td>
</tr>
<tr>
<td>5.</td>
<td>Mode of administration (oral tablet, subcutaneous, intravenous, etc)</td>
</tr>
<tr>
<td>6.</td>
<td>Place of administration (at home, doctor’s office, hospital, etc)</td>
</tr>
<tr>
<td>7.</td>
<td>Same drug during the treatment period (or sequential treatment)</td>
</tr>
<tr>
<td>8.</td>
<td>Mono therapy vs combination therapy (one or two pills)</td>
</tr>
<tr>
<td>9.</td>
<td>Out-of-pocket cost (personal contribution)</td>
</tr>
<tr>
<td>10.</td>
<td>Cost for the society (other health care costs than patient contribution)</td>
</tr>
<tr>
<td>11.</td>
<td>Time on market (recently vs 10 years)</td>
</tr>
<tr>
<td>12.</td>
<td>Branded or generic specification</td>
</tr>
</tbody>
</table>
included a group review of the aggregate score of the initial rankings. In the final phase, participants had the opportunity to reconsider their initial ranking in light of other participants’ views. They were under no pressure to achieve consensus, and all ratings were again made privately. The discussions were facilitated by a medical trainee in rheumatology, observed by a moderator, and tape-recorded. The facilitator tried to ensure that all participants were given an opportunity to contribute. NGT sessions were conducted until the rank order of the most important attributes did not change any further.

For each of the five groups, the individual rankings were summed across participants to derive the rank order at the group level. Some recoding was performed for a few patients who assigned the same number to different attributes. Any change between the initial and final round was examined to indicate the impact of the NGT on ranking. This analysis was carried out at the group level and at the individual level by examining the number of attributes changed by responders and the average of the (absolute) change between rankings of attributes.

Selection of final attributes
Selection of attributes for the DCE was based on group ranking and NGT discussions, followed by discussion among the experts, who decided on the number of attributes that should be included. The NGT sessions were particularly useful to determine the cutoff level beyond which inclusion as an attribute on the final ranking list should be stopped. The final list of attributes was further approved by the working group. No fixed threshold number was used to select attributes for inclusion, although recent reviews have reported that most DCEs used a number of attributes between 4 and 7.18,19

Results
Study sample
After five group discussions (two in The Netherlands and three in Belgium), the rank order of the attributes did not change any further, and no additional groups were invited. The final sample consisted of 26 patients with osteoporosis. As seen in Table 2, the patients represented the full clinical spectrum of age, educational level, osteoporosis diagnosis, fracture history, and treatment. Patient characteristics did not markedly differ between groups.

Final ranking of most important attributes
Figure 1 presents the five most important attributes in the different patient groups. Drug effectiveness was the most important medication attribute, followed by side effects, frequency of administration, and mode of administration, respectively. While out-of-pocket costs, time on market, place of administration (eg, hospital or home), and need for sequential treatment were of some relevance, costs to society, mode of action, combination treatment, and brand/generic specification did not reach the top three most important attributes in any of the groups.

Effect of NGT on rank order
Twenty of the 25 patients (80%) who provided an initial and final ranking changed their ranking after the discussion. However, the average initial and final ranking did not differ importantly, with two exceptions (Table 3). The importance of mode of action was reduced after discussion (from position 5 to position 8), while the out-of-pocket costs increased from position 10 to 5 because, in two Belgian groups, the importance of this attribute increased by 3 and 4 places, respectively, after the discussion. Mode of action was considered by most patients as a way of improving effectiveness and reducing fractures, although drug effectiveness is largely independent of the biological mechanism of drugs. This explanation was provided during the NGT discussions, explaining why this attribute was considered less important in the final ranking.

Individual patient analyses revealed different profiles of respondents (Table 4). Some patients (profile 1) did not change their ranking after discussion, some (profile 2) made minor changes to some parameters, and others (profiles 3–4) made more substantial changes in

Table 2 Participants’ characteristics

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Category</th>
<th>Number of Patients</th>
</tr>
</thead>
<tbody>
<tr>
<td>Women n(%)</td>
<td>24/26 (92%)</td>
<td></td>
</tr>
<tr>
<td>Belgian n(%)</td>
<td>17/26 (66%)</td>
<td></td>
</tr>
<tr>
<td>Age (years) Mean, median, standard deviation</td>
<td>68.0, 67.0, 11.0</td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>41–87</td>
<td></td>
</tr>
<tr>
<td>Diagnosis of osteoporosis Mean, median, standard deviation</td>
<td>25 (96%)</td>
<td></td>
</tr>
<tr>
<td>Osteoporosis since (years) Mean, median, standard deviation</td>
<td>10.2, 8.0, 8.7</td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>0–38</td>
<td></td>
</tr>
<tr>
<td>Education</td>
<td>No, primary or low secondary</td>
<td>9 (37%)</td>
</tr>
<tr>
<td>Secondary school</td>
<td>9 (37%)</td>
<td></td>
</tr>
<tr>
<td>Graduate/University</td>
<td>6 (25%)</td>
<td></td>
</tr>
<tr>
<td>With prior fracture</td>
<td>15 (58%)</td>
<td></td>
</tr>
<tr>
<td>Number of prior fractures Mean, median, standard deviation</td>
<td>1.04, 1.00, 1.22</td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>0–5</td>
<td></td>
</tr>
<tr>
<td>Patients on treatment</td>
<td>25 (96%)</td>
<td></td>
</tr>
<tr>
<td>Patients who took another treatment in the past</td>
<td>9 (35%)</td>
<td></td>
</tr>
<tr>
<td>Patients who experienced adverse events</td>
<td>4 (15%)</td>
<td></td>
</tr>
</tbody>
</table>
their ranking. After discussion, the average absolute change per patient between the ranking of the 12 attributes in the second ranking list compared with that in the first ranking list was 1.3 ± 0.8, meaning that on average each attribute moved (in absolute terms) by 1.3 places. The average number of attributes changed after discussion was 6.8 ± 3.1. As reported in Figure 2, the NGT discussion had the lowest impact on the attributes ranked as the three most important in the initial ranking, while the attributes ranked in the fifth and sixth positions were the most affected by discussion.

Additional findings

First, 15% of patients (4/26) did not correctly rank from 1 to 12 because they assigned the same number to different attributes. Second, the order of presentation of attributes in the rank system and nominal group discussion seemed to influence the ranking. When the attributes were presented in the first positions of the list, they generally obtained their highest score. Third, group analyses suggest that out-of-pocket cost was not in the top four in the two Dutch groups (fifth and eleventh position) reflecting the fact that, in contrast with Belgian patients (ranked as third, fifth and eighth), they have no out-of-pocket contribution for medications. No other major differences were observed between the groups. Finally, only one patient included a missing attribute, ie, drug interactions, which was discussed previously.

Selection of final attributes

Rankings and NGT discussions revealed four important attributes which were consistently identified as important for patients, ie, effectiveness, side effects, mode and frequency of administration. Interestingly, the finding that out-of-pocket cost was considered important in Belgium, but not in The Netherlands, suggests that this attribute could only be included in countries like Belgium where it is relevant. For most patients, time on market was related to safety and fewer side effects, which were already included as attributes. Place of administration is highly correlated with mode of administration and should rather be incorporated in the description of mode of administration. Other attributes were not sufficiently important for inclusion in the DCE based on ranking and discussion. Based on these considerations, we decided to include the four first attributes for the DCE in The Netherlands and the fifth attribute (out-of-pocket cost) in Belgium only.
Discussion

We have demonstrated the feasibility of the NGT in prioritizing attributes for inclusion in DCEs. When many candidate attributes are identified from available sources or patient interviews, this approach may be beneficial for assessing the importance of these attributes to construct the DCE. Therefore, in situations where the number of attributes identified needs to be restricted, a two-stage analysis could be performed, in which a self-exploratory analysis reduces the number of attributes (using NGT for example) and a DCE is conducted with the restricted list of attributes to assess further preferences for the levels of the attributes. Other tools (eg, best-worst scaling, adaptive conjoint analysis where attributes are changed simultaneously) could be alternative approaches.

Starting from a comprehensive list of attributes for osteoporosis medication generated from the literature and expert opinion, we identified which medication attributes are important from the patient perspective. Rankings and discussion revealed four important attributes, ie, effectiveness, side effects, and mode and frequency of administration.

These results are interesting for designing DCE experiments, and are also worthwhile in themselves when aiming to improve therapeutic adherence. Poor adherence with osteoporosis medication is a well-documented problem, which results in significant clinical and economic burden. Barriers to adherence include side effects, inconvenient dosing regimens, lack of information, and cost of medication. Providing patients with adequate information on the treatment options and involving them in decision-making may contribute to optimizing selection of treatment and improving adherence to therapy. Because drug therapies in osteoporosis differ in their side effect profiles as well as mode and frequency of administration, a two-stage analysis could be performed, in which a self-exploratory analysis reduces the number of attributes and a DCE is conducted with the restricted list of attributes to assess further preferences for the levels of the attributes. Other tools (eg, best-worst scaling, adaptive conjoint analysis where attributes are changed simultaneously) could be alternative approaches.

Table 4 Different profiles of responders after NGT discussion

<table>
<thead>
<tr>
<th>Profile 1 (0)</th>
<th>Number of patients</th>
<th>Average absolute change between attributes’ rankings: mean (standard deviation)</th>
<th>Number of attributes changed: mean (standard deviation)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Profile 2 (&gt;0–1)</td>
<td>6</td>
<td>0.85 (0.13)</td>
<td>6.7 (1.4)</td>
</tr>
<tr>
<td>Profile 3 (&gt;1–2)</td>
<td>9</td>
<td>1.76 (0.13)</td>
<td>8.5 (0.7)</td>
</tr>
<tr>
<td>Profile 4 (&gt;3)</td>
<td>5</td>
<td>2.58 (0.27)</td>
<td>10.8 (1.4)</td>
</tr>
</tbody>
</table>

Notes: Profiles of responders were determined based on the average absolute change between attributes’ rankings. Profiles’ classification is provided in parentheses in the first column; the average absolute change between attributes’ rankings was obtained by summing, for each attribute, the absolute change between initial and final ranking of attributes (a positive change (+1) or a negative change (−1) are treated the same (+1)) and dividing by the number of attributes.

Figure 2 Mean absolute change in ranking after NGT discussion according to their rank in the initial round.

Note: Graph shows that the attributes in the first three positions (that differ according to individual patient’s ranking) are the most stable after NGT discussions.

Abbreviation: NGT, nominal group technique.
frequency of administration, and these were considered to be important attributes in our research, sharing this information with patients could lead to optimized selection of treatment and improved adherence with therapy.

In the context of this study, the NGT discussions did not substantially affect rank order of preferences for the attributes in the total group when compared with rank order before the NGT discussion, indicating considerable agreement for the most important attributes. This suggests that a simple ranking exercise (or best-worst scaling) may perhaps be sufficient to determine the most important attributes. However, individual analyses have suggested that 80% of patients changed their ranking after discussion, and this could potentially be reflected in a different group ranking. Therefore, further investigations in other contexts, other diseases, and other decisional issues are needed to determine the added value of the NGT meeting when selecting and prioritizing attributes for DCE or even other purposes.

The approach described here also has the advantage of being rigorous, systematic, and transparent, and so may improve the face validity of DCEs. Many papers have pointed out that conjoint analysis did not justify the selection of attributes very well. Recently, Coast et al explored issues associated with attribute development for DCEs, and contrasted different qualitative approaches in the development of DCEs based on experience generated in interviews. Our study generated further insight by providing additional experience from group discussions. The benefits of conducting qualitative research were also not restricted to the selection of attributes. Discussions have been interesting in terms of refining language and for conducting a Bayesian efficient design. However, application of such a method did not come without a cost. We estimate that the whole process of organizing, running, and analyzing the NGT cost about €10,000 (including about €1500 as an incentive to patients for time spent and 2–3 months of the services of a full-time researcher). However, we believe that the benefits of the approach make it highly cost-effective.

The NGT could also be useful in selecting the initial set of attributes. Participants could first be asked individually to generate a list of important medication attributes, followed by discussion refining the list by adding, merging, or removing attributes, and by the final individual ranking of the most important attributes. This was not done in our study because many potential attributes were already identified by the literature review and we also aimed to assess the impact of the NGT session on rank order. However, our patients had the opportunity to add attributes to the list. Our study could also have some important implications for further research in this area. First, misunderstanding of attributes is frequent, and a clear description and explanation of the attributes is required. Second, ranking many attributes could impose a substantial cognitive burden on respondents. Perhaps it would have been sufficient to ask patients to rank their five most important attributes. Rating scales per attribute could also be an alternative requiring less effort on the respondent’s part, but with more limited information on the relative importance of attributes. Further work should be done to assess and compare ranking/rating exercises. Third, the impact of the NGT discussion was shown to differ substantially between patients. It would be interesting in the future to identify reasons that could explain this. Finally, our study showed that the presenting order of attributes did have an impact on the results. Therefore, we recommend controlling for ordering effects in ranking exercises.

A limitation of this study is that we have not compared the attributes derived from NGT with other approaches (eg, expert opinion, best-worst scaling). The gold standard would be the preference revealed, but this outcome is also difficult to assess. Head-to-head comparisons of different techniques could help to assess and understand differences between approaches, although there may be practical limitations in developing such studies.

Conclusion

A nominal group technique is feasible for selecting attributes for DCEs. Although, in the context of this study, the NGT discussions did not substantially affect the patients’ rank order of preferences for the attributes when compared with rank order before the group discussion, our approach is rigorous, transparent, and improves the face validity of future DCEs. Further work should be done to determine the added value of the NGT session, to assess the optimal ranking/rating method with control of ordering effects, and to compare the attributes selected using the different approaches.

Acknowledgments

The authors thank Marian Curfs and Rita Deroisy for helping us to contact patients, and also all patients for their participation.

Disclosure

Financial support for this study was provided by Amgen. The funding agreement ensured the authors’ independence in designing the study, interpreting the data, and writing and
publishing the report. The authors have no other conflicts of interest to disclose.

References
8. [No authors listed]. Nominal group technique is reliable for deciding research priorities. *BMJ*. 2000;320(7240):E.