Integrating patients' views into health technology assessment: Analytic hierarchy process (AHP) as a method to elicit patient preferences

Marion Danner
Institute for Quality and Efficiency in Health Care

J. Marjan Hummel
University of Twente

Fabian Volz
Institute for Quality and Efficiency in Health Care

Jeannette G. van Manen
University of Twente

Beate Wiegard
Institute for Quality and Efficiency in Health Care

Charalabos-Markos Dintsios
Institute for Quality and Efficiency in Health Care

Hilda Bastian
Institute for Quality and Efficiency in Health Care and *National Institutes of Health*

Andreas Gerber
Institute for Quality and Efficiency in Health Care

Maarten J. IJzerman
University of Twente

Background: Patient involvement is widely acknowledged to be a valuable component in health technology assessment (HTA) and healthcare decision making. However, quantitative approaches to ascertain patients' preferences for treatment endpoints are not yet established. The objective of this study is to introduce the analytic hierarchy process

(AHP) as a preference elicitation method in HTA. Based on a systematic literature review on the use of AHP in health care in 2009, the German Institute for Quality and Efficiency in Health Care (IQWiG) initiated an AHP study related to its HTA work in 2010.

Methods: The AHP study included two AHP workshops, one with twelve patients and one with seven healthcare professionals. In these workshops, both patients and professionals rated their preferences with respect to the importance of different endpoints of antidepressant treatment by a pairwise comparison of individual endpoints. These comparisons were performed and evaluated by the AHP method and relative weights were generated for each endpoint.

Results: The AHP study indicates that AHP is a well-structured technique whose cognitive demands were well handled by patients and professionals. The two groups rated some of the included endpoints of antidepressant treatment differently. For both groups, however, the same six of the eleven endpoints analyzed accounted for more than 80 percent of the total weight.

Conclusions: AHP can be used in HTA to give a quantitative dimension to patients' preferences for treatment endpoints. Preference elicitation could provide important information at various stages of HTA and challenge opinions on the importance of endpoints.

Keywords: Patient involvement, Patient preferences, Analytic hierarchy process (AHP), Health technology assessment (HTA), Medical decision making, Multi-criteria decision analysis (MCDA)

In the past few years, patient involvement and participation have been increasingly introduced in health technology assessment (HTA) and the resulting decision-making processes in various jurisdictions around the world. But why is it so important to incorporate patients' views into HTA, and where and how can this best be done? A survey by the International Network of Agencies for Health Technology Assessment (INAHTA) conducted in 2006 concluded that the involvement of "consumers" in the assessment and decision-making processes within HTA agencies "broadens the perspective of those assessments and of the advice provided to decision makers" (9). In the survey, the term "consumers" included "patients, carers, long-term users of services, organizations representing consumers' interests, and members of the public who are potential recipients of health promotion programs." According to Facey et al. (8), "patients can provide a real world understanding of the illness / condition, and the benefits and disbenefits of using particular technologies." Given that patients are the ultimate "end-users" of health technologies, patient involvement in HTA processes may therefore increase acceptance of HTA and the viability of resulting healthcare decisions.

Patients can be involved at various levels in healthcare decision making, that is, at the micro-, meso-, or macro-level, as experts on their own health. At the micro-level, patients can, for example, participate in shared decision making with their healthcare professionals (23). HTA agencies play a key role at the meso-level, often formulating recommendations as to which technologies should be implemented. Their recommendations may have a direct impact on healthcare policy and reimbursement decisions taken at the macro-level. At these levels, opportunities for patient participation exist and are being continuously explored (4;8;23;25).

There are at least three approaches to the inclusion of patients' views in HTA processes (8;18). First, patients or their representatives can be involved as members in committees and actively contribute to an HTA agency's work processes. Alternatively or additionally, in many countries patients or patient groups are invited to comment on specific issues. Second, certain qualitative methods (e.g., individual in-depth or focus group interviews, consensus panels) may enable researchers and decision makers in HTA agencies to assess and include patients' views in a more systematic and structured manner (e.g., patients as members in DELPHI panels to formulate guideline recommendations). Third, quantitative methods can be used to "measure" patient preferences. For example, the QALY approach, which is predominantly used in health economic evaluation within HTA, is based either on direct elicitation of patient preferences by Standard Gamble, Time Trade-Off and Visual Analogue Scales, or indirectly, by applying generic instruments such as the EQ-5D. Patient involvement in HTA agencies has been implemented to a differing degree in the first two approaches. In the INAHTA survey (9), twenty-one of the thirty-seven agencies involved consumers in several aspects of their HTA programs, for example, in the formulation of assessment topics or in their topic-prioritizing process. A quantification of patient-preferences in the form of utilities (QALY approach) or in monetary terms (willingness-to-pay [WTP] approaches) and integration into health economic evaluations, that is, cost-utility or cost-benefit analyses, is currently standard in many HTA agencies.

The purpose of this study is to introduce the analytic hierarchy process (AHP) as a means of quantifying patient preferences in a new area within the HTA process: to prioritize and weight patient-relevant endpoints for assessment in HTA reports to explore whether the method could be used in IQWiG's benefit- or cost-effectiveness-assessments. First, an AHP study related to two reports by the German Institute for Quality and Efficiency in Health Care (IQWiG) is presented and discussed. Second, further possible areas of application of AHP within HTA are outlined.

THE AHP METHOD: AN INTRODUCTION

Multi-criteria decision analysis (MCDA) is what basically happens in many everyday consumer decisions, in which the different characteristics of a product are taken into account and weighed against each other, for example, the design, color, and gasoline consumption of a new car. MCDA can be used to support decision making in health care and to elicit patient preferences and values in complex decision-making contexts (3;5–7;16;24). AHP is a MCDA technique developed by the mathematician Thomas Saaty in the 1970s (19;20). It was introduced to support strategic decisions in industry. AHP is an approach where a multi-attribute decision problem is first structured into a hierarchy of interrelated elements (17). This hierarchy is a tree-like structure that is used to decompose the decision problem, moving from main criteria to more specific sub- and "sub-sub"-criteria. Pairwise comparisons of these criteria are separately performed at each level of a decision hierarchy, moving through each group of criteria in the hierarchy from the lower-level to the upper-level criteria. Important methodological constraints within AHP regarding the decision hierarchy are the independence and comprehensiveness of criteria at each level (20). Based on matrices of the pairwise comparisons, Saaty's mathematical algorithm as a key element within AHP allows the calculation of an approximative vector ("right eigenvector") representing preference-based weights for each of the decision criteria included (7;20). While the preferences in AHP are recorded on a numbered but ordinal scale, calculation of preference weights is performed by transforming this scale into an approximative cardinal one. For details on the calculation of the eigenvector by means of different matrix multiplication methods, see in particular Dolan et al. 1989 (7).

Based on the weights calculated for each endpoint included for each person, a group geometric mean can be calculated for a group of individuals taking part in the AHP. In addition, because reciprocity and transitivity of preferences is required within AHP, AHP allows for calculation of a measure of consistency for each group of pairwise comparisons. This measure reflects how logical each pairwise comparison is with regard to the remainder of comparisons performed by the same individual. For example, an individual rating A > B, B > C, and C > A would be inconsistent in his or her judgments. This consistency ratio (CR), as a measure of performance within the AHP, has a threshold of 0.1 that should not be exceeded (7).

USE OF AHP IN HEALTH CARE

In 2009, IQWiG commissioned a systematic literature review to obtain information on how and where AHP was used in healthcare decision making. Eighty-five articles were identified of which fifty-five reported on specific applications of AHP in different decision-making contexts (10). Of these, 44 percent focused on management decisions in healthcare organizations, 25 percent on the development of clinical guidelines, 13 percent, respectively, on shared decision making and the development of national healthcare policy, and 5 percent on the development of healthcare innovations. The review supported the notion that AHP was a well-structured decision-making tool, which was easy-to-handle for different groups of individuals participating in the pairwise-comparison procedures. It should be noted in this context, however, that AHP has also repeatedly been subject to methodological criticism, for example, with regard to the preference scales used in AHP or to issues of rank reversal (e.g., (1;2)). Different approaches to the solution of these problems have been suggested or are being explored (1;15;21;22).

METHODS

Patients with a previous diagnosis of major depression, but currently in remission, were recruited by means of German self-help and patient interest groups contacted by the respective department at the German Federal Joint Committee. Patients were invited to participate in IQWiG's AHP workshops by means of patient group websites. Professionals involved in the care of patients with depression (psychiatrists, other specialized physicians, and psychotherapists working in private practice or hospitals) were identified and recruited by means of German scientific societies as well as Internet search engines.

Patients and professionals participating in the moderated AHP workshops were asked to compare pairs of treatment endpoints and score them using radio-controlled keypads. Participants had to rate the importance of one endpoint compared with another on a scale ranging from 1 (reflecting equal importance of the two endpoints) up to 9 (reflecting extremely greater importance of one endpoint over the other). One example of a pairwise comparison is displayed in Figure 1. Professionals in the AHP professional workshop were asked to rate endpoints according to which ones they considered more important in the care of their patients. The study focused on endpoints of antidepressant treatment that had previously been assessed in two IQWiG benefit assessments (12;13). Endpoint selection was based on predefined decision criteria, primarily the strength of the available
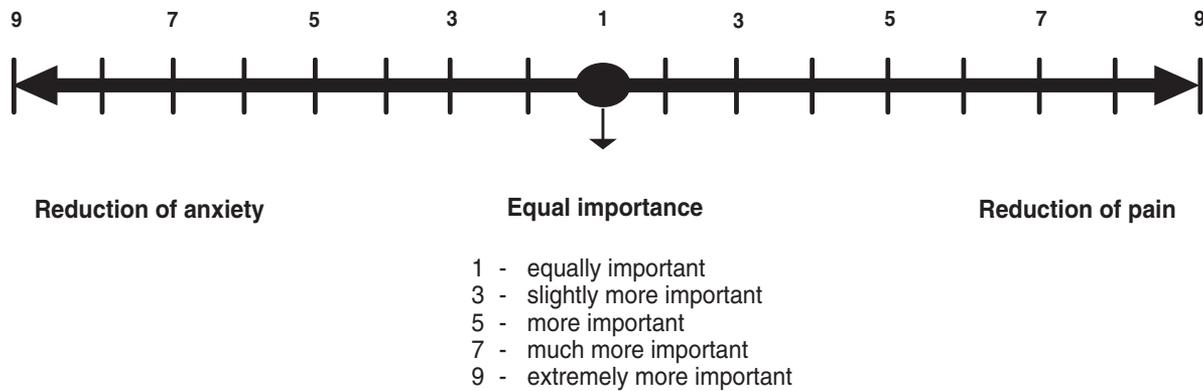
1 - equally important
3 - slightly more important
5 - more important
7 - much more important
9 - extremely more important

Figure 1. Example of a pairwise comparison of two quality-of-life-related treatment endpoints rated on a 9-point scale.

clinical evidence supporting the endpoint as well as methodological constraints of AHP (especially comprehensiveness and independence of decision criteria).

In preparation of the AHP workshops, interviewees received written information on which treatment endpoints were to be included in the workshops and how these were defined. Patients received these explanations of endpoints in lay language. Because the definition of the endpoints included and a common understanding among participants of what they mean are crucial, questions regarding definitions were addressed before the initiation of the pairwise comparison procedures. The definition of endpoints was based on the definitions in the IQWiG benefit assessments, for example, the endpoint "response" was defined as achieving a 50 percent reduction of symptoms in an acute episode of depression. Based on all pairwise comparisons, individual weights for each endpoint were calculated with Saaty's mathematical method (right eigenvector) and the group (geometric) means were then generated separately for the group of patients and the group of professionals (7).

In the workshops, each pairwise comparison between endpoints was accompanied by a short discussion of individual ratings within the respective group. These group discussions facilitated understanding of the motivation of patients and professionals as to why they considered one endpoint more or less important than another one. Furthermore, each pairwise comparison was followed by a "consistency check," that is, the CR was calculated as a measure of performance within AHP. The calculation of weights and the respective CRs were performed by means of the software package Expert Choice 11.

RESULTS

Twelve patients and seven professionals participated in the AHP patient workshop and professional workshop, respectively. Table 1 contains the patient-based and professional-based weights for each individual outcome measure.

Table 1. Priorities of Patients and Professionals Regarding Endpoints of Antidepressant Treatment

Patient-relevant outcome measure	Group priority Patients	Group priority Professionals (position in rank order)
Response	0.324	0.061 (5)
Improvement of cognitive function	0.125	0.062 (4)
Reduction of anxiety	0.118	0.054 (6)
Improvement of social function	0.107	0.090 (3)
Avoidance of relapse	0.091	0.144 (2)
Remission	0.085	0.475 (1)
Reduction of pain	0.054	0.033 (7)
No other serious adverse events	0.039	0.029 (8)
No (attempted) suicide	0.026	0.022 (9)
No other adverse events	0.023	0.020 (10)
No sexual dysfunction	0.007	0.007 (11)

"Response" was rated the most important endpoint by the group of patients, followed by the endpoint "improvement of cognitive function." The endpoint "remission" received the highest weighting within the group of professionals, followed by "avoidance of relapse." The professionals rated "sexual dysfunction" lowest, as did the patients. In both workshop groups, there was good consistency in the results (CR < 0.1). The set of the six most important patient-relevant outcome measures was the same for both groups. Three of these are related to efficacy: response, remission, and no relapse. Three are related to different aspects of quality of life: improvement of social functioning, improvement of cognitive function and reduction of anxiety. These six endpoints covered 85 percent of the overall weights in the patient group and 89 percent in the professional group.

The group discussions held in between the pairwise comparisons offered an insight into the question as to why certain outcome measures were more important to patients and professionals than others. For example, while patients in an acute

episode of major depression focused on a fast response and relief, professionals stated a preference for full remission and prevention of relapse. The results should not, however, be interpreted as a lack of interest by patients in reaching full remission. The patients' vote can be understood as a preference for a drug that helped to achieve a fast response over a drug that would be less effective in this regard, but more effective in reaching full remission after some time. Patients explained this by their experience that most will never function fully again, so that the rapid attainment of a moderate level of functioning in an acute episode of depression would be of more value. Group discussions also showed that the criteria and sub-criteria related to the adverse effects of antidepressants were—when traded off against efficacy or quality-of-life endpoints, considered either "of less consequence" (minor adverse events) or infrequent and therefore posing a "less immediate threat" (serious adverse events such as suicide).

DISCUSSION

This study indicated that the AHP method for pairwise comparisons of individual treatment endpoints was well handled by patients and professionals. It also delivered consistent results regarding the generation of preference-based weights. Finally, it provided valuable insights into the interviewees' reasoning as to why they considered certain endpoints more (or less) important than others. While the study focused on the elicitation of patient preferences, the AHP results for professionals helped to detect deviances from what patients consider important in their treatment. The inclusion of the professional group was considered important to contrast the results of both groups. With its small sample size (twelve patients and seven professionals), the study did not aim to deliver representative results, but to explore the AHP method as a practical tool in the quantification of preferences with respect to treatment endpoints. It is, however, notable in this context that AHP workshops are most effectively organized with small groups of individuals (twenty at most), allowing for effective group discussions to be held in between pairwise comparisons. To make AHP more representative, a precise definition of the subgroup of patients to be targeted would first be needed or stratified workshops (e.g., by age or gender) would need to be conducted. In addition, patient workshops in different regions of Germany could be held to increase the overall number of participants and to produce a better geographic representation of patients. In addition to the in-person workshops, an Internet-based elicitation of preferences could be considered. However, the resources required to conduct AHP surveys in the proposed manner would need to be further explored before extending its application.

A major challenge in this study was the selection of endpoints from the IQWiG benefit assessments for inclusion in the AHP decision hierarchy. While the set of criteria should be as comprehensible as possible in such hierarchy,

conceptional overlap between criteria should be minimized. Because some degree of overlap is inherent in many of the endpoints of antidepressant treatment (e.g., functional impairment is likely to be associated with anxiety), complete avoidance of overlap was impossible. Providing a very precise definition of each endpoint traded-off against each other in the pairwise comparisons helped reduce the potentially negative effects of overlap that might result in an overestimation of a particular endpoint in the final weights (e.g., anxiety in the example above).

Although there are different quantitative methods for eliciting patients' preferences in the form of utilities (e.g., QALY approach) and contingent valuation methods (e.g., WTP) in health economic evaluation within HTA, these methods require patients to state their preferences regarding their health status in aggregate. In contrast, methods for eliciting patient preferences such as the AHP allow for a decompositional valuation of health status by extracting and highlighting the important elements. They permit a reduction in the number of trade-offs to be made at one time by a patient in a complex decision setting to a choice between only two decision criteria in each pairwise comparison within AHP. A strength of the AHP method is the reduced cognitive burden for the patient by decomposing a complex decision problem into a limited number of pairwise comparisons.

The validity of AHP in the context of utility theory and in comparison with other preference assessment tools, such as conjoint analysis, has also been discussed (11;14). While a full discussion of this issue would go beyond the scope of this study, one advantage of AHP may even be that interviewees do not have to comply with the characteristics of the so called "homo-economicus." The pairwise comparison of criteria is in accordance with human behavior, especially if it is based on bounded rationality. Saaty's method of deriving priorities from pairwise comparisons based on matrix multiplication and the eigenvector calculation is mathematically sophisticated. Moreover, AHP provides for control of the logic and quality of decisions by using the inconsistency ratio to check for stability of rankings. A weakness of the method, however, is in particular the risk of rank reversal, for example, when adding or removing decision criteria.

While patients should be involved in HTA to some degree and at different stages of an assessment and the resulting decision-making processes, AHP (or another MCDA method) would most likely be restricted to situations where a quantification of patient preferences can precede or be directly integrated into the HTA and its results, for example, by selecting, prioritizing, or weighting patient-relevant endpoints of treatment. In Germany, prioritization of HTA assessment topics, as well as the appraisal and decision making for the statutory health insurance, take place at the level of a body called the Federal Joint Committee. AHP may be applicable here. Furthermore, the AHP method could help decide which endpoints matter, even before an HTA (benefit assessment) within IQWiG is started. AHP could also serve as a

decision aid in determining which endpoints of an economic evaluation should receive a higher (or lower) weighting in the aggregation of results. In Germany, such an aggregation may be necessary to derive reimbursable prices for new medications and could be part of IQWiG's cost-effectiveness evaluations.

Although AHP has primarily been developed to support decision making, it may also play a role in (i) identifying or prioritizing patient-relevant outcomes when clinical trials are designed and (ii) analyzing the net benefit of health interventions based on AHP results. In the first application, national HTA institutions could provide advice to the pharmaceutical industry as to which treatment endpoints are considered to be of importance. In the second proposed area, it is possible to determine weights for endpoints by developing a hierarchical structure of the outcome measures considered and to base net benefit assessments on these weighted outcome measures.

CONCLUSION

As patient involvement will be further strengthened in many jurisdictions in the future, the consideration of tools such as AHP is essential. Germany seems to be entering a new era by moving toward full acknowledgment of patients' preferences, and by increasing its consideration of different qualitative and quantitative methods for preference elicitation.

CONTACT INFORMATION

Marion Danner, DiplVw, MPH (marion.danner@iqwig.de), Health Economics Department, Institute for Quality and Efficiency in Health Care (IQWiG), Dillenburger Straße 27, 51105 Cologne, Germany

J. Marjan Hummel, PhD (j.m.hummel@utwente.nl), Department of Health Technology and Services Research, University of Twente, Drienerloaan 5, 7500 AE Enschede, The Netherlands

Fabian Volz, DiplKfm (Fabian.volz@iqwig.de), Health Economics Department, Institute for Quality and Efficiency in Health Care (IQWiG), Dillenburger Straße 27, 51105 Cologne, Germany

Jeannette G. van Manen, PhD (j.g.vanmanen@utwente.nl) Department of Health Technology and Services Research, University of Twente, Drienerloaan 5, 7500 AE Enschede, The Netherlands

Beate Wiegard, MA (beate.wiegard@iqwig.de), Health Information Department, Institute for Quality and Efficiency in Health Care (IQWiG), Dillenburger Straße 27, 51105 Cologne, Germany

Charalabos-Markos Dintsios, PhD, MA, MPH (charalabos-markos,dintsios@iqwig.de) Health Economics Department, Institute for Quality and Efficiency in Health Care (IQWiG), Dillenburger Strasse 27, 51105 Cologne, Germany

Hilda Bastian (Hilda.bastian@nih.gov), Institute for Quality and Efficiency in Health Care (IQWiG), Dillenburger Straße 27, 51105 Cologne, Germany; National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, 8600 Rockville Pike, Bethesda, Maryland 20894

Andreas Gerber, PhD, MD, Dipl-Theol, MA, MSc (andreas.gerger@iqwig.de), Health Economics Department, Institute for Quality and Efficiency in Health Care (IQWiG), Dillenburger Straße 27, 51105 Cologne, Germany

Maarten J. IJzerman, PhD (m.j.ijzerman@utwente.nl), Department of Health Technology and Services Research, University of Twente, Drienerloaan 5, 7500 E Enschede, The Netherlands

CONFLICT OF INTEREST

REFERENCES

1. Bana e Costa CA, Vansnick JC. A critical analysis of the eigenvalue method used to derive priorities in AHP. *Eur J Oper Res*. 2008;187:1422-1428.
2. Barzilai J, Golany B. AHP rank several, normalization and aggregation rules. *Inf Syst Oper Res*. 1994;32:57-64.
3. Benaim C, Perennou DA, Pelissier JY, Daures JP. Using an analytical hierarchy process (AHP) for weighting items of a measurement scale: A pilot study. *Rev Epidemiol Sante Publique*. 2010;58:59-63.
4. Bridges JF, Jones C. Patient-based health technology assessment: A vision of the future. *Int J Technol Assess Health Care*. 2007;23:30-35.
5. Dolan JG. Shared decision-making - transferring research into practice: The Analytic Hierarchy Process (AHP). *Patient Educ Couns*. 2008;73:418-425.
6. Dolan JG, Bordley DR. Individualized decision-making using the Analytic Hierarchy Process (AHP): Reliability, validity and clinical usefulness. *Med Decis Making*. 1991;11:322-322.
7. Dolan JG, Isselhardt BJ Jr, Cappuccio JD. The Analytic Hierarchy Process in medical decision making: A tutorial. *Med Decis Making*. 1989;9:40-50.
8. Facey K, Boivin A, Gracia J, et al. Patients' perspectives in health technology assessment: A route to robust evidence and fair deliberation. *Int J Technol Assess Health Care*. 2010;26:334-340.
9. Hailey D, Nordwall M. Survey on the involvement of consumers in health technology assessment programs. *Int J Technol Assess Health Care*. 2006;22:497-499.
10. Hummel JM, IJzermann MJ. *The use of the Analytic Hierarchy Process in health care decision making*. Enschede: University of Twente; 2009.
11. Ijzerman M, van Til J, Snoek G. Comparison of two multi-criteria decision trechniques for eliciting treatment preferences in people with neurological disorders. *Patient*. 2008;1:265-273.
12. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. *Bupropion, Mirtazapin und Reboxetin bei der Behandlung*

der Depression: Abschlussbericht; Auftrag A05-20C; IQWiG-Bericht 68. Köln: IQWiG; 2009. http://www.iqwig.de/download/A05-20C_Abschlussbericht_Bupropion_Mirtazapin_und_Reboxetin_bei_Depressionen.pdf (accessed May 26, 2011).

13. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. *Selektive Serotonin- und Noradrenalin-Wiederaufnahmehemmer (SNRI) bei Patienten mit Depressionen: Abschlussbericht; Auftrag A05-20A; Version 1.1; IQWiG-Bericht 55*. Köln: IQWiG; 2010. http://www.iqwig.de/download/A05-20A_Abschlussbericht_SNRI_bei_Patienten_mit_Depressionen_V1–1.pdf (accessed May 26, 2011).

14. Kallas Z, Lambarraa F, Gil JM. A stated preference analysis comparing the Analytical Hierarchy Process versus choice experiments. *Food Qual Prefer*. 2011;22:181-192.

15. Liang L, Wang G, Hua Z, Zhang B. Mapping verbal responses to numerical scales in the Analytic Hierarchy Process. *Socioecon Plann Sci*. 2008;42:46-55.

16. Liberatore MJ, Nydick RL. The Analytic Hierarchy Process in medical and health care decision making: A literature review. *Eur J Oper Res*. 2008;189:194-207.

17. Mulye R. An empirical comparison of three variants of the AHP and two variants of conjoint analysis. *J Behav Decis Mak*. 1998;11:263-280.

18. Ryan M, Scott DA, Reeves C, et al. Eliciting public references for healthcare: A systematic review of techniques. *Health Technol Assess*. 2001;5:1-186.

19. Saaty TL. A scaling method for priorities in hierarchical structures. *J Math Psychol*. 1977;15:234-281.

20. Saaty TL. Group decision making and the AHP. In: Golden BL, Wasil WA, Harker PT, eds. *The analytic hierarchy process: Applications and studies*. Berlin: Springer; 1989:59-67.

21. Saaty TL. Decision making with the Analytic Hierarchy Process. *Int J Serv Sci*. 2008;1:83-98.

22. Saaty TL, Vargas LG. The Analytic Hierarchy Process: Wash criteria should not be ignored. *Int J Manage Decis Mak*. 2006;7:180-188.

23. Scheibler F, Scheike I, Dintsios CM. Patientenpartizipation bei Festlegung und Gewichtung von Behandlungszielen: Status quo und Entwicklungspotenziale. *Z Evid Fortbild Qual Gesundhwes*. 2008;102:373-377.

24. Scholl A, Manthey L, Helm R. Solving multiattribute design problems with Analytic Hierarchy Process and conjoint analysis: An empirical comparison. *Eur J Oper Res*. 2005;164:760-777.

25. Vogt F, Schwappach DLB, Bridges JFP. Accounting for tastes: A German perspective on the inclusion of patient preferences in healthcare. *Pharmacoeconomics*. 2006;24:419-423.