


Improving Hospital Based Medical Procurement Decisions with Health Technology Assessment and Multi-Criteria Decision Analysis

INQUIRY: The Journal of Health Care
Organization, Provision, and Financing
Volume 58: 1–13
© The Author(s) 2021
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/00469580211022911
journals.sagepub.com/home/inq



Chai Yang, MSc^{1,2}, Yanjun Wang, ME¹, Xiaoxuan Hu, DM¹ ,
Yujun Chen, BE^{2,3}, Liting Qian, MD², Fuchang Li, DM⁴, Wei Gu, BE²,
Qiang Liu, MSc⁵, Di Wang, MD², and Xiaoqing Chai, MD^{2,6}

Abstract

This research proposes a new medical procurement decision-making tool named Evaluation of Medical Technology Procurement (EMTP), which combines Mini-health Technology Assessment (Mini HTA) with the analytic hierarchy process (AHP), as well as the intuitionistic linguistic multi-criteria group decision model for multi-criteria decision analysis (MCDA). This tool was applied to a medical device procurement decision in a large provincial general hospital with more than 5000 beds in China as a case study. Specifically, the AHP evaluation framework is first established to determine the evaluation dimensions and criteria. This goal is achieved by applying the AdHopHTA Mini-HTA template and gathering data from questionnaires completed by experts from 33 major public hospitals in Anhui Province, China. The professionals within the application hospital were invited to evaluate the alternative products in a pairwise comparison and obtain a ranking of their advantages and disadvantages. This goal is achieved using the intuitionistic linguistic fuzzy model to deal with the subjectivity and uncertainty that may be present in the professional evaluation by experts in different fields. At the same time, the Keeney-Raiffa MCDA (KRM) method was used to demonstrate the accuracy of the application results. The results show that our tool can achieve the same effect as the verification method while being more efficient, easier to use, and requiring fewer participants. The advantages and disadvantages of several evaluation methods combined with multi-criteria methods are discussed, including verification methods, pointing out the advantages and limitations of this research tool as well as the prospects for the future.

Keywords

health technology assessment, multi-criteria decision analysis, analytic hierarchy process, hospitals, clinical decision-making

- **What do we already know about this topic?**

At the hospital level, the introduction of new healthcare technologies needs to consider the influence of multiple stakeholders. These complex factors often lead to inefficient group decision-making and lack of scientific.

- **How does your research contribute to the field?**

EMTP creates a unique weight adjustment process by considering the uncertainty and evaluation biases influenced by hospital policies or the characteristics of individual experts. The outcome of the decisions made using our method will vary depending on the needs and goals of the hospital, so that they can achieve a better effect while being more efficient, easier to use, and requiring fewer participants.

- **What are your research's implications toward theory, practice, or policy?**

The research results show that the creation of quick and efficient management tools to compensate for blind spots in decision-making that cannot be covered by HTA is very important for the hospital's acquisition of new technology, and the evaluator's personal psychological characteristics have a significant impact on the results. Without our method, it would be necessary for each person involved in decision-making to conduct more detailed interviews to improve their understanding of the process to obtain accurate results.



Introduction

At the hospital level, the introduction of new health technologies is often very complex, and multiple factors such as risks, costs, results, and effects need to be considered in the process. These factors increase internal and external uncertainty and make decision-making more difficult.¹⁻³ Traditional decision-making, focusing on price, cost, profit, or another single economic goal as defined by the procurement committee, will risk adverse consequences such as harming the interests of patients, reducing the efficacy of treatment, and increasing the burden on the medical insurance fund, leading to a waste of medical and health resources. Current methods can no longer meet the needs of improving the management quality of smart hospitals,⁴ and we propose that it is necessary to link health technology assessment (HTA) with product selection, pricing, and procurement.⁵

Mini-health technology assessment (Mini-HTA) is a tool that applies the principles and methods of traditional HTA and evidence-based medicine, to provide a decision-making framework for hospital decision-makers to introduce relevant health technologies. The tool is based on the actual needs of the hospital and can make a comprehensive systematic evaluation of relevant health technologies.^{6,7} Mini-HTA is implemented through a series of evaluation checklists composed of standard items. The representative one is the evaluation checklist developed by the Danish Center for Health Technology Assessment (DACEHTA), which includes the 4 dimensions of technology, patients, organization, and economy.⁸ Compared with traditional HTA, the Mini-HTA assessment checklist has been improved in terms of assessment time, assessment scope, and assessment content, but is still difficult to implement in many hospitals. This is because hospitals often do not have enough qualified personnel to complete the assessment work, and there is a lack of data to support the decision-making.^{9,10} Therefore, the introduction of new technologies at the hospital level is still not transparent, and managers are often unclear about the role of some important reference standards in the final decision.¹¹ Choosing suitable scoring and weighting techniques to address the above drawbacks is another problem. In addition, HTA generates a report that is only part of the decision-making process, so it cannot easily be used directly for local decision-making.¹² Hospital procurement departments must

analyze the characteristics of procurement decisions based on compliance with health technology access norms and procedures, use information methods and intelligent algorithms to build decision-making models, and make accurate judgments on medical procurement decision-making problems with complex information, rich content, and multiple goals.

At present, more researchers are beginning to explore the integration of Mini HTA with other disciplinary approaches to improve the quality of purchasing decisions. Sloane¹³ developed an HTA application tool using AHP to assist in measuring user perceptions based on pairwise comparisons. Although this method reduces the difficulty of evaluation to a certain extent, there has been no in-depth study of the effect of the weighting distribution of experts, and the evaluation method is still in numerical form. Goetghebeur et al¹⁴ piloted the EVIDEM framework for HTA, where researchers appraising 10 medicines covering 6 therapeutic areas demonstrated its effectiveness. This combined method is frequently used in medical decision-making, but one thing that cannot be ignored is that if there is a continuous relationship between the score of the criterion and the measured value in the model, there may be a problem of information loss. Martelli et al¹⁵ developed an Innovative Device Assessment (IDA) tool that used the PAPRIKA method to provide a more user-friendly interface while contributing to the structure and transparency of HTA decisions. They estimate that the evaluation time can be shortened to 5 to 10 minutes, but the consumption of hidden time is still unknown. In fact, a simple interface requires the planner to spend additional time and apply expert knowledge to prepare and program the guidelines.

These shortcomings in traditional approaches have prompted multi-criteria decision analysis MCDA to be widely accepted, because it has the potential to consider the judgmental criteria of stakeholders.¹⁶ MCDA is a collection of analysis techniques used to support decision making on multiple and conflicting criteria.¹⁷ It provides a wealth of methods derived from practice and already has a broad application base in other fields.¹⁸⁻²³ Health care decision-making can be regarded as an MCDA problem; thus, the importance and use of MCDA in health care decision-making has increased in the past few years. MCDA is regarded as a natural extension of evidence-based medical research and practice such as health technology assessment. By constructing alternatives and quantifying evidence to help decision

¹Hefei University of Technology, Hefei, P.R. China

²The First Affiliated Hospital of University of Sciences and Technology of China, Hefei, P.R. China

³Anhui Provincial Medical Equipment Management and Quality Control Center, Hefei, Anhui, P.R. China

⁴Yunnan Normal University, Kunming, Yunnan, P.R. China

⁵China Pharmaceutical University, Nanjing, Jiangsu, P.R. China

⁶Anhui Provincial Anesthesia Quality Control Center, Hefei, Anhui, P.R. China

Received 3 February 2021; revised 4 May 2021; revised manuscript accepted 14 May 2021

Corresponding Author:

Xiaoxuan Hu, School of Management, Hefei University of Technology, No. 193 Tunxi Road, Hefei, Anhui Province 230009, P.R. China.

Email: xiaoxuanhu@hfut.edu.cn

makers make the best choice while eliminating conflicts among stakeholders,²⁴ MCDA is “Especially suitable for scenarios that combine hard data with subjective preferences or require multiple decision makers to weigh in.”²⁵ In addition, the use of MCDA can help suppliers clarify what factors are valued by hospitals and help them focus on providing the most important data to decision makers.²⁶

Atanassov²⁷ proposed the use of intuitionistic fuzzy sets (IFS), which are characterized by degrees of both membership and non-membership. IFS is playing an increasingly important role internationally in dealing with MCDA problems because it has the advantage of reflecting imprecise reasoning in human cognitive processes.^{28,29} The use of linguistic terms are another way to model human cognition. In general, experts deal with many problems using linguistic terms to express qualitative information such as “bad,” “tolerable,” “average,” and “good.”³⁰ In some complex decision-making environments, experts express their opinions by combining linguistic terms with IFS, so that imprecise information from subjective and objective environments can be better represented. As a result, the intuitionistic linguistic number (ILN) was developed, which is a combination of intuitionistic fuzzy numbers and linguistic terms,³¹ which can comprehensively describe the fuzzy features of things. As ILN incorporates the uncertainty of human cognition into MCDM, it has attracted a lot of attention in both theoretical and practical fields. Also, some studies of intuitionistic linguistic information have been proposed and applied to MCDA.³²

In MCDA, aggregation operators provide a method to obtain comprehensive evaluation results with respect to different criteria. Recently, the aggregation operator has become an indispensable tool and a growing research field in the process of seeking information fusion in decision making. In cases where the criteria of alternatives are difficult to evaluate, experts tend to solve decision-making problems by making pairwise comparisons of alternatives and eventually establishing preference relations among the alternatives. Each preference value represents the expert’s degree of preference for one alternative over another. HTA has direct effects on patient’s health and lives, so the evaluations of the medical procurement suppliers must be accurate and reliable. The intuitionistic linguistic preference relations can effectively simulate conditions where experts cannot compare some pairs of alternatives, helping experts to avoid giving biased opinions, which makes it a good fit for the selection of medical suppliers.

In contrast to most research, which focuses on decision criteria, the overarching goal of this study is to develop a practical tool for medical procurement decision-making called evaluation of medical technology procurement (EMTP) that can be applied in a systematic manner, and that can accommodate all relevant decision criteria to be considered. This research makes 3 main contributions. First, EMTP combines the AHP method in MCDA and the fuzzy preference relationship multi-criteria group decision-making method in the intuitive linguistic environment. Here, AHP is

used to establish the evaluation level, while the operator, namely the intuitionistic linguistic power generalized compensative weighted averaging (ILPGCWA),³³ is used to aggregate the evaluation information between criteria and decision makers in the evaluation process. This is done by means of intuitionistic linguistic preference relations, which can consider a very fine range of compensatory effects by setting possible combinations of parameters. Second, EMTP creates a unique weight adjustment process, as it considers the uncertainty and the evaluation tendency influenced by hospital policies or other characteristics of experts. The outcome of the decision will vary depending on the needs and goals of the hospital, and it can achieve a better effect while being more efficient, easier to use, and requiring fewer participants. Third, EMTP can be applied in a systematic manner, which can accommodate all relevant decision criteria under consideration.

Materials and Methods

Data Collection

This research was conducted in a large provincial general hospital, which has 5450 open beds; the anesthesia workload in the operating room reached more than 60000 cases in 2019.³⁴ This hospital is the leading unit of the provincial medical device industry association and is the provincial center for anesthesia quality control. This site was chosen as the research and data collection hub to ensure the scientific and authoritative evaluation of the questionnaire. We applied the EMTP to the procurement evaluation of different brands of IPWE. Such equipment helps maintain the normal body temperature of patients during the perioperative period and is an indispensable medical device during surgical anesthesia.³⁵⁻³⁷ The study was approved by the hospital’s Ethics Committee, and all participants provided written informed consent before completing the study.

The questionnaire survey was conducted by members of the medical equipment committee of 33 tertiary hospitals in Anhui Province. These members were mainly the director of the clinical department (anesthesiology) and the director of the hospital administrative department (purchasing department). Two rounds of questionnaires were issued. The first round of questionnaires (see Supplemental Appendix 1) was used to determine the evaluation dimensions and criteria, which provided the basis for the EMTP method and the verification of KRM. The second round of the questionnaire (see Supplemental Appendix 2) was to score the weight of each evaluation dimension independently of the product brand, which provides the basis for the KRM verification.

EMTP Method Implementation

The EMTP can be divided into 3 parts according to the workflow and main participants (see Figure 1). The process is

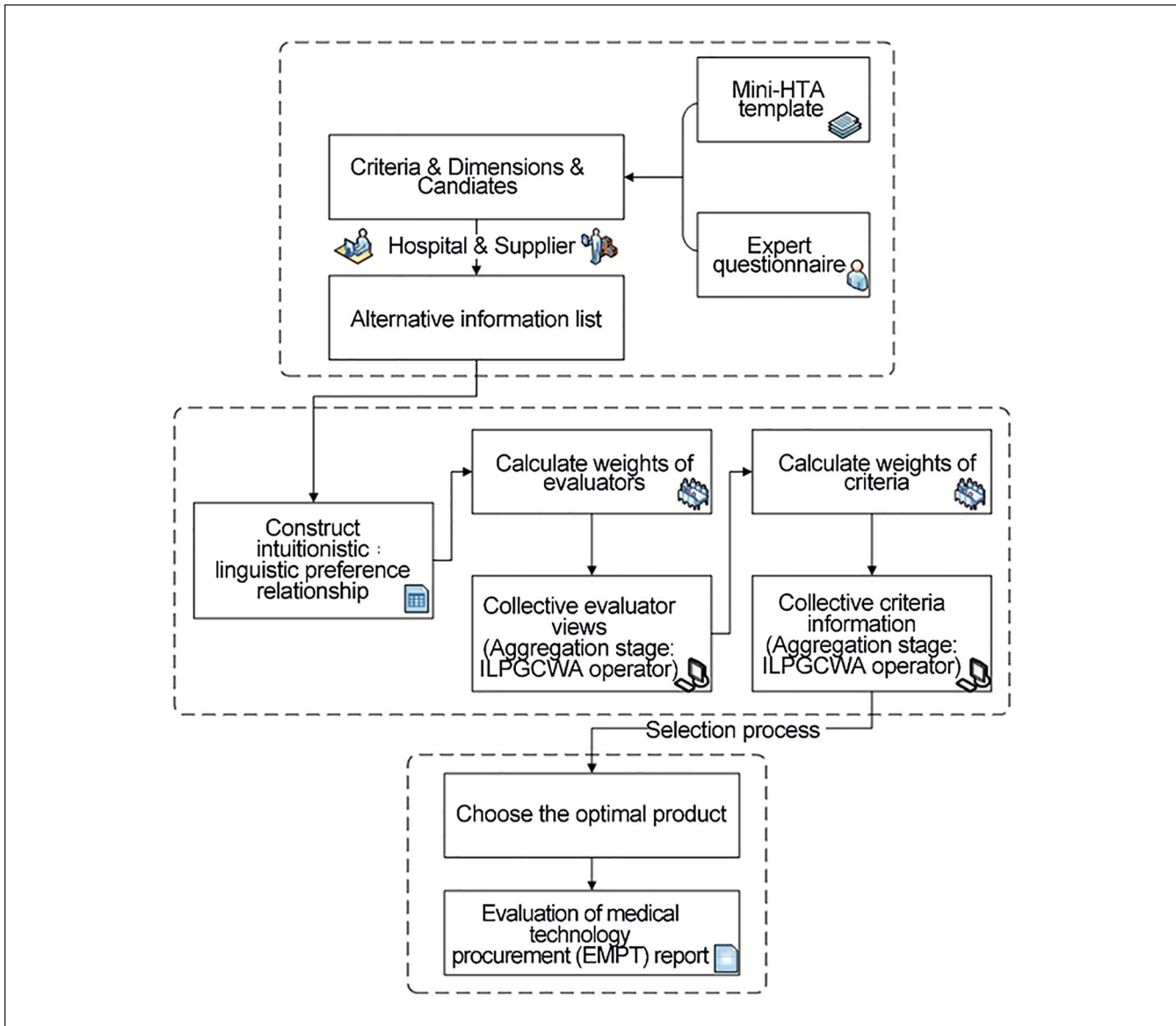


Figure 1. Method flow chart.

carried out from top to bottom. The first part is the preparation for decision-making, which is used to set-up the specific content of the expert's score. This serves the second part. The second part is the specific scoring and calculation process, which presents the results of different products or programs to the decision maker. The third part is the final decision of the decision maker, regarding which product to choose or give up.

Determine the Evaluation Product and its Dimensions and Criteria

The research team chose the AdHopHTA Mini-HTA template³⁸ as the reference source for the first round of expert questionnaires. This template has been approved in several countries as well as by the European Union. In this part,

based on the first round of expert questionnaires and the Mini-HTA checklist, the evaluation dimensions (second-level indicators) under the decision objective were determined and are denoted as $D = \{D_1, D_2, \dots, D_n\}$, and the specific evaluation criteria of each dimension (the third level indicators), which are denoted as C_c ($c = 1, 2, \dots, m$), respectively, were selected and entered into the final evaluation system. The product supplier and relevant decision makers of HB-HTA fill in the specific parameters of each product under each standard for the next evaluation service.

Evaluation by Decision Makers

Obtain evaluation information given by experts. This section requires the hospital to provide evaluation information

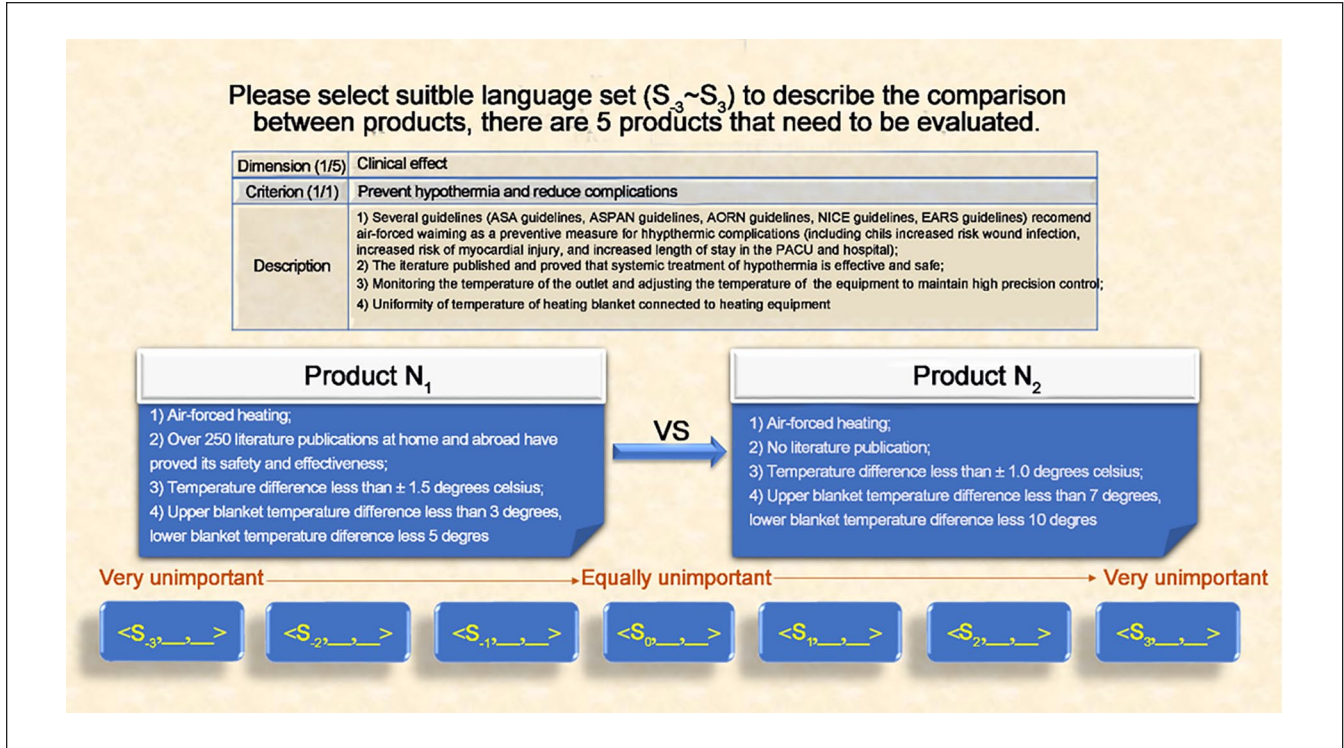


Figure 2. Expert evaluation interface.

according to its own needs by decision makers from different departments, such as the financial department, doctors, ethics committee members, and patients. Based on $C_c (c = 1, 2, \dots, m)$ criteria, $Y_y (y = 1, 2, \dots, t)$ evaluators make pairwise comparisons of $N_n (n = 1, 2, \dots, r)$ brands of insulation equipment, and give comparative evaluation information to establish a decision matrix:

$$\left(B_c^y \right)_{r \times r} = \left(\beta_{ij}^{(y,c)} \right)_{r \times r} = \left\langle S_{g(ij)}, u \left(\beta_{ij}^{(y,c)} \right), v \left(\beta_{ij}^{(y,c)} \right) \right\rangle_{r \times r} \quad (1)$$

$(i, j = 1, 2, \dots, r)$.

where the evaluation information is described by intuitive linguistic preference relations, and possible values of the linguistic terms are represented by a set S ,³⁹ where $S = \{S_{g(ij)} | g = -f, \dots, -1, 0, 1, \dots, f\}$, and f is a positive integer.

Among them, $u \left(\beta_{ij}^{(y,c)} \right)$ and $v \left(\beta_{ij}^{(y,c)} \right)$ represent the degree of membership and non-membership, respectively. They satisfy the condition that $u \left(\beta_{ij}^{(y,c)} \right) \in (0, 1), v \left(\beta_{ij}^{(y,c)} \right) \in (0, 1), 0 \leq u \left(\beta_{ij}^{(y,c)} \right) + v \left(\beta_{ij}^{(y,c)} \right) \leq 1, u \left(\beta_{ij}^{(y,c)} \right) = v \left(\beta_{ji}^{(y,c)} \right), u \left(\beta_{ii}^{(y,c)} \right) = 1$ for all $i, j = 1, 2, \dots, r$.

In this case, after obtaining the list of alternative products, we set up a pairwise comparison linguistic set S , which is created as 7 levels from S_{-3} to S_3 , where S_0 is “equally important,” and the importance degree extends from S_0 to the 2

opposite extremes. Here, S_{-3} and S_3 represent “very unimportant” and “very important,” respectively (see Figure 2). The next task is to fill in the degree of membership and the degree of non-membership in the space after each linguistic term in each selection. The procurement division of Anhui Provincial Hospital invited the director of the anesthesia department (set to Y_1) and the director of the procurement department (set to Y_2). They evaluated different alternatives in pairs after obtaining the preference matrix on the decision platform, as shown in Figure 2.

Determine weight and aggregate information. The weights and criteria in the framework are variable because different hospitals may focus on different priorities. Based on this situation, decision-makers will be able to assign initial weights in our method. The focus of our research is to adjust the initial weight based on the expert’s judgments and evaluation criteria set by the hospital according to its own needs. The final weight is determined by the initial weight as well as the degree of membership and non-membership. In this research, we gave equal initial weight to all the experts. The initial weights of the product guidelines are calculated through the questionnaire in the first part. The distribution of these weights is consistent with the follow-up verification method and have been recognized by all decision makers participating in the research hospital. The advantage of this is that the results can be effectively verified and compared.

In the following, we use the intuitionistic linguistic power generalized compensative weighted averaging (ILPGCWA) operator.³³

Definition 1³³ Let $\beta_i = \langle s_{\theta(\beta_i)}, (u(\beta_i), v(\beta_i)) \rangle$ be a collection of ILNs. An intuitionistic linguistic power generalized compensative weighted averaging (ILPGCWA) operator on $(\beta_1, \beta_2, \dots, \beta_n)$ is given as follows:

$$ILPGCWA_{\lambda, p}(\beta_1, \beta_2, \dots, \beta_n) = \left\langle \begin{array}{l} S \\ \left(\log_{\lambda} \left(\sum_{i=1}^n w_i^* \lambda^{(u(\beta_i))^p} \right) \right)^{\frac{1}{p}}, \\ \left(\log_{\lambda} \left(\sum_{i=1}^n w_i^* \lambda^{(u(\beta_i))^p} \right) \right)^{\frac{1}{p}}, \\ \left(1 - \left(\log_{\lambda} \left(\sum_{i=1}^n w_i^* \lambda^{(1-v(\beta_i))^p} \right) \right)^{\frac{1}{p}} \right) \end{array} \right\rangle$$

where p is a parameter such that $p \in [-\infty, \infty]$, $p \neq 0$,

$0 < \lambda \leq \infty$, $w_i^* = \frac{w_i(1+T(\beta_i))}{\sum_{i=1}^n w_i(1+T(\beta_i))}$, $w_i^* \in [0, 1]$, and $\sum_{i=1}^n w_i^* = 1$,

and $w = (w_1, w_2, \dots, w_n)$ the weight vector of β_i ($i = 1, 2, \dots, n$)

with $w_i \in [0, 1]$ and $\sum_{i=1}^n w_i = 1$. $T(\beta_i) = \sum_{j=1, j \neq i}^n Sup(\beta_i, \beta_j)$

and $Sup(\beta_i, \beta_j) = 1 - d(\beta_i, \beta_j)$, $d(\beta_i, \beta_j)$ is defined in.³³

Theorem 1 Let $\beta_i = \langle s_{\theta(\beta_i)}, u(\beta_i), v(\beta_i) \rangle$ be a collection of ILNs, and $w = (w_1, w_2, \dots, w_n)$ be the weight vector of β_i ($i = 1, 2, \dots, n$), with $w_i \in [0, 1]$ and $\sum_{i=1}^n w_i = 1$. Then

1) (Idempotency): If $\beta_i = \beta$ ($i = 1, 2, \dots, n$), then

$$ILPGCWA_{\lambda, p}(\beta_1, \beta_2, \dots, \beta_n) = \beta.$$

2) (Boundary): $\beta^- \leq ILPGCWA_{\lambda, p}(\beta_1, \beta_2, \dots, \beta_n) \leq \beta^+$,

where $\beta^- = \langle \min_i \{s_{\theta(\beta_i)}\}, \min_i \{u(\beta_i)\}, \max_i \{v(\beta_i)\} \rangle$,

$$\beta^+ = \langle \max_i \{s_{\theta(\beta_i)}\}, \max_i \{u(\beta_i)\}, \min_i \{v(\beta_i)\} \rangle.$$

The specific application based on the ILPGCWA operator steps are as follows:

Step 1. Determine the final weight of evaluators and aggregate the evaluation information

When comparing 2 products, if the evaluator supports a certain product's criterion or the degree of matching is higher, their weight is increased. Conversely, if the evaluator is not sure which product should be supported, their weight

will be reduced. Since the evaluators give different degrees of support when scoring pairs of brands under different criteria, the weights obtained are also different. The weight calculation formula is:

$$w_{ij}^{y,c} = \frac{e_y (1 + T(\beta_{ij}^{(y,c)}))}{\sum_{y=1}^t e_y (1 + T(\beta_{ij}^{(y,c)}))}. \quad (2)$$

Here, e_y is the initial weight of decision maker y .

$$T(\beta_{ij}^{(y,c)}) = \sum_{l=1, l \neq y}^t Sup(\beta_{ij}^{(y,c)}, \beta_{ij}^{(l,c)}). \quad (3)$$

where $Sup(\beta_{ij}^{(y,c)}, \beta_{ij}^{(l,c)}) = 1 - d(\beta_{ij}^{(y,c)}, \beta_{ij}^{(l,c)})$, $d(\beta_{ij}^{(y,c)}, \beta_{ij}^{(l,c)})$ is defined in.³³

Also, $Sup(\beta_{ij}^{(y,c)}, \beta_{ij}^{(l,c)})$ represents $\beta_{ij}^{(y,c)}$'s support for $\beta_{ij}^{(l,c)}$

which is the degree of support for the comparative linguistic evaluation values of any 2 alternative products i and j under the same criterion c between different experts y and l .

Step 2. The ILPGCWA is used to aggregate the information from each expert

$$\beta_{ij}^{(c)} = ILPGCWA_{\lambda, p}(\beta_{ij}^{(1,c)}, \beta_{ij}^{(2,c)}, \dots, \beta_{ij}^{(n,c)}) = \left\langle \begin{array}{l} S \\ \left(\log_{\lambda} \left(\sum_{y=1}^t w_y^{(y,c)} \lambda^{(u(\beta_{ij}^{(y,c)}))^p} \right) \right)^{\frac{1}{p}}, \\ \left(\log_{\lambda} \left(\sum_{y=1}^t w_y^{(y,c)} \lambda^{(u(\beta_{ij}^{(y,c)}))^p} \right) \right)^{\frac{1}{p}}, \\ \left(1 - \left(\log_{\lambda} \left(\sum_{y=1}^t w_y^{(y,c)} \lambda^{(1-v(\beta_{ij}^{(y,c)}))^p} \right) \right)^{\frac{1}{p}} \right) \end{array} \right\rangle. \quad (4)$$

Among them, λ and p are 2 set parameters, where λ represents the attitude of the evaluator, which satisfies $0 \leq \lambda \leq \infty$, $\lambda \neq 1$. When $0 < \lambda < 1$ this represents a negative attitude, when $\lambda \rightarrow 1$, it represents a neutral attitude, when $\lambda > 1$, this represents a positive attitude of the evaluator in the decision-making process. Parameter p represents the evaluator's preference in the comparison between different products in the evaluation process, $p \in (-\infty, +\infty)$, $p \neq 0$. In this case, before starting the evaluation, we used the interviews to understand and synthesize the views of all the participating experts on their role, in addition to providing feedback information in the context of their own professions (business administration, and anesthesiology) regarding the evaluation of the various

components of the health technology. We understand that overall, the experts have a very positive attitude toward the new evaluation method, and there is no obvious preference for each product. According to this, we suppose $\lambda = 8, p = 10^{(-10)}$.

The assignment of weights in our method is done by the decision makers. We have adopted an equal weight distribution across the clinical and procurement departments, and a setting of different weights and criteria that is consistent with the weight distribution in the verification method. This was approved by the decision makers of all participating research hospitals. The advantage of this is that the quality of statistics and calculation is consistent, and the results obtained are more convincing.

Step 3. Determine the final criteria weights and aggregate the criteria information

Similarly, the criteria weight is calculated based on the information given by the evaluator. The higher the degree of support for the criteria, the greater the weight. As the degree of support for the criteria is different when scoring between different brands, the weighting obtained is also different.

$$w_{ij}^{(c)} = \frac{w_j \left(1 + T\left(\beta_{ij}^{(c)}\right)\right)}{\sum_{c=1}^m w_j \left(1 + T\left(\beta_{ij}^{(c)}\right)\right)}. \quad (5)$$

Among them, w_j is the initial weight corresponding to criterion c_j , and meets the condition $\sum_{c=1}^m w_c = 1$.

Step 4. The ILPGCWA is used to aggregate the criteria information

$$\beta_{ij} = ILPGCWA_{\lambda, p} \left(\beta_{ij}^{(1)}, \beta_{ij}^{(2)}, \beta_{ij}^{(m)} \right) = \left\{ \begin{array}{l} S \\ \left(\log_{\lambda} \left(\sum_{c=1}^m w_j^{(c)} \lambda^{\left(\beta_{ij}^{(c)} \right)^p} \right) \right)^{\frac{1}{p}}, \\ \left(\log_{\lambda} \left(\sum_{c=1}^m w_j^{(c)} \lambda^{\left(u\left(\beta_{ij}^{(c)}\right) \right)^p} \right) \right)^{\frac{1}{p}}, \\ \left(1 - \left(\log_{\lambda} \left(\sum_{c=1}^m w_j^{(c)} \lambda^{\left(1 - v\left(\beta_{ij}^{(c)}\right) \right)^p} \right) \right)^{\frac{1}{p}} \right) \end{array} \right. \quad (6)$$

Finally, the β_i corresponding to each plan and its ranking were obtained using the formula $\beta_i = \frac{\sum_{j=1}^r \beta_{ij}}{n}$, after which the optimal product was selected by comparing the obtained values of β_i . We developed a MATLAB R2018a program to calculate the final results.

KRM method validation. The assessment results were verified in comparison with the Keeney and Raiffa⁴⁰ MCDA method. This has good practice and process references from the International Society for Pharmacoeconomics and Outcomes Research (ISPOR).^{41,42} The relationship between the measured value and the score of the benefit criterion in this method is continuous, reflecting the performance range of the alternatives whose weights are evaluated⁴³ and the weight of the criterion. It is suitable for verification of the proposed method because the verifiers participating in this decision have enough knowledge to make the relevant expert judgments.

Scoring alternatives. According to the evaluation dimensions, criteria, and product comparisons obtained in the first step, the score for each dimension on each alternative product was measured by constructing a partial value function.^{17,44} The criterion score range for each dimension was 0 to 1 point. The best plan on each dimension scored 1 point. Alternative plans for each term were graded between 0 and 1 point according to the reviewer's preference. The evaluator assigned scores according to the performance of each product in all dimensions, according to their own experience.

Criterion weighting. The swing weighting method⁴⁵ was used to weight each criterion. The most important criterion was assigned a value of 100. Then, the values for this criterion were compared with the scores on the other criteria in pairs to determine the relative importance of each criterion, with values between 0 and 100 being assigned.

Calculating aggregate scores. The score for a single sample was calculated by a marginal preference function using the following formula:

$$V(n) = \sum_{c=1}^m w_c v_c(n). \quad (7)$$

In the formula, $V(n)$ is the overall value of product N , w_c is the relative importance of criterion C , and $v_c(n)$ is the score for the product on the c -th criterion.

Results

A total of 82 questionnaires were received from the 33 Anhui provincial hospitals. These included 52 from anesthesiology departments and 30 from the procurement and medical engineering departments between them.

Dimension and Criterion Screening Results

Among the 52 questionnaires obtained from anesthesiology departments, the evaluation criteria that received the highest number of votes (50) were clinical complications and after-sales service. Among the 30 questionnaires returned from

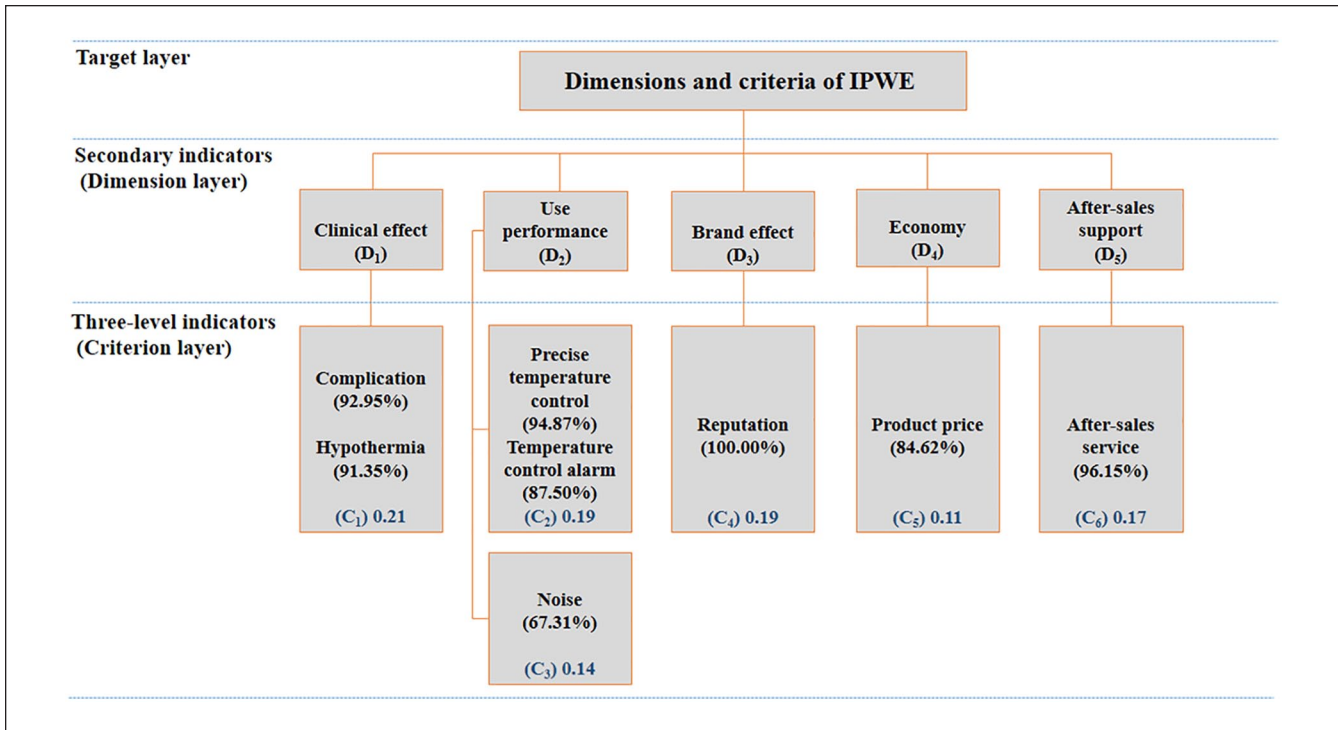


Figure 3. The final choice of dimensions and criteria.

non-clinical departments, the evaluation criteria that received the most votes (30) included brand awareness and after-sales service. Out of the 82 questionnaires, the largest number of votes was given to after-sales service (80). These results show that regardless of being a clinical or a non-clinical department, after-sales service is an important factor in purchasing a patient warming system. Some differences were observed in the votes from anesthesiology and non-clinical departments on other dimensions. To balance the number of non-clinical and clinical departments, the scores were standardized. Next, the experts performed a ranking according to the selection rate of different dimensions and criteria in the questionnaire. Then, combined with their own experience, they selected 10 criteria for the final evaluation. These are required to be mutually exclusive so that the 10 criteria were eventually merged into 5 dimensions and 6 criteria. The weights of the criteria were 0.21, 0.19, 0.14, 0.19, 0.11, and 0.16 (see Figure 3).

Product Selection Results

Of the 33 hospitals involved in the Anhui survey, only 4 have not purchased a patient warming system, while 29 have purchased various types and brands of IPWE, we listed some brands that have not been purchased by any of the hospitals. After the experts performed the selection, 5 brands of insulation products were chosen for the final evaluation. In order to avoid any possibility of commercial partiality, the letters N_1 , N_2 , N_3 , N_4 , and N_5 were used to indicate the names of the manufacturers.

Matrix Scoring Results Based on ILPGCWA

The 2 experts who participated in the evaluation conducted a pair-wise evaluation of 5 different products and filled out 12 judgment matrices based on 6 criteria. Subsequently, the ILPGCWA operator was used to summarize the opinions of the 2 experts, and the aggregate evaluation matrix E_1 to E_6 for the 6 criteria was obtained. Here, we take E_1 as an example (see Table 1).

Then we used the ILPGCWA operator to summarize the relationship between different criteria.

The aggregation results of 5 products were as follows:

$$\beta_{ij} = \begin{bmatrix} \langle S_{1.6331}, 0.5888, 0.3216 \rangle \\ \langle S_{1.4919}, 0.6017, 0.3108 \rangle \\ \langle S_{1.4866}, 0.6037, 0.3086 \rangle \\ \langle S_{1.4866}, 0.5987, 0.3090 \rangle \\ \langle S_{2.1058}, 0.3777, 0.6223 \rangle \end{bmatrix}$$

Finally, we used the formula $\beta'_i = \frac{\sum_{j=1}^n \beta'_{ij}}{n}$ to obtain the score for each product. The results are as follows:

$$y = \begin{bmatrix} 1.0347 \\ 0.9629 \\ 0.9627 \\ 0.9587 \\ 0.7954 \end{bmatrix}$$

Table 3. Comprehensive Score Results for Thermal Insulation Products from Various Brands in 10 Hospitals.

Dimension	Criterion (C)	N_1	N_2	N_3	N_4	N_5
Clinical effect	Complication	0.92	0.81	0.66	0.67	0.47
	Hypothermia					
Product performance	Precise temperature control	0.91	0.77	0.75	0.64	0.36
	Temperature control alarm					
Brand effect	Noise	0.88	0.77	0.71	0.72	0.8
	Reputation	0.92	0.81	0.62	0.61	0.6
Economy	Product price	0.76	0.75	0.7	0.73	0.67
After-sales support	After-sales service	0.92	0.76	0.55	0.55	0.2
Weighted total score		0.89	0.78	0.66	0.64	0.49

for different attributes. For example, in this study, the decision makers in medical engineering are no more certain of their judgment than decision makers in anesthesiology. Thus, even if the initial weights are the same initially, EMTP will make adjustments to their weights due to the professional attributes of the decision makers. Regarding this issue, we recommend that hospital managers pay attention to 2 aspects in the future: (1) whether the initial weights of different stakeholders who evaluate different products are scientifically allocated and (2) whether the professional level of the evaluator is appropriate for the knowledge requirements.

In the multi-attribute group decision-making problem, the aggregation of evaluation information is a critical step. The existing aggregation operators only act in a discrete state, fail to reflect their continuity and cannot well simulate the compensation behavior that exists in the process of human aggregation. The ILPGCWA operator uses 2 parameters to control the degree of compensation in the process of aggregation to decrease the gap between theoretical and experimental results and achieve continuity in the process of aggregation. The minimum compensation degree represents a pessimistic view, and the largest degree of compensation represents an optimistic view. In the evaluation process, as the degree of compensation related to these parameter values increases, the actual evaluation results will also change accordingly. By studying the influence of the combination of different parameters λ and p on aggregation, we can observe in Table 2 that when $\lambda \rightarrow 1$, that is, when experts are neutral on the evaluation action as a whole, no matter what value p takes, their degree of preference for products has little effect on the ranking results of the best products. This means the judgment of the best product N_1 is basically the same as the judgment result of the KRM method, and when $\lambda \rightarrow \infty$, the change of the preference parameter p of the product is that it will have a significant impact on the final selection result of the product. To a certain extent, this can guide the hospital's purchasing decision-makers: if the evaluator is more active in the evaluation process, then attention should be paid to controlling the degree of preference for products from different departments or stakeholders before reaching a consensus, and when the decision makers are less

active in the evaluation activities themselves, they may not be able to play the role of various stakeholders in the decision-making goals. At this time, leaders need to communicate with them in time to improve the hospital's procurement decision-making results.

We used the KRM method as a verification tool to test the effectiveness of the EMTP calculations. The disadvantage of swing weighting is that it relies on experts giving independent, accurate scores for each scheme. In this case, it was possible to set difficulty barriers for different evaluation participants, thus making the scoring more difficult and time-consuming.⁴⁷ In this application, in addition to the number of experts required to set the target layer of the solution, KRM requires more evaluators to ensure the stability of the results, which also makes the work of verification take longer, about 2 weeks in this case. There are only 2 EMTP participants in this evaluation, and the entire evaluation only takes about 1 hour. It does not require professional skills and data processing technology to achieve the same effect as KRM, thereby improving the efficiency of decision-making and selection.

Conclusions

To meet the requirements of hospital managers in different regions for purchasing decision-making, we established a new decision-making method, EMTP. This method, combined with MCDA and Mini-HTA, can take the hesitancy and uncertainty of the evaluators into account, and uses their opinions as evaluation information to select the optimal products. At the same time, the ILPGCWA operator provides 2 additional parameters called λ and p to reflect compensation capability in the decision process. This provides an objective, efficient, and rapid method for making hospital medical procurement decisions, and can make the evaluation process more flexible and robust. We applied this method in a large Chinese hospital's purchasing decision for IPWE and the plausibility of the results were verified by comparison with KRM. The results show that the preferred product selected by EMTP is consistent with the preferred product selected by KRM within a certain range of parameters. We

discussed the uncertainty and personal preferences of the experts. The psychological disposition of the experts, influenced by hospital policies or other factors, can be expressed more accurately through membership and parameters, thus affecting the weighting and ranking of the results. Therefore, to obtain accurate results, it may be necessary to fully consider the specific environment of decision-making, understand the field where the expert is located, carefully determine the initial weights, and select the appropriate parameters in combination with the dynamic weights that have been adjusted by the experts. We should take into consideration the expert's personal tendencies and professional knowledge before decision-making, to balance the degree of compensation of collective decision-making parameters. Of course, the relevant professional knowledge training of the evaluators should be strengthened wherever possible. Furthermore, the Mini-HTA tool as a guiding framework can be replaced by other industry guidelines. This structured approach has the potential for application in many industries that demand group decision-making in the presence of ambiguous data.

One limitation of this study is that although we demonstrated the effectiveness of our method by comparing it against a verification method, we only evaluated 1 medical technique. Today's medical equipment committees also need to incorporate broader standards into their assessments. However, there were only 2 decision makers using the framework. Although the results obtained here are not significantly different from those of the verification method, this does not represent the opinions of other hospital decision makers. In future work, we will use the framework in more hospital decision-making departments to further improve the representation of judgments and opinions.

Future studies should continue this line of research bearing in mind the following aspects: first, further work should develop more objective and dynamic weighting methods to determine the index weights corresponding to each candidate brand based on current data and accumulated knowledge. Second, candidate brand sets that can be adapted dynamically should be considered, so as to enable consideration of the changing composition of new and old products. Third, to meet the requirements of individual consistency and group consensus during the evaluation process, the individual wishes of the evaluator were ignored when adjusting for their personal views. Therefore, in future research, a feedback mechanism that considers the personal wishes of the evaluator should be studied, so that the adjusted evaluation viewpoint not only considers the consistency and consensus requirements, but also allows for the personal wishes of the evaluator, that is, whether the evaluator accepts the adjustment proposal.

Author Contributions

Substantially contributed to conception or design: CY, YJW, XXH, FCL, DW, and XQC; Contributed to acquisition, analysis,

or interpretation of data: CY, YJW, LTQ, YJC, WG, QL, and XQC; Drafted the manuscript for important content: CY, YJW, and QL; Critically revised the manuscript for important intellectual content: CY, XXH, FCL, WG, DW, and XQC; Gave final approval: All authors.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the National Natural Science Foundation (91846107); Key research and development projects in Anhui province (1804h08020286).

ORCID iD

Xiaoxuan Hu  <https://orcid.org/0000-0002-2579-1385>

Supplemental Material

Supplemental material for this article is available online.

References

1. Durbach IN, Stewart TJ. Modeling uncertainty in multi-criteria decision analysis. *Eur J Oper Res*. 2012;223(1):1-14. doi:10.1016/j.ejor.2012.04.038
2. Retèl VP, Grutters JPC, van Harten WH, Joore MA. Value of research and value of development in early assessments of new medical technologies. *Value Health*. 2013;16(5):720-728. doi:10.1016/j.jval.2013.04.013
3. Barberá-Tomás D, Consoli D. Whatever works: uncertainty and technological hybrids in medical innovation. *Technol Forecast Soc Change*. 2012;79(5):932-948. doi:10.1016/j.techfore.2011.12.009
4. Seo HJ, Park JJ, Lee SH. A systematic review on current status of health technology reassessment: insights for South Korea. *Health Res Policy Syst*. 2016;14(1):82. doi:10.1186/s12961-016-0152-x
5. Surgey G, Chalkidou K, Reuben W, Suleman F, Miot J, Hofman K. Introducing health technology assessment in Tanzania. *Int J Technol Assess Health Care*. 2019;36(2):80-86. doi:10.1017/s0266462319000588
6. Govender M, Mueller DB, Basu D. Purchasing of medical equipment in public hospitals: the mini-HTA tool. *S Afr Med J*. 2011;101(11):807-808.
7. Kidholm K, Ehlers L, Korsbek L, Kjærby R, Beck M. Assessment of the quality of mini-HTA. *Int J Technol Assess Health Care*. 2009;25(1):42-48. doi:10.1017/s0266462309090060
8. Banta D, Kristensen FB, Jonsson E. A history of health technology assessment at the European level. *Int J Technol Assess Health Care*. 2009;25(S1):68-73. doi:10.1017/s0266462309090448
9. Teerawattananon Y, Teo YY, Dabak S, Rattanavipapong W, et al. Tackling the 3 big challenges confronting health technology assessment development in Asia: a commentary.

- Value Health Reg Issues.* 2020;21:66-68. doi:10.1016/j.vhri.2019.07.001
10. Oortwijn W, van Oosterhout S, Kapiriri L. Application of evidence-informed deliberative processes in health technology assessment in low- and middle-income countries. *Int J Technol Assess Health Care.* 2020;36(4):440-444. doi:10.1017/s0266462320000549
 11. Baltussen R, Niessen LW. Priority setting of health interventions: the need for multi-criteria decision analysis. *Cost Eff Resour Alloc.* 2006;4(1):1-9. doi:10.2139/ssrn.943814
 12. Cleemput I, Franken M, Koopmanschap M, le Polain M. European drug reimbursement systems' legitimacy: five-country comparison and policy tool. *Int J Technol Assess Health Care.* 2012;28(4):358-366. doi:10.1017/s0266462312000529
 13. Sloane EB. Using a decision support system tool for health-care technology assessments. *IEEE Eng Med Biol Mag.* 2004;23(3):42-55. doi:10.1109/memb.2004.1317981
 14. Goetghebeur MM, Wagner M, Khoury H, Levitt RJ, Erickson LJ, Rindress D. Bridging health technology assessment (HTA) and efficient health care decision making with multicriteria decision analysis (MCDA). *Med Decis Making.* 2011;32(2):376-388. doi:10.1177/0272989x11416870
 15. Martelli N, Hansen P, van den Brink H, et al. Combining multicriteria decision analysis and mini-health technology assessment: a funding decision-support tool for medical devices in a university hospital setting. *J Biomed Inform.* 2016;59:201-208. doi:10.1016/j.jbi.2015.12.002
 16. Marsh K, Lanitis T, Neasham D, Orfanos P, Caro J. Assessing the value of healthcare interventions using multi-criteria decision analysis: a review of the literature. *Pharmacoeconomics.* 2014;32(4):345-365. doi:10.1007/s40273-014-0135-0
 17. Belton V, Stewart T. *Multiple Criteria Decision Analysis: An Integrated Approach.* Springer Science & Business Media; 2002.
 18. Köksalan MM, Wallenius J, Zionts S. *Multiple Criteria Decision Making from Early History to the 21st Century.* World Scientific; 2011.
 19. Gitinavard H, Zarandi MHF. A mixed expert evaluation system and dynamic interval-valued hesitant fuzzy selection approach. *Int J Math Comput Sci.* 2016;10(7):341-349. doi:10.5281/zenodo.1125079
 20. Gitinavard H, Mousavi SM. A new balancing and ranking method based on hesitant fuzzy sets for solving decision-making problems under uncertainty. *Int J Eng.* 2015;28(2):214-223. doi:10.5829/idosi.ije.2015.28.02b.07
 21. Ebrahimnezhad S, Gitinavard H, Sohrabvandi S. A new extended analytical hierarchy process technique with incomplete interval-valued information for risk assessment in IT outsourcing. *Int J Eng.* 2017;30(5):739-748
 22. Dong JY, Chen Y, Wan SP. A cosine similarity based QUALIFLEX approach with hesitant fuzzy linguistic term sets for financial performance evaluation. *Appl Soft Comput.* 2018;69:316-329. doi:10.1016/j.asoc.2018.04.053
 23. Wan SP, Yan J, Zou WC, Dong JY. Generalized Shapley Choquet integral operator based method for interactive interval-valued hesitant fuzzy uncertain linguistic multi-criteria group decision making. *IEEE Access.* 2020;8:202194-202215. doi:10.1109/access.2020.3034107
 24. Thokala P, Duenas A. Multiple criteria decision analysis for health technology assessment. *Value Health.* 2012;15(8):1172-1181. doi:10.1016/j.jval.2012.06.015
 25. Dolan JG. Multi-criteria clinical decision support: a primer on the use of multiple criteria decision making methods to promote evidence-based, patient-centered healthcare. *Patient.* 2010;3(4):229-248. doi:10.2165/11539470-000000000-00000
 26. Marsh K, Caro JJ, Muszbek N. Does the future belong to MCDA? *ISPOR Connections* 2012;18(6):9-11.
 27. Atanassov KT. Intuitionistic fuzzy sets. *Fuzzy Sets Syst.* 1986;20(1):87-96. doi:10.1016/s0165-0114(86)80034-3
 28. Gitinavard H, Akbarpour Shirazi M. An extended intuitionistic fuzzy modified group complex proportional assessment approach. *Ind Syst Eng.* 2018;11(3):229-246.
 29. Hamzeh AM, Mousavi SM, Gitinavard H. Imprecise earned duration model for time evaluation of construction projects with risk considerations. *Automat Constr.* 2020;111:102993. doi:10.1016/j.autcon.2019.102993
 30. Lan J, Chen Y, Ning M, Wang Z. A new linguistic aggregation operator and its application to multiple attribute decision making. *Oper Res Perspect.* 2015;2:156-164. doi:10.1016/j.orp.2015.09.001.
 31. Wang JQ, Li HB. Multi-criteria decision-making method based on aggregation operators for intuitionistic linguistic fuzzy numbers. *Control Decis.* 2010;25(10):1571-1574.
 32. Dong JY, Wan SP. Arithmetic aggregation operators for interval-valued intuitionistic linguistic variables and application to multi-attribute group decision making. *Iran J Fuzzy Syst.* 2016;13(1):1-23. doi:10.22111/ijfs.2016.2284
 33. Wang L, Wang Y, Sangaiah AK, Liao B. Intuitionistic linguistic group decision-making methods based on generalized compensative weighted averaging aggregation operators. *Soft Comput.* 2017;22(22):7605-7617. doi:10.1007/s00500-017-2734-0.
 34. The First Affiliated Hospital of USTC. Published 2020. Accessed June, 2020. http://www.ahslyy.com.cn/cn/About/index_87.html
 35. National Institute for Health and Care Excellence (NICE). Accessed June, 2020. Hypothermia: prevention and management in adults having surgery. 2008. Updated 2016. <https://www.nice.org.uk/Guidance/CG65>
 36. Herlehy AM. Influencing safe perioperative practice through collaboration. *AORN J.* 2011;94(3):217-218. doi:10.1016/j.aorn.2011.07.008
 37. Sessler DI. Perioperative thermoregulation and heat balance. *Lancet.* 2016;387(10038):2655-2664. doi:10.1016/s0140-6736(15)00981-2
 38. AdHopHTA. Published 2012. Accessed June, 2020. <http://www.adhophta.eu/>
 39. Zadeh LA. The concept of a linguistic variable and its application to approximate reasoning—I. *Inf Sci.* 1975;8(3):199-249. doi:10.1016/0020-0255(75)90036-5
 40. Keeney RL, Raiffa H. *Decisions with Multiple Objectives: Preferences and Value Tradeoffs.* Wiley; 1976.
 41. Thokala P, Devlin N, Marsh K, et al. Multiple criteria decision analysis for health care decision making—an introduction: report 1 of the ISPOR MCDA emerging good practices task force. *Value Health.* 2016;19(1):1-13. doi:10.1016/j.jval.2015.12.003
 42. Marsh K, Ijzerman M, Thokala P, et al. Multiple criteria decision analysis for health care decision making—emerging good

- practices: report 2 of the ISPOR MCDA emerging good practices task force. *Value Health*. 2016;19(2):125-137. doi:10.1016/j.jval.2015.12.016
43. Communities and Local Government (CLG). Multi-criteria analysis: a manual. Published 2009. Accessed October, 2020. https://www.researchgate.net/publication/30529921_Multi-Criteria_Analysis_A_Manual
 44. Keeney RL, Raiffa H, Rajala DW. Decisions with multiple objectives: preferences and value trade-offs. *IEEE Trans Syst Man Cybern*. 1979;9(7):403. doi:10.1109/tsmc.1979.4310245
 45. Dieter GE, Schmidt CL. *Engineering Design*. 4th ed. McGraw-Hill Higher Education; 2009.
 46. Wang XF, Wang JQ, Yang WE. Multi-criteria group decision making method based on intuitionistic linguistic aggregation operators. *J Intell Fuzzy Syst*. 2014;26(1):115-125. doi:10.3233/ifs-120719
 47. Radaelli G, Lettieri E, Masella C, et al. Implementation of EUnetHTA core Model® in Lombardia: the VTS framework. *Int J Technol Assess Health Care*. 2014;30(1):105-112. doi:10.1017/s0266462313000639