The Use of Delphi Method and Analytical Hierarchy Process in the Establishment of Assessment Tools in Premature Ejaculation: The Scoring System for Premature Ejaculation Treatment Outcomes

Jianzhong Zhang[1], Taoli Han[2], Zhonglin Cai[1], Ying Wang[2], Xuejun Shang[3], Bin Yang[4], and Hongjun Li[1] (iD)

Abstract

There is a lack of objective tools to comprehensively evaluate premature ejaculation (PE) treatment results clinically. We aimed to describe the development of a novel scoring system for PE treatment results as an example of using the Delphi method and an analytical hierarchy process for complex decision-making in the field of sexual medicine. A Delphi question survey was adopted to collect expert opinions from 47 Chinese specialists in andrology/urology on the assessment of PE treatment outcomes based on four primary properties, that is, the improvement in intravaginal ejaculation latency time, a couple's mental status, the ability to control ejaculation, and sexual intercourse satisfaction. Different weights on those primary properties were assigned to create a mathematical hierarchy matrix and then perform an analysis. The scores were assigned according to the calculated weights. The ratio among the combined weights of the four primary properties was 1:3:2:3. The total numerical score was 36. Scores above 27, between 26 and 18, and below 17 indicated significant improvement, moderate improvement, and no improvement in PE, respectively, with selected treatments. The scoring system with 36 points can be used by physicians, patients, and their sexual partners to comprehensively and objectively assess quantitative PE treatment results.

Introduction

Premature ejaculation (PE) is a common male sexual dysfunction with the prevalence estimates from 4% to 39% according to the EAU guidelines (https://uroweb.org/guideline/sexual-and-reproductive-health/). Currently, there are three questionnaires for the clinical diagnosis of PE, for example, the Premature Ejaculation Diagnostic Tool (PEDT) (Symonds et al., 2007), the Arabic Index of Premature Ejaculation (AIPE) (Arafa & Shamloul, 2007), and the Chinese index of sexual function for premature ejaculation (CIPE) (Cai et al., 2014). These questionnaires are used for the diagnosis of PE based on the self-assessment of patients with complaints of PE using their own (subjective) opinion. Currently, clinical measurement by a

stopwatch and/or subjective estimation of the intravaginal ejaculation latency time (IELT) is commonly used to

[1]Department of Urology, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences, Beijing, China
[2]National Institute for Viral Disease Control and Prevention, Chinese Center for Disease Control and Prevention, China
[3]Department of Andrology, Jinling Hospital Affiliated to Southern Medical University, Nanjing, China
[4]Department of Urology, Affiliated Hospital of the Qingdao University, Qingdao, China

Corresponding Author:
Hongjun Li, Department of Urology, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences, No.1 Shuaifuyuan, Beijing 100730, China.
Email: lihongjun@pumch.cn

evaluate the outcome of PE drug treatment together with improvement in sexual satisfaction reported by patients and their sexual partners. There are few methods for the comprehensive assessment of PE treatment results that also include reports from sexual partners (Serefoglu et al., 2011a, 2011b). Based on the current subjective self-assessment tools, a clinical evaluation may lead to inadequate drug approvals or treatment regimes. Regarding the quality of the data, a comprehensive score-based system might contribute to an improved evaluation of the drug treatment results of PE and is urgently needed.

The Delphi method systematically collects the opinions of participants and/or experts on a certain topic, especially for topics with many subjective properties. It is a structured interactive communication method originally developed as a systematic and interactive forecasting method that relies on the opinion of a panel of experts (Boulkedid et al., 2011; Jorm, 2015). The Delphi method has been adapted for the development of scoring or indexing systems, especially for events that lack objective measures or that depend on experts' opinions (Cai et al., 2014). The method provides a powerful likelihood of avoiding systemic errors during a decision-making process for a complex issue or event.

According to the definition from the International Society for Sexual Medicine (ISSM) for PE, which was published in 2014, PE included three major features: (1) an ejaculation that occurs always or nearly always before or within 1 minute of vaginal contact (lifelong PE) or a ejaculation that occurs within less than 3 minutes (acquired PE); (2) uncontrollable ejaculation in every or almost every vaginal intercourse after insertion; and (3) psychological effects associated with PE such as anxiety, distress, depression, and/or avoidance of sexual intimacy (Althof et al., 2014; Segraves, 2010). The definition of PE includes many subjective properties, and the Delphi method is suitable for this condition to determine the properties involved in the scoring system. Additionally, an analytical hierarchy process (AHP) was applied to determine the weights of the properties in the scoring system for the assessment of PE treatment outcomes.

Materials and Methods

Procedure of Development of the Scoring System for PE Treatment Outcomes

Two facilitators were assigned. Their tasks were as follows: (1) reviewing and analyzing relevant literature of PE treatments and measures within the last 10 years, understanding essential components in the process of the development of a scoring system, listing key factors and possible problems during the process, and selecting properties for PE treatment assessment; (2) designing and

completing a survey questionnaire for the Delphi survey; (3) selecting and inviting experts across mainland China; (4) collecting and summarizing the results from each round of questions; and (5) gathering all information, creating the database, and performing statistical analysis accordingly.

The search strategy of the literature related to PE diagnosis and treatment was facilitated with two key elements, keywords and database. The keywords and combinations were used such as "Premature Ejaculation (PE)," "Acquired PE (APE)," "Lifelong PE (LPE)," "Ejaculatory Dysfunction (ED)," "Intravaginal Ejaculation Latency Time (IELT)," "Treatment," "Orgasm," "Anxiety," "Depression," "Distress," "Satisfaction Degree," "Systematic Review," "Guidelines," "Practice Guidelines," "Consensus," "Standards," and "Clinical Research." The database included Medline, Embase, Web of Science, and Cnki (Chinese database). Publications including reviews, guidelines, consensus, and clinical studies (published before February 2017) were reviewed and analyzed. The papers were defined irrelevant and excluded if it was (1) lack of information on the evaluation of treatment outcomes of PE; (2) redundant reports; (3) published before 2007; and (4) non-Chinese or non-English literature because of the limit of language command. The searching and selection processes were listed in Supplemental Figure S1.

Experts in mainland China specializing in urology and andrology were invited to participate in the study. All of them originated from centralized hospitals in different metropolitan regions and have had substantial experience and knowledge of the diagnosis and clinical treatment of PE. The inclusion criteria of the experts enrolled in the Delphi survey were: (1) they are physicians and clinical scientists; (2) they have clinical experiences in the relevant andrology/urology field for more than 10 years; (3) they have professional titles above the intermediate levels, indicating that their expertise is recognized in the diagnosis and treatment of male sexual dysfunction (especially premature ejaculation); (4) they are doing clinical and basic researches actively and have publications related; and (5) they are willing to participate in this study and have their commitment to complete two rounds of survey independently. The one who was not met with all criteria above was not invited to the study. None of the experts had withdrawn during two rounds of the Delphi survey process.

The core group consisted of seven experts in the field who had a meeting to finalize the design of the Delphi survey. The primary and secondary properties entered into the questionnaire and rounds of questions for the survey were based on information provided by the facilitators. The first round of questions was initiated in April 2017 and

Table 1. Definition of Weight of Properties.

Weight of property	Definition
1	Two elements bear equal importance when comparing these two
3	Element 1 bears more importance than element 2 when comparing these two
5	Element 1 bears more obviously importance than elements 2 and 3 when comparing these three
7	Element 1 bears much more importance than elements 2, 3 and 4 when comparing these four
9	Element 1 is the most important among the five elements when comparing them
2, 4, 6, 8	Intermediate value for judgement
Reciprocal	If the importance of element i and element j is A_{ij}, then the ratio of importance of element j and element i is $A_{ji} = 1/A_{ij}$

the second round in July 2017. The two rounds of Delphi survey questionnaires were sent to individual participants by emails in sequential order and required to complete answers to the questions independently. This ensures that enrolled experts were not influenced by others' opinions. The results collected from the two rounds of questions and the preliminary analysis of results showed that there was no disagreement among all 47 experts, which met with pre-determined numbers of rounds for the Delphi method. After the two rounds, the survey was finished.

A Hierarchy Modeling and Analytical Process

Based on the collected experts' opinions, a hierarchy was modeled. The priorities were established among the properties of the hierarchy through the establishment of a series of judgments by pairwise comparisons of two properties. In addition, the weights of the properties were defined (see Table 1 for weight scale). For instance, a set of four primary properties of PE, that is, the improvement of IELT, the couple's mental status, the ability to control ejaculation, and sexual intercourse satisfaction reported by both patients and their sexual partners, were assigned into A1, A2, A3, and A4 to yield a set of overall priorities for the hierarchy and to perform pairwise comparisons for each property (Figure 1).

The hierarchy matrix that we created was examined mathematically. Formula 1 was used to calculate the consistency index (CI) for the consistency test. The average random consistency index (RI) was obtained from Table 2, and formula 2 was used to calculate the consistency ratio (CR). If CR < 0.10, the consistency of the hierarchy matrix is acceptable. If CR > 0.10, the matrix must be revised. In addition to the consistency test for different properties, the consistency test was also performed for the overall properties.

Data Analysis

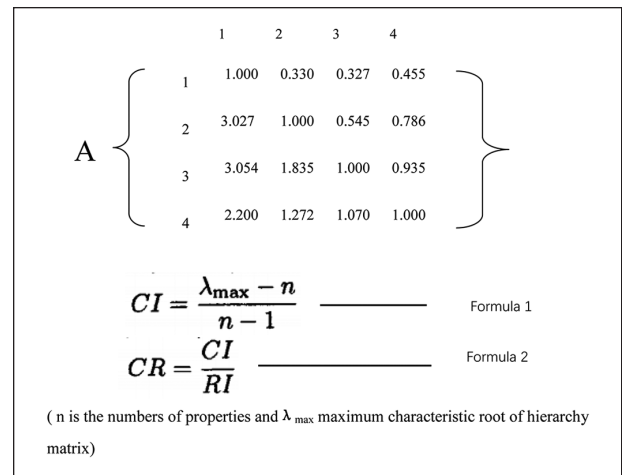Based on the Delphi principle, judgment from experts is pivotal for the development of a scoring system for the

$$A \begin{Bmatrix} & 1 & 2 & 3 & 4 \\ 1 & 1.000 & 0.330 & 0.327 & 0.455 \\ 2 & 3.027 & 1.000 & 0.545 & 0.786 \\ 3 & 3.054 & 1.835 & 1.000 & 0.935 \\ 4 & 2.200 & 1.272 & 1.070 & 1.000 \end{Bmatrix}$$

$$CI = \frac{\lambda_{max} - n}{n - 1} \qquad \text{Formula 1}$$

$$CR = \frac{CI}{RI} \qquad \text{Formula 2}$$

(n is the numbers of properties and λ_{max} maximum characteristic root of hierarchy matrix)

Figure 1. The hierarchy matrix for pairwise comparisons of primary properties of PE.

Table 2. The Average Random Consistency Index.

N	1	2	3	4	5	6	7
RI	0	0	0.52	0.89	1.12	1.24	1.36

assessment of PE treatment outcomes. Therefore, basic information was collected from the experts in the field and analyzed descriptively, including age, agenda, knowledge of their specialty, and years of practice, to provide confident links between their clinical experiences, knowledge of PE, and feedback opinions from the Delphi surveys. The reliability and objectivity of the final results, which were derived from the analytical hierarchy process, depended on four factors.

(1) The positive constant of experts was determined by the recovery rate of questionnaires and the completion of answers to each question. This constant reflects the degree of their attention to and understanding of the study.

(2) The concentration tendency of experts' opinions on the weights of primary and other properties.

Table 3. The Weights Assigned to the Experts' Factors on Premature Ejaculation.

Judging criteria	Influence on expert judgment			Academic Title	Weight	Familiarity Level	Weight	Working years	Weight
	High	Moderate	Low						
Theoretical analysis	0.3	0.2	0.1	Senior	1	Very	1	≥30	1.0
Clinical experiences	0.5	0.4	0.3	Assoc-senior	0.8	Better	0.8	20–29	0.8
Peer's recognition	0.1	0.1	0.1			Moderate	0.6	11–19	0.6
Perceptions	0.1	0.1	0.1			General	0.4	≤10	0.4
						Slightly	0.2		

This tendency was expressed as the average, median, and frequency of full scores.

(3) The harmonization tendency of experts' opinions represents the discrepancy of the experts' judgment for each property. The harmony coefficient was used for the harmonization tendency and was expressed by the Kendall rank coefficient.

(4) The clinical authority of the experts was expressed as the authorization coefficient. The authorization coefficient was determined by judgment and levels of comprehensiveness/familiarity with selected properties of PE from each expert. Weights assigned to the experts' factors on PE are listed in Table 3.

Generally, judgment precision is positively correlated with the clinical authority of experts, although the questionnaire included a criterion of self-evaluation on the authority as a reference. The authority is equal to the following: (judgment coefficient + comprehensive/familiar levels)/2.

The experts' rating was based on the importance of indicators or items in the two rounds of rating tables. The rating criteria of selected indicators were as follows: "not relevant at all, "not important," "acceptable," "relatively important," and "extremely important," which converted to the points of 1, 2, 3, 4, and 5, respectively. The arithmetic mean, *SD*, and coefficient of variation of rating points were calculated for each indicator or item. If the rating average is high and the coefficient of variation is small, it indicates that the indicator or item has good applicability and weighs more in the evaluation system. The arithmetic means of the importance scores of indicators and items required greater than 70% of the full scale of marks (3.5 points) and the coefficient of variation less than 0.5 in the two rounds of the Delphi method.

The data were collected, and the datasheets were created in Microsoft Excel. The weights, the maximum characteristic roots of the hierarchy matrix, the CI, and the CR were calculated and entered for the analytic hierarchy process. SPSS 22.0 software was used for hierarchy and statistical analysis after distribution and consistency tests were performed on the data. A significance level was set at $p < .05$.

Results

Literature Review and Property Selection

The diagram of the literature selection process was listed in Supplemental Figure S1. Assessment indicators were extracted, and the following aspects were included.

(1) Patients reported their global impression of change (PGIC), which represents their overall improvement in sexual life quality after full consideration of the side effects of treatment. The PGIC divides both the therapeutic effect ("significant effect," "effective," "slightly effective," and "no change or deterioration") and side effects ("none," "light," "medium," and "heavy") into four grades, and the final efficacy index score is the efficacy score divided by the side effect score.

(2) The improvement in IELT. Most studies evaluate the therapeutic effect of PE by directly comparing the numerical changes in IELT before and after treatment.

(3) The improvement in the number of vaginal insertions. Some patients with short IELT are more impressed with the number of insertions. The number of insertions can also reflect the length of IELT.

(4) The improvement in control over ejaculation (PCOE). PCOE is one of the most common indicators to describe the efficacy of PE treatment. Some scales of PE also include the evaluation of PCOE, such as premature ejaculation profile (PEP), index of premature ejaculation (IPE), and CIPE scales.

(5) The improvement in the mental states of patients, including anxiety, depression, or distress, caused by PE. Some studies directly evaluated the therapeutic effect through the improvement of personal depression related to ejaculation and interpersonal difficulties. Other studies use psychological questionnaires to evaluate anxiety and

Table 4. Draft of the Scale for the Evaluation of PE.

Primary properties	Secondary properties
1. Improvement of duration of sexual intercourse	1.1 Improvement of IELT
	1.2 Improvement of the times of vaginal insertions
2. Improvement of the couples' mental status	2.1 Improvement of the distress, anxiety, and depression caused by PE
	2.2 Improvement of the relations between the couples
	2.3 Improvement of the female mental status.
3. Improvement of the ability to control ejaculation	3.1 Improvement in the ability to prolong the ejaculation time
	3.2 Lower prevalence of ejaculation before the patients want to
4. Improvement of the sexual intercourse satisfaction	4.1 Improvement of the patient's self-satisfaction
	4.2 Improvement of the female sexual partner's satisfaction

IELT = intravaginal ejaculation latency time; PE = premature ejaculation.

depression, such as the Self-rating Depression Scale, the Patient Health Questionnaire-9, the Self-rating Anxiety Scale, and the Generalized Anxiety Disorder scale.

(6) The improvement in sexual life satisfaction from the perspective of both the patients and their female sexual partners.

(7) The improvement in the scores of PE-related questionnaires, including the CIPE5, the IPE, the PEDT, the AIPE, and the PEP. In addition to the indicators mentioned above, these questionnaires also cover erectile function, sexual desire, female orgasm, and other aspects.

(8) The improvement in the results of the electro-physiological examination. These examinations are not recommended as a routine examination in the current guidelines and are used only in some studies related to penile electrophysiology. Due to the lack of application, the number of relevant studies is limited.

The above properties were evaluated and selected according to the principles of being numeric, comprehensive, comparable, real time, and dynamic, as well as according to a combination of qualitative and quantitative analysis. Four primary properties and nine secondary properties were selected, and a preliminary discussion draft of the evaluation system for the treatment effect of patients with PE was made (Table 4).

The Delphi Method

Among the 47 experts, the median age was 48 (range 34–65). Males accounted for 97.9% of the experts, and females accounted for 2.1%. Senior urologists and associated senior urologists accounted for 72.3% and 27.7% of the experts, respectively. The years of clinical practice in andrology/urology were as follows: <10 years: 2.13%; 10–20 years: 23.4%; 20–30 years: 46.81%; and 30–40 years: 27.66%. Regarding judgments and opinions of PE

properties given by the experts, 86.3% of experts completed the questions based on their clinical experience and knowledge, while 3.9% had opinions based on their individual perceptions.

The recovery rates for the two rounds of questions were both 100%, and each question was completed by the participants so that the positive constant of experts was 1. The average judgment coefficient was 0.773, the comprehensive/familiar coefficient was 0.817, and the authority of experts was 0.795. The harmony coefficient of experts was expressed using the Kendall coefficient. The Kendall rank coefficient for the primary and secondary properties of PE was 0.335 and 0.355 ($p < .001$ and $p < .003$), respectively, indicating that experts' recognition of the structure and rationale of PE properties was highly accordant, that is, with primary and secondary properties bearing different weights.

In the first round of the Delphi method, secondary properties of "1.2: Improvement in the frequency of vaginal insertion" and "3.2: Lower prevalence of ejaculation before the patients want to" were deleted because these two indicators are at a low level in rank and importance. In the second round of the Delphi method, all the experts agreed with the revised list of properties, and hierarchy analysis was further performed. The primary property "1: Improvement in the duration of sexual intercourse" was replaced by "1.1: Improvement in IELT", and the secondary property "3.1: Improvement in the ability to prolong the ejaculation time" was deleted because the primary property "3: Improvement in the ability to control ejaculation" was repeated.

Hierarchy Analysis

The calculated λ_{max} and CR were 4.064 and 0.024, respectively. The consistency of the hierarchy matrix was approved (CR < 0.1). A series of consistency tests generated corresponding weights of primary and secondary properties, which were converted into scores for the assessment of PE treatment outcomes (Table 5).

Table 5. Properties and Its Weights in the Scoring System for Assessment on PE Treatment Outcomes.

Primary property	Secondary property (Weights)	Combined weights	
Improvement of IELT		0.108	
Improvement of the couples' mental status	Improvement of the distress, anxiety and depression caused by PE (0.219)	0.066	0.301
	Improvement of the relations between the couples	0.115	
	Improvement of the female mental status.	0.120	
Improvement of the ability to control ejaculation		0.244	
Improvement of sexual intercourse satisfaction	Patient's self-satisfaction (0.654)	0.226	0.346
	Female sexual partner's satisfaction (0.346)	0.120	

IELT = intravaginal ejaculation latency time; PE = premature ejaculation.

The Assigned Scores

The scores were assigned based on the calculated weights of all properties. The score ratio of weights from four primary properties was 1:3:2:3 (Table 4). If a basic score of 4 was given to the IELT, the scores for the improvement in the couple's mental status, the ability to control ejaculation, and sexual intercourse satisfaction would be 12, 8, and 12, respectively. Three secondary properties under the couple's improved relationship had a similar ratio of the weights such that 4 scores were given to each, while 8 and 4 scores were given to patients' self-reported satisfaction and partner-reported satisfaction with improvement in sexual intercourse, respectively, following the same rule. The total score was 36 (Table 6). Scores above 27 indicate significant improvement, scores between 26 and 18 indicate moderate improvement, and scores below 17 indicate no improvement in PE by selected treatments. The patient himself will give scores on improvements in IELT, the couples' relationship, and the ability to control ejaculation, while the patient and his sexual partner would provide scores on the improvement in their satisfaction with sexual intercourse.

Discussion

PE has been traditionally recognized as a disorder of the complex sensory pathway, including delay in the transmission of ejaculation neural stimuli in the central nervous system and penile sensitivity in the peripheral nervous system. The etiology of PE can include many factors (Gao et al., 2014; Laumann et al., 1999; Porst et al., 2007; Serefoglu et al., 2011a, 2011b), and organic disorders in the central or peripheral nervous system can only explain the etiology of a small number of patients with PE. Subjective properties such as psychorelational factors have received increasing attention from scholars in recent years and are thought to be the main or unique determinants of PE. Zhang et al. evaluated the clinical features of patients with PE in a

recent cross-sectional study. The results indicated that PE can be multifactorial and is associated with several psychosocial factors. In addition, couples' attitudes toward PE are of great importance and can influence whether patients with PE seek medical treatment (Zhang et al., 2019). The subjective aspects of PE are important for the etiology of PE, and several scholars have even proposed in recent years an evolutionary theory that PE is a symptom rather than a kind of disease (Palmieri et al., 2012; Puppo & Puppo, 2016).

Based on the current data, there were many indicators utilized in the evaluation of PE treatment effects, including both objective indicators (such as IELT) and subjective indicators (such as degree of satisfaction and ability to control ejaculation). One of the most commonly used indicators is the change in IELT (an objective indicator). Currently, various physicians prefer to use the change in IELT as a single parameter to describe the efficacy of the treatment (Bar-Or et al., 2012). However, for some patients with PE, IELT has been significantly improved after treatment, but their satisfaction is still very low. Simple prolongation of intercourse time does not always mean improvement in the ability to control ejaculation and relief of all distress derived from sexual activity for both patients and their sexual partners (McMahon, 2014; Symonds et al., 2003). From the clinical experience of the current expert panel, when physicians evaluate the treatment effects of PE, patient satisfaction and ability to control ejaculation are equally or more important than IELT. Under this circumstance, these subjective indicators should be focused on the evaluation of PE treatments.

Sexual dysfunction, including erectile dysfunction and PE, may be related to the female sexual partner; thus, the shared decision-making by the physician, the patient himself, and his sexual partner is critical for better treatment and prognosis (Burnett et al., 2018; Li et al., 2016). Both patients and their sexual partners were considered important for the assessment of improvements in sexual intercourse satisfaction.

Table 6. The Scoring System for Assessment of PE treatment outcomes.

PART 1. IELT (4 points)				
No change	Slight increase	Moderate	Significant	Complete
0	1	2	3	4

PART 2. Improvement of the couples' mental status (12 points)				
1. Improvement of the distress, anxiety and depression caused by PE				
No change	Slight better	Moderate	Significant	Complete
0	1	2	3	4
2. Improvement of the relations between the couples				
No change	Slight better	Moderate	Significant	Complete
0	1	2	3	4
3. Improvement of the female mental status				
No change	Slight better	Moderate	Significant	Complete
0	1	2	3	4

PART 3. Improvement of the ability to control ejaculation (8 points)				
No change	Slight better	Moderate	Significant	Complete
0	2	4	6	8

PART 4. Improvement of sexual intercourse satisfaction (12 points)				
1. Patient's self-satisfaction				
No change	Slight better	Moderate	Significant	Complete
0	2	4	6	8
2. Female sexual partner's satisfaction				
No change	Slight better	Moderate	Significant	Complete
0	1	2	3	4

IELT = intravaginal ejaculation latency time; PE = premature ejaculation.

There is a lack of systemic and standard measures for the clinical efficacy of PE treatment, which may frustrate both physicians and patients (Simões Paço & Jorge Pereira, 2016). In the current study, four primary properties were enrolled, including the improvement in intravaginal ejaculation latency time, the couple's mental status, the ability to control ejaculation, and sexual intercourse satisfaction. These four properties have evaluated the treatment effect both from the subjective aspects (couple's mental status, the ability to control ejaculation, and sexual intercourse satisfaction) and the objective aspect (IELT) comprehensively. The scoring system has synergized possible subjective and objective indicators in the evaluation of PE treatment, making the current system less biased.

The weight of each property was calculated by the current evaluation system. Notably, the weight ranking results showed that the most commonly used property, IELT, has a lower weight (weight: 0.108) and is only half of the weight of the primary property "improvement in the ability to control ejaculation" (weight: 0.244) and "improvement in sexual intercourse satisfaction" (weight: 0.226). Most of the experts who participated in the consultation thought that improvements in the ability to control ejaculation and sexual intercourse satisfaction were more important. As a common disease of andrology, PE is a psychological

disorder. It is very important that men's direct feelings can affect the outcomes of the treatment, and the weight of these two indicators is larger and reasonable. In addition, patients and their female partners were included in the evaluation system to evaluate the improvement in sexual life satisfaction. A comprehensive assessment of the situation of female partners can increase the objectivity of the assessment system. The multidimensional PE efficacy evaluation index can meet the requirements of clinicians, patients, and their female partners at the same time.

Evaluation of PE treatment outcomes is mainly based on experts' opinions about the reliability and precision of selected properties and their weights in the scoring system. Among 47 Chinese experts who participated in this study, senior urologists accounted for 72%, and 98% of them had over a decade of clinical experience in andrology. Two rounds of Delphi question surveys and a hierarchy analysis were used to estimate the consistent tendency of experts' opinions and measure the authority of participants (Baiocchi et al., 2019; Bardach et al., 2018; Boulias et al., 2018). The results demonstrated that the authority coefficient and the harmony coefficient of the experts on selected primary and secondary properties for the assessment of PE treatment were both high, indicating that experts' voting on the rationales

and weights of PE properties was highly accordant, that is, primary and secondary properties having different levels of importance in the scoring system. The analysis also showed a consistency ratio of 0.024 (far below 0.10), demonstrating that the hierarchy matrix and model used for this study were confidently fitted and that the selected properties and their weights were statistically acceptable.

A novelty of the current study was that the Delphi method combined with AHP was applied for the first time in sexual medicine as an example to develop an assessment tool for PE treatment.

This study was mainly focused on method development and its process as an example of clinical tool development in sexual medicine. The scoring system for the assessment of PE treatment outcomes has not been applied clinically. Clinical experiences and the criteria for PE treatment assessment were based on a single ethnic group of Chinese experts. However, this numerical scoring system is well applicable to all PE patients for their treatment assessment globally since the selection of the properties and indicators used in the study was based on literature mostly from international journals and database rather than Chinese ones. The relatively high weights on couple-related properties might reflect the prevalent cultural values and recognition of PE and marriage. The generalization of the scoring system still needs to be verified clinically. During two rounds of Delphi question surveys, we observed that some experts' opinions were influenced by others who were usually at the senior level and that some experts were reluctant to change their initial opinions on the weights of primary or secondary properties likely influenced by self-esteem based on their ranks. These were limitations of this study. A clinical study is needed to verify and/or amend the scoring system for the assessment of PE treatment outcomes.

Declaration of Conflicting Interests

Funding

Ethical approval obtained

This study was approved by the Ethics Committee of Peking Union Medical College Hospital (approval no. S214).

ORCID iD

Hongjun Li 🔟 https://orcid.org/0000-0001-6040-5251

Supplemental Material

Supplemental material for this article is available online.

References

Althof, S. E., McMahon, C. G., Waldinger, M. D., Serefoglu, E. C., Shindel, A. W., Adaikan, P. G., Becher, E., Dean, J., Giuliano, F., Hellstrom, W. J., Giraldi, A., Glina, S., Incrocci, L., Jannini, E., McCabe, M., Parish, S., Rowland, D., Segraves, R. T., Sharlip, I., & Torres, L. O. (2014). An update of the international society of sexual medicine's guidelines for the diagnosis and treatment of Premature Ejaculation (PE). *Sexual Medicine*, *2*(2), 60–90. https://doi.org/10.1002/sm2.28

Arafa, M., & Shamloul, R. (2007). Development and evaluation of the Arabic Index of Premature Ejaculation (AIPE). *The Journal of Sexual Medicine*, *4*(6), 1750–1756. https://doi.org/10.1111/j.1743-6109.2006.00213.x

Baiocchi, G. L., Giacopuzzi, S., Marrelli, D., Reim, D., Piessen, G., Matos da Costa, P., Reynolds, J. V., Meyer, H. J., Morgagni, P., Gockel, I., Lara Santos, L., Jensen, L. S., Murphy, T., Preston, S. R., Ter-Ovanesov, M., Fumagalli Romario, U., Degiuli, M., Kielan, W., Mönig, S., . . . De Manzoni, G. (2019). International consensus on a complications list after gastrectomy for cancer. *Gastric Cancer: Official Journal of the International Gastric Cancer Association and the Japanese Gastric Cancer Association*, *22*(1), 172–189. https://doi.org/10.1007/s10120-018-0839-5

Bardach, N. S., Burkhart, Q., Richardson, L. P., Roth, C. P., Murphy, J. M., Parast, L., Gidengil, C. A., Marmet, J., Britto, M. T., & Mangione-Smith, R. (2018). Hospital-based quality measures for pediatric mental health care. *Pediatrics*, *141*(6), e20173554. https://doi.org/10.1542/peds.2017-3554

Bar-Or, D., Salottolo, K. M., Orlando, A., Winkler, J. V. & Tramadol ODT Study Group. (2012). A randomized double-blind, placebo-controlled multicenter study to evaluate the efficacy and safety of two doses of the tramadol orally disintegrating tablet for the treatment of premature ejaculation within less than 2 minutes. *European Urology*, *61*(4), 736–743. https://doi.org/10.1016/j.eururo.2011.08.039

Boulias, C., Ismail, F., Phadke, C. P., Bagg, S., Bureau, I., Charest, S., Chen, R., Cheng, A., Ethans, K., Fink, M., Finlayson, H., Gulasingam, S., Guo, M., Haziza, M., Hosseini, H., Khan, O., Lang, M., Lapp, T., Leckey, R., . . . Winston, P. (2018). A Delphi-based consensus statement on the management of anticoagulated patients with

botulinum toxin for limb spasticity. *Archives of Physical Medicine and Rehabilitation*, *99*(11), 2183–2189. https://doi.org/10.1016/j.apmr.2018.04.023

Boulkedid, R., Abdoul, H., Loustau, M., Sibony, O., & Alberti, C. (2011). Using and reporting the Delphi method for selecting healthcare quality indicators: A systematic review. *PloS One*, *6*(6), e20476. https://doi.org/10.1371/journal.pone.0020476

Burnett, A. L., Nehra, A., Breau, R. H., Culkin, D. J., Faraday, M. M., Hakim, L. S., Heidelbaugh, J., Khera, M., McVary, K. T., Miner, M. M., Nelson, C. J., Sadeghi-Nejad, H., Seftel, A. D., & Shindel, A. W. (2018). Erectile dysfunction: AUA guideline. *The Journal of Urology*, *200*(3), 633–641. https://doi.org/10.1016/j.juro.2018.05.004

Cai, Y., Wang, T., & Liu, J. H. (2014). Questionnaires on male sexual dysfunction and their clinical application. *Zhonghua Nan Ke Xue*, *20*(9), 840–845.

Gao, J., Xu, C., Liang, C., Su, P., Peng, Z., Shi, K., Tang, D., Gao, P., Lu, Z., Liu, J., Xia, L., Yang, J., Hao, Z., Zhou, J., & Zhang, X. (2014). Relationships between intravaginal ejaculatory latency time and national institutes of health-chronic prostatitis symptom index in the four types of premature ejaculation syndromes: A large observational study in China. *The Journal of Sexual Medicine*, *11*(12), 3093–3101. https://doi.org/10.1111/jsm.12696

Jorm, A. F. (2015). Using the Delphi expert consensus method in mental health research. *The Australian and New Zealand Journal of Psychiatry*, *49*(10), 887–897. https://doi.org/10.1177/0004867415600891

Laumann, E. O., Paik, A., & Rosen, R. C. (1999). Sexual dysfunction in the United States: Prevalence and predictors. *JAMA*, *281*(6), 537–544. https://doi.org/10.1001/jama.281.6.537

Li, H., Gao, T., & Wang, R. (2016). The role of the sexual partner in managing erectile dysfunction. *Nature Reviews Urology*, *13*(3), 168–177. https://doi.org/10.1038/nrurol.2015.315

McMahon, C. G. (2014). Management of ejaculatory dysfunction. *Internal Medicine Journal*, *44*(2), 124–131. https://doi.org/10.1111/imj.12344

Palmieri, A., Imbimbo, C., Longo, N., Fusco, F., Verze, P., Maletta, A., Creta, M., & Mirone, V. (2012). Ejaculatory abstinence influences intravaginal ejaculatory latency time: Results from a prospective randomized trial. *Urologia Internationalis*, *88*(4), 459–462. https://doi.org/10.1159/000337057

Porst, H., Montorsi, F., Rosen, R. C., Gaynor, L., Grupe, S., & Alexander, J. (2007). The Premature Ejaculation Prevalence and Attitudes (PEPA) survey: prevalence, comorbidities, and professional help-seeking. *European Urology*, *51*(3), 816–824. https://doi.org/10.1016/j.eururo.2006.07.004

Puppo, V., & Puppo, G. (2016). Comprehensive review of the anatomy and physiology of male ejaculation: Premature ejaculation is not a disease. *Clinical Anatomy (New York, N.Y.)*, *29*(1), 111–119. https://doi.org/10.1002/ca.22655

Segraves, R. T. (2010). Considerations for an evidence-based definition of premature ejaculation in the DSM-V. *The Journal of Sexual Medicine*, *7*(2 Pt 1), 672–679. doi:10.1111/j.1743-6109.2009.01682.x.

Serefoglu, E. C., Yaman, O., Cayan, S., Asci, R., Orhan, I., Usta, M. F., Ekmekcioglu, O., Kendirci, M., Semerci, B., & Kadioglu, A. (2011). The comparison of premature ejaculation assessment questionnaires and their sensitivity for the four premature ejaculation syndromes: Results from the Turkish society of andrology sexual health survey. *The Journal of Sexual Medicine*, *8*(4), 1177–1185. https://doi.org/10.1111/j.1743-6109.2010.02183.x

Serefoglu, E. C., Yaman, O., Cayan, S., Asci, R., Orhan, I., Usta, M. F., Ekmekcioglu, O., Kendirci, M., Semerci, B., & Kadioglu, A. (2011). Prevalence of the complaint of ejaculating prematurely and the four premature ejaculation syndromes: Results from the Turkish Society of Andrology Sexual Health Survey. *The Journal of Sexual Medicine*, *8*(2), 540–548. https://doi.org/10.1111/j.1743-6109.2010.02095.x

Simões Paço, J., & Jorge Pereira, B. (2016). New therapeutic perspectives in premature ejaculation. *Urology*, *88*, 87–92. https://doi.org/10.1016/j.urology.2015.11.003

Symonds, T., Perelman, M., Althof, S., Giuliano, F., Martin, M., Abraham, L., Crossland, A., Morris, M., & May, K. (2007). Further evidence of the reliability and validity of the premature ejaculation diagnostic tool. *International Journal of Impotence Research*, *19*(5), 521–525. https://doi.org/10.1038/sj.ijir.3901567

Symonds, T., Roblin, D., Hart, K., & Althof, S. (2003). How does premature ejaculation impact a man s life? *Journal of Sex & Marital Therapy*, *29*(5), 361–370. https://doi.org/10.1080/00926230390224738

Zhang, J., Li, F., Li, H., Zhang, Z., Yang, B., & Li, H. (2019). Clinical features of and couple's attitudes towards premature ejaculation: a multicenter cross-sectional study. *Aging Male*. https://doi.org/10.1080/13685538.2019.1640194