Patient Preference and Adherence

Dovepress

open access to scientific and medical research

Open Access Full Text Article

ORIGINAL RESEARCH

Comparing the Self-Reported Acceptability of Discrete Choice Experiment and Best-Worst Scaling: An Empirical Study in Patients with Type 2 Diabetes Mellitus

Fuming Li[1,2], Shimeng Liu[1,2], Yuanyuan Gu[3], Shunping Li[4], Ying Tao[1,2], Yan Wei[1,2], Yingyao Chen[1,2]

[1]School of Public Health, Fudan University, Shanghai, People's Republic of China; [2]National Health Commission Key Laboratory of Health Technology Assessment (Fudan University), Shanghai, People's Republic of China; [3]Macquarie University Centre for the Health Economy, Macquarie Business School & Australian Institute of Health Innovation, Macquarie University, Macquarie Park, Macquarie Park, New South Wales, Australia; [4]Centre for Health Management and Policy Research, School of Public Health, Cheeloo College of Medicine, Shandong University, Jinan, Shandong, People's Republic of China

Correspondence: Shimeng Liu; Yingyao Chen, School of Public Health, Fudan University, 130 Dongan Road, Xuhui, Shanghai, 200032, People's Republic of China, Tel +86 198 2182 4471; +86 135 6450 8981, Email smliu19@fudan.edu.cn; yychen@shmu.edu.cn

Purpose: Discrete choice experiment (DCE) and profile case (case 2) best-worst scaling (BWS) present uncertainties regarding the acceptability of quantifying individual healthcare preferences, which may adversely affect the validity of responses and impede the reflection of true healthcare preferences. This study aimed to assess the acceptability of these two methods from the perspective of patients with type 2 diabetes mellitus (T2DM) and examine their association with specific characteristics of the target population.

Patients and Methods: This cross-sectional study was based on a nationally representative survey; data were collected using a multistage stratified cluster-sampling procedure between September 2021 and January 2022. Eligible adults with confirmed T2DM voluntarily participated in this study. Participants completed both the DCE and case 2 BWS (BWS-2) choice tasks in random order and provided self-reported assessments of acceptability, including task completion difficulty, comprehension of task complexity, and response preference. Logistic regression and random forest models were used to identify variables associated with acceptability.

Results: In total, 3286 patients with T2DM were included in the study. Respondents indicated there was no statistically significant difference in completion difficulty between the DCE and BWS-2, although the DCE scores were slightly higher (3.07 ± 0.68 vs 3.03 ± 0.67, P = 0.06). However, 1979 (60.2%) respondents found the DCE easier to comprehend. No significant preferences were observed between the two methods (1638 (49.8%) vs 1648 (50.2%)). Sociodemographic factors, such as residence, monthly out-of-pocket costs, and illness duration were significantly associated with comprehension complexity and response preference.

Conclusion: This study yielded contrasting results to most of previous studies, suggesting that DCE may be less cognitively demanding and more suitable for patients with T2DM from the perspective of self-reported acceptability of DCE and BWS. This study promotes a focus on patient acceptability in quantifying individual healthcare preferences to inform tailored optimal stated-preference method for a target population.

Plain Language Summary

- Stated preference methodologies such as the discrete choice experiment (DCE) and case 2 best-worst scaling (BWS-2) are gaining popularity as methods for quantifying individual preferences in healthcare. However, the acceptability of the two methods to participants must be considered in practice to reduce cognitive burden and ensure the validity of preference elicitation.
- DCE was perceived to be less cognitively burdensome than BWS-2. In contrast to patients who thought that DCE was more acceptable, BWS-2 was more accepted by rural patients, patients who lived with the disease for a longer period, and those who had lower monthly out-of-pocket costs.
- These findings demonstrate potential differences in the acceptability of DCE and BWS-2 for patients with type 2 diabetes mellitus. To improve efficiency, it would be useful for researchers to consider the optimal stated preference method for identifying target populations according to sociodemographic and disease-related characteristics.

Received: 23 March 2024
Accepted: 16 August 2024
Published: 30 August 2024

Keywords: patient-reported outcomes, acceptability, discrete choice experiment, best-worst scaling, preference, type 2 diabetes

Introduction

Incorporating patients' values, preferences and unmet medical needs to inform decision making may ultimately result in clinical care, health technology lifecycle and reimbursement decisions that better reflect the broader public preferences on disease-specific management.[1–3] This has led to a growing interest in measuring and quantifying patient and public preferences in healthcare through various stated-preference methods, particularly discrete choice experiment (DCE) and best-worst scaling (BWS).[4–6] DCE and BWS were both grounded in random utility theory, which posit that the utility of service attributes could be decomposed into an explainable or systematic component and a random component, and that the random unobservable component exists because of the inherent variability within and between individuals.[7,8] Within a DCE, respondents are asked to select between a set of alternative profiles, each characterized by several attributes and their levels; respondents' choices subsequently determine how preferences are influenced by each attribute level, as well as their relative importance.[9,10] BWS could elicit additional information on the least preferred option, with three variants allowing respondents to identify the best and the worst attributes (object case, case 1), attribute levels (profile case, case 2), or combinations of multiple attribute levels (multi-profile, case 3).[8,11] To understand the relative ranking or prioritization of the content levels of attributes, case 2 BWS (BWS-2) appears to be the preferred presentation, as respondents do not require to consider the value of the profile as a whole.[12,13]

According to the random utility theory, patients' healthcare-related preferences are elicited by capturing their intentions expressed in hypothetical situations. When faced with multiple trade-offs between two or more alternatives, patients often require sufficient time and knowledge to comprehend complex information and carefully consider their personal preferences. This becomes particularly crucial when contemplating uncertain future scenarios and outcome states they have never experienced before.[14] The excessive cognitive burden imposed by decision fatigue undermines the assumption of utility maximization and leads to inconsistent and poorly considered choices, particularly among individuals with lower health literacy.[15–17] The concept of acceptability is closely related to participants' rational responses within the stated-preference methods, which requires a clear understanding of both the choice context and choice task. This entails the participants being willing to engage in and capable of completing the task. Failure to meet this critical requirement for eliciting preferences may adversely affect the validity of the responses and impede the reflection of true healthcare preferences.[18]

Limited empirical research has been conducted to determine the acceptability of both methods for eliciting healthcare preferences, with existing studies being inconclusive. Flynn et al conducted the first comparison of BWS-2 with DCE and found that whilst the vast majority of (older) respondents provided usable data from the BWS-2 task, only around one half did so for the DCE.[12] Rogers et al demonstrated a better understanding of and ability to complete BWS-2 tasks for valuing health states compared to DCE.[19] The cognitive burden of BWS-2 tasks may be lower as individuals only need to focus on one set of attribute levels in each choice task, as opposed to multiple in DCE.[20] However, a systematic review suggested that respondents may favor DCE, as it was perceived to have lower self-reported task difficulty and was preferred over BWS-2 in a priority setting.[13] Studies by Himmler et al (valuing quality of life measures among older people) and Soekhai et al (treatment preference for patients with neuromuscular diseases) also confirmed that DCE is less cognitively burdensome than BWS-2.[21,22] Furthermore, another study suggested that there may not be a significant difference in completion rates between DCE and BWS-2.[4] The aforementioned studies serve as a valuable reminder that the acceptability of stated preference methods may vary considerably depending on the specific decision context, and selecting the ideal method to capture a patient perspective is essential for optimal therapeutic decision making and resource allocation in health or community-based care services.

Type 2 diabetes mellitus (T2DM) is a common chronic disease and one of the greatest public health challenges to an aging society. T2DM has a high global prevalence and is rising across all regions.[23] As of 2021, China has the largest number of patients with T2DM; this number is anticipated to exceed 257 million by 2045, which will pose unprecedented challenges to the country's healthcare and social care system.[24] Medications if required could control blood glucose levels in patients with T2DM, thereby delaying disease progression or complications.[23] The American Diabetes Association recommends that a patient-centered approach should be used to guide the choice of pharmacologic agents as well, considering multidimensional factors associated

with patients' preferences.[25] The present study was designed within the context of preference elicitation for second-line anti-hyperglycemic medicine among patients with T2DM to specifically document stated-preference methods for assessing acceptability from the perspectives of patients with T2DM. These treatments were chosen because the diverse properties of glucose-lowering agents can guide individualized treatment choices for patients with T2DM during long-term disease self-management, further influencing their confidence and adherence to treatment.[26] Additionally, patients with T2DM may be accompanied by a measure of age-related cognitive decline,[27] and it is important for them to reduce the cognitive burden and ensure validity of medication preference elicitation. Hence, we aim to achieve two objectives: (1) to describe the difference between perceived acceptability of DCE compared to BWS-2 in patients with T2DM; and (2) to identify factors related to sociodemographic characteristics, health status, and disease treatment that may influence the acceptability of these preference elicitation methods.

Material and Methods

Survey Design and Participants

We used a multistage stratified cluster-sampling procedure that considered the geographical region and economic development status for data collection between September 2021 and January 2022. In stage one, we selected six provinces and municipalities that represented the socioeconomic status and lifestyles of five major geographical regions in China. In stage two, we selected a provincial capital city (developed city) and one or two non-provincial capital cities (underdeveloped cities) from each geographical region in China. In stage three, we randomly selected one or two hospitals and two or three primary-care institutions from each city. The inclusion criteria were as follows: (1) age ≥18 years, (2) confirmed T2DM, and (3) voluntary participation. Following the rule of thumb proposed by Johnson and Orme, we settled on a minimum sample size of 250.[28,29] The questionnaire for patients with T2DM comprised three parts: (1) patient sociodemographic characteristics and information related to health status and disease treatment; (2) six DCE choice tasks, and six BWS-2 choice tasks; and (3) self-reported acceptability of the two methods.

We followed recommended guidelines proposed by the International Society for Pharmacoeconomics and Outcomes Research for DCE and BWS study design and analysis.[5] Regarding attributes development, an initial list of attributes was obtained through a literature review, and the seven attributes included in the final survey were identified through multistage expert focus-group discussions and a case 1 BWS experiment, including treatment efficacy (reduction in HbA1c), hypoglycemia events risk, gastrointestinal events risk, weight change in 6 months, cardiovascular benefits, route of administration, and monthly out-of-pocket costs; the levels were further developed based on the clinical practice of T2DM in China. The final set of DCE and BWS-2 questions included in the survey was generated using D-optimal algorithm by SAS to construct a fractional factorial experimental design. The resulting design included 48 unique DCE and BWS-2 questions, assembled into eight blocks of six tasks each. The final experimental design was evaluated for level balance, minimal overlap, and orthogonality, as shown in Figure S1 (examples of the DCE and BWS-2 choice tasks). Each DCE task offered respondents a choice between two hypothetical medication profiles; patients were asked which medication they would prefer, and each BWS task asked respondents to choose the attribute that was most or least important to them in one hypothetical medication profile. In the DCE, a strictly dominant option (alternative A is preferred over alternative B) is included to evaluate the internal validity of the data. In the BWS, the validity task included only three attribute levels: 0, 100, and 600 Chinese Yuan. Respondents were expected to select the first option as the best feature and the last option as the worst. Details on the attributes and levels of development, as well as the experimental design of medications, are available elsewhere.[30]

After completing the choice tasks for DCE and BWS-2, the respondents self-reported the acceptability of the methods. The conceptual framework of acceptability outlines three key variables: completion difficulty, comprehension complexity and response preference. A 5-point Likert scale (1 = very simple, and 5 = very difficult) was used to rate the difficulty of completing a series of choice tasks associated with DCE and BWS-2. For comprehension complexity and response preference, the respondents were asked to make a trade-off between the two methods and consider which was easier to comprehend and which they preferred.

Data Collection

We collected data face-to-face to ensure the quality of data collection. The DCE and BWS-2 instructions were explained in detail to the investigators, who received training from the research team. Cognitive pre-tests were administered from June to July 2021 to ensure the validity and comprehensibility of the survey instrument. Participants (n=12) were provided with pre-test questions and given the option to verbally share comments for revision. During the formal survey (September 2021 to January 2022), participants completed the questionnaire anonymously after providing written informed consent. Each participant was randomly assigned six DCE and six BWS-2 choice tasks in one of the eight blocks. Participants who had difficulty filling in the questionnaire, such as those with impaired eyesight or were illiteracy, were given the option to complete the questionnaire verbally. The completion of the questionnaire took approximately 10–20 minutes, and each participant was reimbursed for their time. All completed questionnaires were returned directly to the investigators, and patients' identities and information were strictly protected by the researchers.

Statistical Analysis

Patient-specific characteristics and the tasks' completion difficulties were assessed and reported as the mean ± the standard deviation. The difficulties were compared using the Wilcoxon rank sum test. All continuous variables were then transformed into categorical variables and presented as frequencies and percentages. Cohen's kappa was employed to assess intra-rater reliability in consideration of potential variations in completion difficulty among individual respondents, as well as variations between understanding complexity and response preference. To compare categorical variables between the DCE and BWS-2 groups, univariate analysis was conducted using the chi-square test or the Cochran-Mantel-Haenszel tests.

To further examine the potential factors associated with the acceptability of the eliciting preference methods, multivariate logistic regression analyses were used based on statistically significant variables in the chi-square tests, and odds ratios (ORs) and corresponding 95% confidence intervals (CIs) were reported. According to the variance expansion factor (VIF) and tolerance results, there was no evidence of multicollinearity in the logistic regression model (VIF > 2, tolerance < 0.5). The statistical analyses were performed using SAS version 9.4 (SAS Institute Inc., Cary, NC, USA).

To assess the robustness of the multivariate analysis and explore the variable importance more accurately, we used random forest (RF) as a nonparametric machine-learning algorithm to construct classification models for predicting the choice of method and determining variables' importance. We analyzed the full dataset to develop these RF models, and significant variables from the chi-square tests were entered into the models as potential classifiers. The parameters' default settings were used for classification in the algorithm. We used the out-of-bag (OOB) proportion of the data left outside of the algorithm as validation data to compute the classification error, which was ultimately averaged over all trees. We also reported the OOB estimation error. The variable importance analysis considers the mean decrease in accuracy as an evaluation indicator to predict the acceptability of the methods by testing the OOB sample. The RF was implemented using R software version 4.2.2 (R Development Core Team, Vienna, Austria) "random Forest" package.

To validate the robustness of the base-case analysis, we also performed sensitivity analyses to include inconsistent responses in the preference analysis and analyzed the order of method to investigate whether the sequence of completing the DCE or BWS-2 would have any impact on acceptability. The level of statistical significance was set to 0.05 for all statistical analyses.

Results

Respondents' Sociodemographic and Disease-Related Characteristics

In total, 3487 eligible patients with T2DM completed the survey. A total of 201 respondents (5.8%) failed the validity test owing to their choice of an inferior option, leaving 3286 respondents for the final analysis. The self-reported socio-demographic and disease-related characteristics of the respondents are summarized in Table S1. The mean age of the respondents was 61.4 ± 11.9 years, and the median duration of disease was 9.5 ± 7.1 years. Of the respondents, 50.2% were female, 49.5% had a body mass index (BMI) ≥24, and the majority were urban residence (64.2%) and outpatients (69.5%), making the sample nationally representative.[31]

Table 1 Acceptability in Completing DCE and BWS-2 Tasks (n=3286)

Question		DCE No. (%)	BWS-2 No. (%)
Completing difficulty	Very simple	52 (1.6)	70 (2.1)
	Simple	424 (12.9)	423 (12.9)
	Neither simple nor difficult	2137 (65.0)	2186 (66.5)
	Difficult	595 (18.1)	545 (16.6)
	Very difficult	78 (2.4)	62 (1.9)
Easier to comprehend		1979 (60.2)	1307 (39.8)
Preferred to response		1638 (49.9)	1648 (50.1)

Abbreviations: DCE, discrete choice experiment; BWS-2, best-worst scaling case 2.

Description of the Self-Reported Acceptability in DCE and BWS-2

Table 1 shows the responses to the questions regarding the acceptability of the two methods. Respondents indicated there was no statistically significant difference in completion difficulty between DCE and BWS-2, although the DCE scores were slightly higher (3.07 ± 0.68 vs 3.03 ± 0.67, P = 0.06). This indicates that more respondents thought the DCE might be more difficult to complete than the BWS-2. However, regarding comprehension complexity, approximately 60.2% of respondents tended to choose DCE as being easier to comprehend, and there was almost no difference in response preference for the methods. The weighted Cohen's kappa value for intra-rater agreement was 0.62 (95% CI: 0.60–0.65) for the difficulty of tasks, and 0.67 (95% CI: 0.64–0.69) for agreement between comprehension complexity and response preference, which showed the possible disagreement among individual responders regarding their choices (Figure 1).

Factors Associated with Comprehension Complexity

The variable assignments are provided in Table S2. In the univariate analysis, seven variables (residence, annual income, illness duration, BMI, route of administration, monthly out-of-pocket costs, and gastrointestinal events) were statistically significant (P < 0.05) and included in the multivariate model (Table S3). Results displayed in Table 2 indicate that the identifying factors of comprehension complexity were significantly associated with residence (OR=1.46, 95% CI=1.25–1.71), annual income (OR=1.27, 95% CI=1.16–1.39), illness duration (OR=1.30, 95% CI=1.11–1.53), BMI (OR=0.55, 95% CI=0.38–0.81), preferred route of administration (OR=0.79, 95% CI=0.63–0.98), monthly out-of-pocket costs

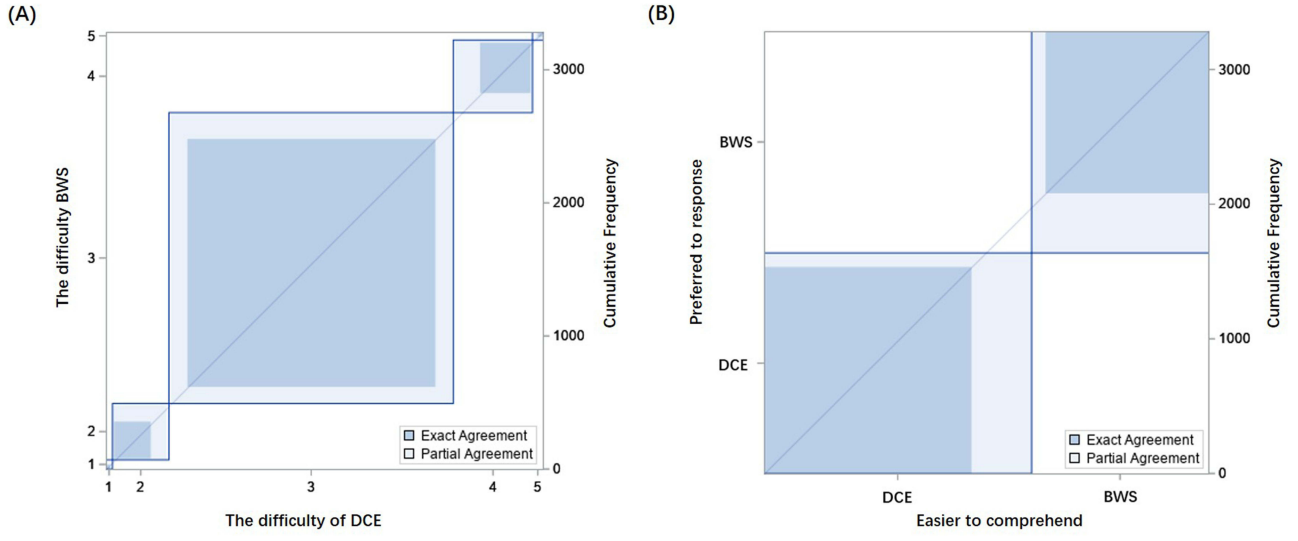

Figure 1 Agreement between the self-reported acceptability of DCE and BWS-2 (N=3286). (**A**) agreement between the difficulty of the two methods, (**B**) agreement between comprehension complexity and response preference of the two methods.
Abbreviations: DCE, discrete choice experiment; BWS-2, case 2 best-worst scaling.

Table 2 Association of Sociodemographic and Disease-Related Factors with Comprehension Complexity Between DCE and BWS-2[a]

Variable		Estimate	*S.E.*	*P* Value	OR (95% CI)
Intercept		−1.42	0.27	< 0.001	
Residence		0.38	0.08	< 0.001	1.46 (1.25–1.71)
Annual income		0.24	0.05	< 0.001	1.27 (1.16–1.39)
Illness duration		0.26	0.08	< 0.01	1.30 (1.11–1.53)
BMI (Ref: ≥24)	<18.5	−0.34	0.13	< 0.01	0.55 (0.38–0.81)
	≥18.5 and <24	0.09	0.07	0.21	0.86 (0.74–0.99)
Preferredroute of administration		−0.24	0.11	0.03	0.79 (0.63–0.98)
Monthly out-of-pocket costs		−0.10	0.03	< 0.001	0.90 (0.86–0.95)
Gastrointestinal events (Ref: Yes)	Unclear	−0.01	0.08	0.88	0.86 (0.66–1.13)
	No	−0.12	0.05	0.02	0.77 (0.66–0.91)

Notes: [a]The reference level for the odds ratios of dichotomous variables is represented by the first group, while for multicategorical variables, the reference level is represented by the last group.
Abbreviations: *S.E.*, standard error; OR, odds ratio; CI, confidence interval; BMI, body mass index.

(OR=0.90, 95% CI=0.86–0.95), and gastrointestinal events (OR=0.77, 95% CI=0.66–0.91). Specifically, respondents who were rural residents, had a higher level of annual income, had a longer disease duration, were overweight or obese, preferred oral administration, had lower monthly out-of-pocket costs, and experienced gastrointestinal events were more likely to consider BWS-2 easier to comprehend; respondents who demonstrated the opposite were more likely to consider DCE as being easier to comprehend.

The RF analysis was used to examine the seven variables associated with comprehension complexity in the univariate analysis. The overall OOB estimate of the error rate for the full dataset was 39.2%. The importance scores are shown in Figure 2(A), and following predictors with positive scores were arranged in descending order of importance: residence, monthly out-of-pocket costs, gastrointestinal events, annual income, and preferred route of administration.

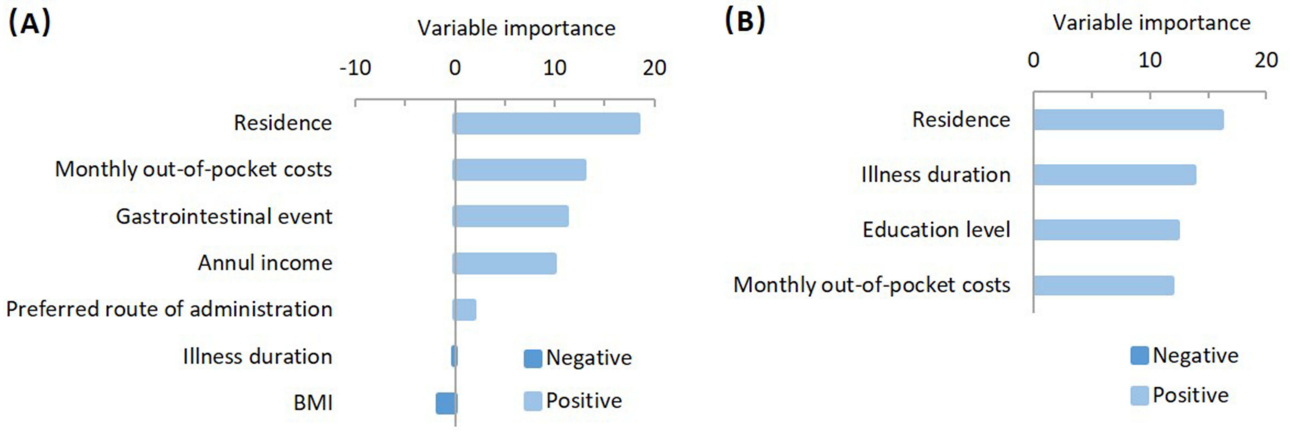

Figure 2 Random forest to classify comprehension complexity and response preference. (**A**) permutation variable importance measures for comprehension complexity among DCE and BWS-2, (**B**) permutation variable importance measures for response preference among DCE and BWS-2.
Notes: Variables with positive importance values increased the accuracy of the random forest algorithm, whereas negative values decreased the accuracy.
Abbreviation: BMI, body mass index.

Table 3 Association of Sociodemographic and Disease-Related Factors with Response Preference Between DCE and BWS-2[a]

Variable	Estimate	*S.E.*	*P* Value	OR (95% CI)
Intercept	−0.42	0.24	0.08	
Residence	0.24	0.08	< 0.01	1.27 (1.09–1.48)
Education level	−0.08	0.06	0.16	0.92 (0.82–1.03)
Illness duration	0.27	0.08	< 0.001	1.31 (1.12–1.52)
Monthly out-of-pocket costs	−0.06	0.02	< 0.01	0.94 (0.90–0.98)

Notes: [a] The reference level for the odds ratios of dichotomous variables is represented by the first group, while for multicategorical variables, the reference level is represented by the last group.
Abbreviations: *S.E.*, standard error; OR, odds ratio; CI, confidence interval.

Factors Associated with Response Preference

Four variables (residence, education level, illness duration, and monthly out-of-pocket costs) in the univariate analysis were included in multivariable modeling (Table S4), which found that respondents with a rural residence (OR=1.27, 95% CI=1.09–1.48), longer disease duration (OR=1.30, 95% CI=1.12–1.52), and less monthly out-of-pocket costs (OR=0.94, 95% CI=0.90–0.98) preferred the BWS-2 over the DCE. No significant association was found between the educational level and response preference (Table 3).

The overall OOB estimate of the error rate for the full dataset in the RF analysis was attenuated to 46.8% using only four important covariates. Figure 2(B) shows that all variables had positive scores and were useful for classifying the response preferences of the methods: residence, illness duration, education level, and monthly out-of-pocket costs.

Sensitivity Analyses

Table S1 displays the sociodemographic and disease-related characteristics of all respondents, including those excluded based on the validity test. These characteristics were not significantly different from the base-case characteristics, and their inclusion in the sensitivity analyses yielded results that were generally consistent with those of the base-case analysis (Tables S5–S9 and Figure S2). Furthermore, this analysis revealed a notable association between fasting blood glucose levels and acceptability of the methods. There was no statistical difference between the DCE and BWS-2 completion difficulty scores (3.06 ± 0.69 vs 3.03 ± 0.68, P = 0.09). Additionally, we discovered that the order of the methods affected completion difficulty, although there was no difference in comprehension complexity or response preference (Table S10).

Discussion

Comparative studies examining the merits of DCE and BWS, have gained significant attention, emphasizing the need for further exploration and understanding of their comparative advantages. This study's objective was to investigate the acceptability of DCE and BWS. Additionally, we aimed to identify the key sociodemographic characteristics and disease-related factors that are most influential in determining the acceptability of these methods by performing a preference study using a nationally representative T2DM sample in China. To our knowledge, there has been limited research examining factors related to the acceptability of stated-preference methods using a large patient sample. In this study, we discovered that DCE and BWS-2 had similar levels of difficulty. Interestingly, the DCE was found to be easier to comprehend, but there was no substantial difference in the response preference between the two methods.

Several studies have addressed acceptability in various ways; our findings were somewhat different from those in previous literature.[4,13,32,33] Specifically, we observed that there was no statistically significant difference in completion difficulty between the DCE and BWS-2, despite the DCE scores being slightly higher. Interestingly, DCE is associated with a relatively simpler cognitive process in terms of comprehension complexity. These seemingly contradictory

findings can be explained by examining the qualitative information provided by the respondents.[34] The DCE task allows for a realistic decision-making scenario involving multiple attributes and levels that requires thorough analysis and decision-making based on a comprehensive framework. However, the BWS task involves selecting the most/least important attribute/level combination that is furthest apart on the latent utility scale, which is a more abstract process and can sometimes be less time-consuming. For researchers, BWS can offer more comprehensive insights into potential utility functions and enhance the statistical efficiency of preference elicitation.[35] However, this may not hold true for respondents involved in decision-making processes. Individuals without prior experience in a specific application area may find it challenging to identify extreme preferences from a set of choices, thereby hindering the accurate reflection of realistic preference weights. The notion of "comprehension" is frequently recognized as a vital aspect in establishing the validity of DCE and making informed decisions. Decreased cognitive burden associated with DCE may lead to a reduction in decision uncertainty. However, it is worth noting that research has demonstrated that "failing" consistency tests does not necessarily imply irrational or uninformed responses, nor does this indicate a lack of understanding of a task. This perspective is further supported by the results of our sensitivity analyses.[18,32]

Our study findings align with Janssen et al's study on patients with T2DM, which indicated no significant difference in response preference between DCE and BWS.[36] In Janssen's study, pre-test interviews revealed that participants did not exhibit a preference for either elicitation method. However, it is important to note that participants in other studies showed distinct preferences for one task over another. For instance, in a preference survey on funding for new health technologies, the Australian public favored the DCE, while children in the United Kingdom preferred the BWS to assess dental caries.[19,32] To explore potential discrepancies and determine whether these preferences are unique to healthcare priority-setting or more widely applicable, further disease-specific studies or replication of our study in different settings are crucial.

Regarding the sociodemographic and T2DM-related factors for the acceptability of DCE and BWS-2, we found that residence, illness duration and monthly out-of-pocket costs were significantly associated with acceptability, while additional factors such as annual income and gastrointestinal events were specifically related to comprehension complexity. Rogers et al considered using BWS as a more suitable method for valuing health states in children, prompting its selection in another study focused on generating value sets for adolescent caries-specific oral health-related quality of life.[19,37] Similarly, another study examining social care outcomes opted for BWS owing to its perceived lower cognitive burden and greater suitability for collecting preference data from service users.[20] Based on our model analysis results, we assigned the DCE to urban residents in the preference survey among patients with T2DM to mitigate their cognitive burden. The logistic regression and RF models complemented each other, enhancing our ability to identify the factors associated with acceptability from various perspectives. Key drivers, such as residence and monthly out-of-pocket costs identified by the RF model can be targeted for further surveys. These results may inform future research on the most suitable method of stated preferences for target T2DM populations with specific sociodemographic characteristics in the context of limited healthcare resources. No significant association was found between education level and acceptability of the methods. However, the RF model highlighted the importance of considering this variable for classification purposes. One plausible explanation is that certain vulnerable populations received assistance from the investigators in completing the DCE and BWS choice tasks. The investigators clarified the meaning of the tasks and provided guidance for eliciting preferences.

This study has some limitations that must be acknowledged. First, because our study involved a face-to-face survey, we did not collect the exact completion times of the DCE and BWS choice tasks, which could be a metric of the acceptability of the two methods. However, completion time alone may not fully capture the acceptability of DCE and BWS, as individuals who spend more time on tasks might have been engaging in more careful trade-offs during their decision-making processes. Additionally, the face-to-face survey method may have introduced some bias, particularly for participants who received assistance in completing tasks. Second, because of the inherent design limitations of cross-sectional research, causal inferences could not be derived. While our study employed a self-report form to investigate patient acceptance of stated-preference methods, which included a large and representative sample obtained through stratified cluster random sampling, it only partially reflected patients' perspectives and did not reveal the underlying reasons behind their choices. To gain a deeper understanding, additional qualitative research, such as interviews, is necessary to explore patients' perspectives. Finally,

caution should be exercised when interpreting results based on self-reporting measures, as they are subject to inherent limitations and potential biases. It is also important for future research to explore the comparative merits of DCE and BWS from additional perspectives, such as by evaluating the concordance of different methods using the generalized multinomial logit model. This would provide more empirical evidence for choosing the most suitable preference elicitation method.

Nevertheless, our study provides valuable insights for identifying optimal healthcare preference methods that adapt to the cognitive abilities of patients with T2DM, informing future studies that focus on patient acceptability in healthcare decision-making. This, in part, allows patients to make a thorough trade-off with limited cognitive capacity, thus plausibly eliciting evidence of healthcare preference, improving the healthcare decision-making process, and promoting patient-centered care.

Conclusion

We conducted an empirical study based on a sample of Chinese patients with T2DM to examine the relative self-reported acceptability of two stated-preference methods, DCE and BWS-2. Specifically, respondents perceived the difficulty of completing the DCE and BWS-2 to be similar, but the DCE was easier for respondents to comprehend than the BWS-2, and no significant differences were found in response preference between the two methods. We also observed that respondents' sociodemographic and disease characteristics partially influenced the acceptability of the methods. This study promotes a focus on patient acceptability in quantifying individual healthcare preferences to inform tailored optimal stated-preference methods for a target population within the context of limited healthcare resources.

Data Sharing Statement

All data generated or analyzed during this study are included in this published article and supplementary information files.

Ethics Approval and Informed Consent

This study was approved by the ethics review board of the School of Public Health, Fudan University (Reference No. IRB# 2021-07-0911), and the research adhered to the tenets of the Declaration of Helsinki. All study participants provided written informed consent.

Author Contributions

All authors made a significant contribution to the work reported, whether that is in the conception, study design, execution, acquisition of data, analysis and interpretation, or in all these areas; took part in drafting, revising or critically reviewing the article; gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agree to be accountable for all aspects of the work.

Funding

This study was funded by the National Natural Science Foundation of China (Grant No. 72074047). The funder had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Disclosure

The authors have no relevant conflicts of interest to disclose for this work.

References

1. Meara A, Crossnohere NL, Bridges JFP. Methods for measuring patient preferences: an update and future directions. *Curr Opin Rheumatol*. 2019;31(2):125–131. doi:10.1097/bor.0000000000000587
2. van Overbeeke E, Forrester V, Simoens S, Huys I. Use of Patient Preferences in Health Technology Assessment: perspectives of Canadian, Belgian and German HTA Representatives. *Patient*. 2021;14(1):119–128. doi:10.1007/s40271-020-00449-0
3. Giusti A, Nkhoma K, Petrus R, et al. The empirical evidence underpinning the concept and practice of person-centred care for serious illness: a systematic review. *Brit Med J Glob Health*. 2020;5(12):e003330. doi:10.1136/bmjgh-2020-003330

4. van Dijk JD, Groothuis-Oudshoorn CG, Marshall DA, IJ MJ. An Empirical Comparison of Discrete Choice Experiment and Best-Worst Scaling to Estimate Stakeholders' Risk Tolerance for Hip Replacement Surgery. *Value Health*. 2016;19(4):316–322. doi:10.1016/j.jval.2015.12.020
5. Bridges JF, Hauber AB, Marshall D, et al. Conjoint analysis applications in health--a checklist: a report of the ISPOR Good Research Practices for Conjoint Analysis Task Force. *Value Health*. 2011;14(4):403–413. doi:10.1016/j.jval.2010.11.013
6. Mühlbacher AC, Zweifel P, Kaczynski A, Johnson FR. Experimental measurement of preferences in health care using best-worst scaling (BWS): theoretical and statistical issues. *Health Econ Rev*. 2016;6(1):5. doi:10.1186/s13561-015-0077-z
7. Lancsar E, Louviere J. Conducting discrete choice experiments to inform healthcare decision making: a user's guide. *Pharmacoeconomics*. 2008;26 (8):661–677. doi:10.2165/00019053-200826080-00004
8. Hollin IL, Paskett J, Schuster ALR, Crossnohere NL, Bridges JFP. Best-Worst Scaling and the Prioritization of Objects in Health: a Systematic Review. *Pharmacoeconomics*. 2022;40(9):883–899. doi:10.1007/s40273-022-01167-1
9. Salloum RG, Shenkman EA, Louviere JJ, Chambers DA. Application of discrete choice experiments to enhance stakeholder engagement as a strategy for advancing implementation: a systematic review. *Implement Sci*. 2017;12(1):140. doi:10.1186/s13012-017-0675-8
10. Soekhai V, de Bekker-Grob EW, Ellis AR, Vass CM. Discrete Choice Experiments in Health Economics: past, Present and Future. *Pharmacoeconomics*. 2019;37(2):201–226. doi:10.1007/s40273-018-0734-2
11. Cheung KL, Wijnen BF, Hollin IL, et al. Using Best-Worst Scaling to Investigate Preferences in Health Care. *Pharmacoeconomics*. 2016;34 (12):1195–1209. doi:10.1007/s40273-016-0429-5
12. Flynn TN. Valuing citizen and patient preferences in health: recent developments in three types of best-worst scaling. *Expert Rev Pharmacoecon Outcomes Res*. 2010;10(3):259–267. doi:10.1586/erp.10.29
13. Whitty JA, Oliveira Gonçalves AS. A Systematic Review Comparing the Acceptability, Validity and Concordance of Discrete Choice Experiments and Best-Worst Scaling for Eliciting Preferences in Healthcare. *Patient*. 2018;11(3):301–317. doi:10.1007/s40271-017-0288-y
14. Elwyn G, Frosch D, Thomson R, et al. Shared decision making: a model for clinical practice. *J Gen Intern Med*. 2012;27(10):1361–1367. doi:10.1007/s11606-012-2077-6
15. Rice T. The behavioral economics of health and health care. *Annu Rev Public Health*. 2013;34(1):431–447. doi:10.1146/annurev-publhealth -031912-114353
16. Chu JN, Sarkar U, Rivadeneira NA, Hiatt RA, Khoong EC. Impact of language preference and health literacy on health information-seeking experiences among a low-income, multilingual cohort. *Patient Educ Couns*. 2022;105(5):1268–1275. doi:10.1016/j.pec.2021.08.028
17. Kang JH. Influences of decision preferences and health literacy on temporomandibular disorder treatment outcome. *BMC Oral Health*. 2022;22 (1):385. doi:10.1186/s12903-022-02420-x
18. Pearce A, Harrison M, Watson V, et al. Respondent Understanding in Discrete Choice Experiments: a Scoping Review. *Patient*. 2021;14(1):17–53. doi:10.1007/s40271-020-00467-y
19. Rogers HJ, Marshman Z, Rodd H, Rowen D. Discrete choice experiments or best-worst scaling? A qualitative study to determine the suitability of preference elicitation tasks in research with children and young people. *J Patient Rep Outcomes*. 2021;5(1):26. doi:10.1186/s41687-021-00302-4
20. Potoglou D, Burge P, Flynn T, et al. Best-worst scaling vs. discrete choice experiments: an empirical comparison using social care data. *Soc Sci Med*. 2011;72(10):1717–1727. doi:10.1016/j.socscimed.2011.03.027
21. Himmler S, Soekhai V, van Exel J, Brouwer W. What works better for preference elicitation among older people? Cognitive burden of discrete choice experiment and case 2 best-worst scaling in an online setting. *J Choice Model*. 2021;38:100265.100265. doi:10.1016/j.jocm.2020.100265
22. Soekhai V, Donkers B, Johansson JV, et al. Comparing Outcomes of a Discrete Choice Experiment and Case 2 Best-Worst Scaling: an Application to Neuromuscular Disease Treatment. *Patient*. 2023;16(3):239–253. doi:10.1007/s40271-023-00615-0
23. International Diabetes Federation. IDF Diabetes Atlas, 10th edn. 2021. Available from: https://www.diabetesatlas.org. Accessed 5, June, 2023.
24. GBD 2021 Diabetes Collaborators. Global, regional, and national burden of diabetes from 1990 to 2021, with projections of prevalence to 2050: a systematic analysis for the Global Burden of Disease Study 2021. *Lancet*. 2023;402(10397):203–234. doi:10.1016/s0140-6736(23)01301-6.
25. Doyle-Delgado K, Chamberlain JJ, Shubrook JH, Skolnik N, Trujillo J. Pharmacologic Approaches to Glycemic Treatment of Type 2 Diabetes: synopsis of the 2020 American Diabetes Association's Standards of Medical Care in Diabetes Clinical Guideline. *Ann Intern Med*. 2020;173 (10):813–821. doi:10.7326/m20-2470
26. Thrasher J. Pharmacologic Management of Type 2 Diabetes Mellitus: available Therapies. *Am J Med*. 2017;130(6s):S4–17. doi:10.1016/j. amjmed.2017.04.004
27. Luo A, Xie Z, Wang Y, et al. Type 2 diabetes mellitus-associated cognitive dysfunction: advances in potential mechanisms and therapies. *Neurosci Biobehav Rev*. 2022;137:104642. doi:10.1016/j.neubiorev.2022.104642
28. Orme B Sample size issues for conjoint analysis studies. *Sequim: Sawtooth Software Technical Paper*. 1998.
29. Johnson R, Orme B Getting the most from CBC. *Sequim: Sawtooth Software Research Paper Series, Sawtooth Software*. 2003:1–7.
30. Liu S, Liu J, Si L, et al. Patient preferences for anti-hyperglycaemic medication for type 2 diabetes mellitus in China: findings from a national survey. *Brit Med J Glob Health*. 2023;8(4):e010942. doi:10.1136/bmjgh-2022-010942
31. Wang L, Peng W, Zhao Z, et al. Prevalence and Treatment of Diabetes in China, 2013–2018. *J Am Med Assoc*. 2021;326(24):2498–2506. doi:10.1001/jama.2021.22208
32. Whitty JA, Ratcliffe J, Chen G, Scuffham PA. Australian Public Preferences for the Funding of New Health Technologies: a Comparison of Discrete Choice and Profile Case Best-Worst Scaling Methods. *Med Decis Making*. 2014;34(5):638–654. doi:10.1177/0272989x14526640
33. Severin F, Schmidtke J, Mühlbacher A, Rogowski WH. Eliciting preferences for priority setting in genetic testing: a pilot study comparing best-worst scaling and discrete-choice experiments. *Eur J Hum Genet*. 2013;21(11):1202–1208. doi:10.1038/ejhg.2013.36
34. Whitty JA, Walker R, Golenko X, Ratcliffe J. A think aloud study comparing the validity and acceptability of discrete choice and best worst scaling methods. *PLoS One*. 2014;9(4):e90635. doi:10.1371/journal.pone.0090635
35. Netten A, Burge P, Malley J, et al. Outcomes of social care for adults: developing a preference-weighted measure. *Health Technol Assess*. 2012;16 (16):1–166. doi:10.3310/hta16160
36. Janssen EM, Segal JB, Bridges JF. A Framework for Instrument Development of a Choice Experiment: an Application to Type 2 Diabetes. *Patient*. 2016;9(5):465–479. doi:10.1007/s40271-016-0170-3
37. Rogers HJ, Sagabiel J, Marshman Z, Rodd HD, Rowen D. Adolescent valuation of CARIES-QC-U: a child-centred preference-based measure of dental caries. *Health Qual Life Outcomes*. 2022;20(1):18. doi:10.1186/s12955-022-01918-w