Enhancing Women's Health Literacy through Artificial Intelligence Chatbots: A Discrete Choice Experiment Study

Jing Wang, Hewei Min, Tao Li, Jiaheng Li, Yang Jiang, Yibo Wu, Xinying Sun

Table of Contents

Enhancing Women's Health Literacy through Artificial Intelligence Chatbots: A Discrete Choice Experiment Study

Jing Wang[1]; Hewei Min[1]; Tao Li[2]; Jiaheng Li[3]; Yang Jiang[4]; Yibo Wu[1]; Xinying Sun[1]

[1]School of Public Health, Peking University, Beijing, China Beijing CN
[2]Department of gynaecology and obstetrics, The Fourth Central Hospital of Baoding City, Hebei, China; Hebei CN
[3]School of Basic Medical Sciences, Hebei University, Hebei, Baoding, China; Hebei CN
[4]North China University of Science and Technology, Hebei,Tangshan, China; Hebei CN

Corresponding Author:
Xinying Sun
School of Public Health, Peking University, Beijing, China
38 Xueyuan Road, Haidian District, Beijing, China
Beijing
CN

Abstract

Background: Over 96% of adult women face health issues, with 70% experiencing conditions like infections. Mobile health education is increasingly popular but faces challenges in personalization and readability. Artificial intelligence chatbots provide tailored support, and a Discrete Choice Experiment can help in understanding user preferences to improve chatbot design.

Objective: This study explores the preferences of women towards artificial intelligence chatbots to improve health education communication, user experience, and health literacy.

Methods: A discrete choice experiment was conducted, identifying six main attributes of artificial intelligence chatbots: response accuracy, legibility, service cost, background information collection, information utility, and content provision. 957 participants from a hospital in Hebei Province participated, choosing between two hypothetical chatbots or opting for neither (a no-choice option). The conditional logit model was used to estimate user preferences.

Results: Including 957 participants for analysis, participants are more willing to choose a chatbot with 100% response accuracy(?=0.940, P<.001), very easy to understand information(?=0.907, P<.001), service fee of ¥0/month (?=-0.095, P<.001), information utility is very practical(?=1.085, P<.001) and the provision of information content is disease knowledge(?=0.752, P<.001). Whether or not to allow the collection of background information (only question-answering information) has no significant impact on women's choice preferences.

Conclusions: Artificial intelligence chatbots should focus on high response accuracy, clear content, free access, privacy protection, practical information, and disease knowledge to attract users and enhance health education.

Preprint Settings

1) Would you like to publish your submitted manuscript as preprint?

✔ **Please make my preprint PDF available to anyone at any time (recommended).**
   Please make my preprint PDF available only to logged-in users; I understand that my title and abstract will remain visible to all users.
   Only make the preprint title and abstract visible.
   No, I do not wish to publish my submitted manuscript as a preprint.

2) If accepted for publication in a JMIR journal, would you like the PDF to be visible to the public?

✔ **Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).**
   Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain v
   Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in <a href="http

Original Manuscript

Enhancing Women's Health Literacy through Artificial Intelligence Chatbots: A Discrete Choice Experiment Study

Jing Wang[a†], Hewei Min[a†], Tao Li[b], Jiaheng Li[c], Yang Jiang[d], Jingbo Zhang[e], Yibo Wu[a] Xinying Sun[a]*

a School of Public Health, Peking University, Beijing, China;

b Department of gynaecology and obstetrics, The Fourth Central Hospital of Baoding City, Hebei, China;

c School of Basic Medical Sciences, Hebei University, Hebei, Baoding, China;
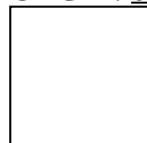
d North China University of Science and Technology, Hebei,Tangshan, China;

e Center for Evidence-Based Chinese Medicine, Beijing University of Chinese Medicine, Beijing, China;

* xysun@bjmu.edu.cn;

Tel: +86 13691212050

ORCID: 0000-0001-6638-1473

uthors contributed equally to this work.

Abstract

Background

Over 96% of adult women face health issues, with 70% experiencing conditions like infections. Mobile health education is increasingly popular but faces challenges in personalization and readability. Artificial intelligence chatbots provide tailored support, and a Discrete Choice Experiment can help in understanding user preferences to improve chatbot design.

Objective

This study explores the preferences of women towards artificial intelligence chatbots to improve health education communication, user experience, and health literacy.

Methods

A discrete choice experiment was conducted, identifying six main attributes of artificial intelligence chatbots: response accuracy, legibility, service cost, background information collection, information utility, and content provision. 957 participants from a hospital in Hebei Province participated, choosing between two hypothetical chatbots or opting for neither (a no-choice option). The conditional logit model was used to estimate user preferences.

Results

Including 957 participants for analysis, participants are more willing to choose a chatbot with 100% response accuracy(β=0.940, P<.001), very easy to understand information(β=0.907, P<.001), service fee of ¥0/month (β=-0.095, P<.001), information utility is very practical(β=1.085, P<.001) and the provision of information content is disease knowledge(β=0.752, P<.001). Whether or not to allow the collection of background information (only question-answering information) has no significant impact on women's choice preferences.

Conclusion

Artificial intelligence chatbots should focus on high response accuracy, clear content, free access, privacy protection, practical information, and disease knowledge to attract users and enhance health education.

Keywords: women's health; artificial intelligence chatbots; discrete choice experiment; willingness to pay; health education; health literacy

Introduction

According to a survey by the World Health Organization (WHO), over 96% of adult women globally encounter health conditions specific to women. Annually, approximately 9 million women succumb to these health issues, equating to roughly 20 deaths per minute, with an annual increase rate of 8%. Data from China's 2020 census reveal that the country has a female population of around 680 million, among whom about 70% are affected by vaginal infections or other women's health conditions; this rate increases to 90% among women of childbearing age. The Chinese Health Statistics Yearbook reports that obstetrics and gynecology outpatient visits number between 500-600 million annually, with over 55% attributed to reproductive tract infections, leading to over 20 billion yuan in healthcare expenses each year[1].

Compared to other health conditions, issues related to women's health are generally more preventable and sensitive. Improving health literacy among women plays a key role in preventing these conditions and mitigating their progression. Promoting awareness and education not only benefits the women themselves but also has significant implications for society as a whole[2].

There are various methods to promote health knowledge among women. These include regular physical examinations provided by institutions, distributing informational materials on women's health, and offering medical consultations and educational training. Health-related content is also available through special television programs, radio segments[3], magazines, and newspapers. With the advancement of technology, how women access health advice has been transforming. Mobile health education, based on applications and social media platforms, enhances health literacy and awareness among women[4-9]. In contrast to traditional methods, mobile health education offers greater accessibility, personalization, and cost-effectiveness, overcoming time and geographical barriers and saving human resources. This makes it a primary means of disseminating health information in recent years[10-12].

However, mobile health education still face many challenges, including content customization, real-time interaction, and personalized health advice provision. The diversity of content, accuracy of information, and level of understanding of health knowledge greatly affect the effectiveness of education. Especially for women with low health literacy, it is crucial to design content that is highly readable and practical. At this juncture, artificial intelligence (AI) chatbots emerge as a novel solution for mobile health education[13]. AI, which comprises programs and algorithms that perform tasks traditionally requiring human intelligence, is extensively applied in healthcare, showing promise in both specialized and general settings[14,15]. AI chatbots, as one of the key AI applications, provide interactive platforms for delivering health information and services[16]. Using natural language processing (NLP)[17], these systems convert user inquiries into a machine-readable format, leveraging a broad spectrum of medical data to establish a knowledge base that supports health-related interactions[18]. AI chatbots can offer appropriate health advice and facilitate health discussions, thereby supporting communication and patient empowerment[19]. With the expansion of mobile internet access, AI chatbots offer a practical and readily accessible tool for women with limited healthcare access or low health literacy, enhancing patient-centered care and self-management[20]. Their use notably increased during the COVID-19 pandemic, demonstrating positive support in various health domains[21]. Specifically, AI chatbots effectively address women's health topics with tailored and comprehensive guidance, enhancing the dissemination of health knowledge on public platforms.

Artificial intelligence (AI) chatbots have shown promising applications in the realm of women's

health, particularly in the management of gynecological conditions. They offer personalized support, educational resources, and assistance in self-assessing symptoms. Research indicates that AI chatbots can effectively reduce anxiety and enhance patient engagement in areas such as cancer care and reproductive health management[22]. However, existing studies have not sufficiently investigated the customization of these AI tools for specific gynecological needs. This study aims to fill that gap by employing a Discrete Choice Experiment (DCE) to understand user preferences. To improve user experience and the accuracy of AI chatbot information, it is crucial to comprehend user preferences. A suitable method for measuring these preferences is the Discrete Choice Experiment (DCE), a robust survey technique that presents respondents with multiple choices among hypothetical treatments[23]. The choices made by respondents are influenced by the alternatives presented. By offering options that consist of various attributes at differing levels and repeatedly asking participants to select their preferred option or the one that maximizes utility[24], we seek to understand how users value different aspects of AI chatbot use in women's health. Quantifying these preferences enables the identification of preferred attributes and provides insight into how women perceive and value AI chatbots for their health needs[25].

Methods
Study Design
In this study, we investigated women's preferences for attributes of the AI chatbots using the discrete choice experiment (DCE) method. DCE is a questionnaire-based stated preference approach to study people's preferences for certain characteristics or conditions by asking subjects to make choices about hypothetical scenarios or alternatives they prefer[26,27], DCE can be used to obtain quantitative data about people's preferences, gain in-depth insights and predictive power, and is often applied in market research, social sciences, medicine and other fields. Selected attributes in the DCE scheme can contain price attributes for different price levels to estimate WTP[28].

Development of the Discrete Choice Experiment
In the DCE section, after the preliminary literature investigation, we conducted 3 rounds of expert consultation and a pre-survey of 100 people (the pre-survey locations and methods were the same as the formal survey, and the results were not included in the text). Ultimately, we identified six attributes of AI chatbots: (1) response accuracy (2) legibility (3) service cost (4) whether to allow background collection of information (question-answering information only) (5) information utility (6) provision of information content. Each attribute and its level are shown in Table 1.

Table 1: DCE attributes and levels

Attribute Number	property name	Attribute Level
1	Response accuracy	60%; 70%; 80%; 90%; 100%
2	Legibility	Very easy to understand; easier to understand; harder to understand; difficult to understand
3	service cost	¥0/month; ¥5/month; ¥10/month; ¥15/month; ¥20/month
4	Whether to allow the collection of background information (only question-answering information)	Yes; No

5	Information Utility	Very practical; More practical; Less practical; Very impractical
6	Provision of information content	Knowledge of diseases; daily health care; knowledge of maternity; gender knowledge

Based on the identified attribute levels, we created choice sets containing various attribute levels for respondents to select from through orthogonal tests. Given that there are 2 to 5 attribute levels for each of the 6 attributes in our study, employing a full factorial design would result in 3,200 combinations ($5 \times 4 \times 5 \times 2 \times 4 \times 4 = 3,200$), which is clearly impractical for generating such a large number of choice sets. Consequently, we utilized a fractional factorial design to determine the optimal number of choice sets. Following the principles of orthogonality, balance, and minimal overlap, we established 24 random tasks, 1 fixed task, and 3 conceptual options for each task ("choose option A"; "choose option B"; "choose neither"). The 24 tasks were randomly divided into 4 groups, with each respondent completing 7 choice tasks (6 random tasks plus 1 repeated task). Respondents selected their preferred combination of attributes for the AI chatbots in each scenario or opted for none, thereby minimizing bias resulting from forced choices. According to the sample size formula for discrete choice experiments (DCE): n = (500c)/(ta), where "c" represents the largest number of levels for an attribute, "t" denotes the number of choice sets in a block, and "a" indicates the number of alternatives, and assuming that 5% of respondents choose "no option", the sample size for this study is determined to be 220. Table 2 shows an example of a DCE program.

Table 2: An example scenario of the choice-based conjoint in the questionnaire.

Attribute	Option A	Option B
Response accuracy	70%	100%
Legibility	easier to understand	harder to understand
Service cost	¥10/month	¥0/month
Whether to allow the collection of background information (only Q&A information)	No	No
Information Utility	Very practical	Very practical
Provision of information content	Knowledge of diseases	Knowledge of diseases
	Option A Option B None of them	

Study population and data collection

This study began in December 2022 with a questionnaire survey of women from a hospital in Hebei Province, which lasted for 2.5 months. Research information was collected through electronic questionnaires. The questionnaire is divided into two parts, the first part is basic demographic characteristics, including gender, age, education level, and usual residence. The second part is a survey on the preference of AI chatbots. In total, we distributed a total of 1281 questionnaires and successfully returned 1216 valid questionnaires, with an effective rate of 94.9%. The inclusion and exclusion criteria are as follows:

Inclusion criteria:
1. Age ≥ 18 years.
2. Sex: Female.
3. Possesses the nationality of the People's Republic of China.

4. Own a smartphone and have the ability to use WeChat and follow WeChat public numbers.

5. Ability to complete web-based questionnaires on your own or with the assistance of a surveyor.

6. Possess basic literacy skills to enable normal communication and interaction.

7. Voluntary participation in the study and willingness to complete an informed consent form.

Exclusion Criteria:

1. Individuals who are delirious or mentally abnormal.
2. Individuals with cognitive dysfunction.
3. Individuals who are participating in other similar research projects.
4. Individuals who are unwilling to cooperate.

Ethical Considerations

The study has been approved by the Ethics Committee of the Baoding No. 4 Central Hospital (approval number 2022013), and all study subjects have signed informed consent forms to ensure that they clearly understand the purpose and process of the study and that participate voluntarily.

Statistical analysis

Statistical analyses were performed using IBM SPSS Statistics 26,0 (IBM Corporation, Armonk, NY, USA) and Stata 15.0. Descriptive statistics of demographic variables were performed by frequency counts (composition ratios). Conditional logit models (CLOGIT) were used to quantify the relative levels of attribute preferences used by the AI chatbots via Stata 15.0. Different levels of each attribute were dummy-coded and one of the levels was set as the reference level. In a Discrete Choice Experiment study, the calculated results provide important statistical information. These results include coefficients, p-values, standard errors, and 95% confidence intervals (95% CI). We also calculated the study participants' payment willingness (WTP) for different attribute levels of the AI chatbots to more intuitively reflect the strength of the respondents' attribute preferences for the AI chatbots.

Results

Characteristics of Respondents

After excluding unqualified questionnaires based on the exclusion conditions, 957 questionnaires were included in the analysis. Figure 1 depicts the flowchart of the study process. The socio-demographic characteristics of the study sample are presented in Table 3. All participants were female, with 75.86% employed, 2.82% students, 1.15% retirees, and 20.17% in unstable or no employment. Most participants resided in urban areas (66.14%). Age distribution was 16.61% (18-25), 45.04% (26-35), 29.26% (36-45), and 9.09% (>45). Educational levels included 27.80% junior high or below, 22.05% special/senior high, and 50.16% college or above. A majority were married (86%), while 12.02% were single, 1.67% divorced, and 0.31% widowed. Household per capita monthly income was 40.33% (3001-5000 yuan), 35.32% (1001-3000 yuan), 17.45% (>5000 yuan), and 6.90% (≤1000 yuan).
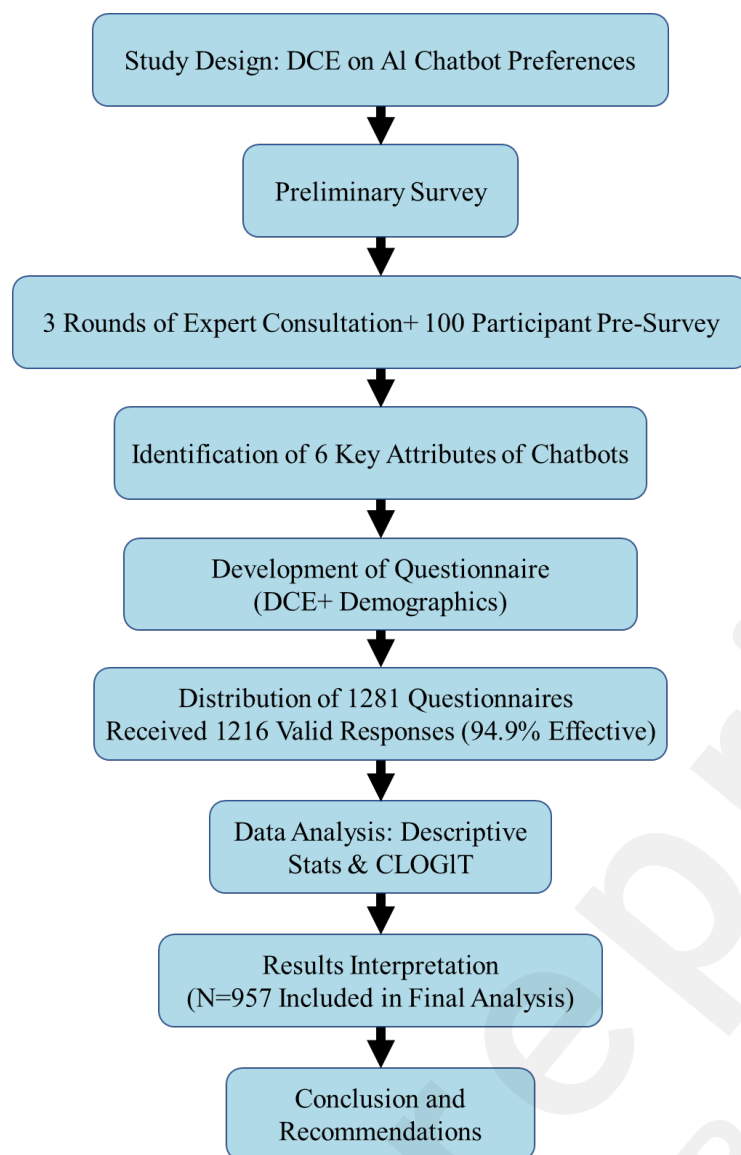
Figure 1. Flowchart of Study Process

Table 3: General characteristics of the subjects(n=957)

Items	Number of cases (%)
Occupation	
Employed	726 (75.86)
Student	27 (2.82)
Retired	11 (1.15)
Unstable Occupation/Freelancer/ Unemployed/Not in Employment	193 (20.17)
Location	

	Urban	633 (66.14)
	Rural	324 (33.86)
Age		
	18-25	159 (16.61)
	26-35	431 (45.04)
	36-45	280 (29.26)
	>45	87 (9.09)
Education level		
	Junior high school or lower	266 (27.80)
	Special school or senior high school	211 (22.05)
	Junior college or higher	480 (50.16)
Marital Status		
	Single	115 (12.02)
	Married (including first marriage, remarried, or remarried after divorce)	823 (86.00)
	Divorced	16 (1.67)
	Widowed	3 (0.31)
Income*		
	≤1000	66 (6.90)
	1001-3000	338 (35.32)
	3001-5000	386 (40.33)
	￿5000	167 (17.45)

* Per capita monthly household income

The Percent Importance of AI Chatbot Attributes

To analyze the overall preferences of study participants, the discrete choice model was utilized to determine the relative importance of each AI chatbot attribute. This importance reflects the extent to which each attribute influences the differences in preferences. It is calculated based on the range of preference weights across each attribute's levels, which provides a measure of its impact on decision-making. The resulting relative importance values, which collectively sum to 100%, indicate the degree to which each attribute is valued by respondents. A higher value corresponds to a greater influence on preferences. As shown in Table 4 and Figure 2, the information utility of the chatbot is considered the most important attribute (28.123%). Response accuracy and legibility are also regarded as significant, at 24.365% and 23.510%, respectively. The importance of information

content provision stands at 19.466%, while cost and whether background information collection is allowed hold similar importance, at 2.462% and 2.074%, respectively.

Table 4: Importance of Attributes（N=957）

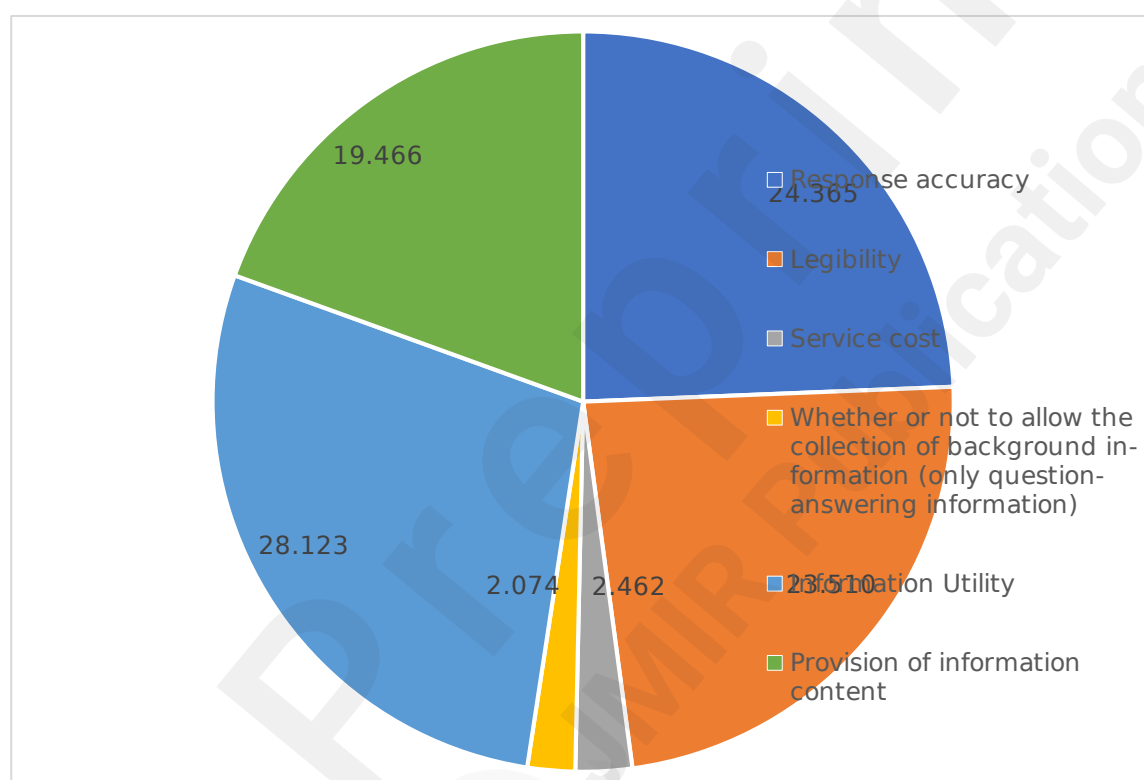| Attribute	Percent importance(%)
Response accuracy	24.365
Legibility	23.510
Service cost	2.462
Whether or not to allow the collection of background information (only question-answering information)	2.074
Information Utility	28.123
Provision of information content	19.466

Figure 2. Importance of Attributes

DCE Results

Based on the results of the CLOGIT model analysis, the preference of the survey respondents for the women's health AI chatbots can be determined. Table 5 displays the results, showing the relative importance of each attribute level for consumer preference (coefficients), the significance of the coefficients (P-value), the degree of confidence in the coefficient estimates (standard error), and the confidence interval of the coefficient estimates (95% CI). From the analysis, it is evident that specific attributes and levels have an impact on respondents' preferences. These attributes include the accuracy of the answers, the readability of the information, the cost of the service, the utility of the information, and the content attributes of the provided information. The participants exhibited a preference for attributes such as 100% response accuracy, very easy to understand readability, ¥0/month service cost, very practical information utility, and provision of information content as knowledge of the disease in the AI chatbots. Allowing the collection of background information

(only question-answering information) does not significantly affect women's preferences.

Table 5: CLOGIT results of participants' preferences for attributes of AI chatbots

Attributes and Levels	Coefficient	P-value	Standard error	95% CI
Response accuracy				
60%[a]				
70%	0.561	.001	0.168	0.232-0.891
80%	0.511	0.002	0.167	0.184-0.838
90%	0.637	<.001	0.161	0.321-0.953
100%	0.940	<.001	0.161	0.624-1.255
Legibility				
Harder to understand [a]				
Difficult to understand.	0.207	0.205	0.163	-0.113-0.528
Easier to understand	0.725	<.001	0.137	0.456-0.993
Very easy to understand.	0.907	<.001	0.139	0.634-1.180
Service cost	-0.095	<.001	0.007	-0.108-0.082
Whether or not to allow the collection of background information (only question-answering information)				
Yes [a]				
No	0.080	0.37	0.090	-0.095-0.254
Information Utility				
Very impractical [a]				
Less practical	0.180	0.241	0.153	-0.121-0.481
More practical	0.584	<.001	0.164	0.262-0.907
Very practical	1.085	<.001	0.129	0.832-1.338
Provision of information content				
Gender knowledge [a]				
Knowledge of Maternity	0.247	0.108	0.153	-0.141-0.447
daily health care	0.664	<.001	0.137	0.396-0.932
Knowledge of Diseases	0.751	<.001	0.136	0.485-1.018

[a]: Reference level

WTP Results

Willingness to pay (WTP) is a valid indicator of how much money a person is willing to sacrifice to choose one diagnostic attribute level over another (i.e., the reference attribute level). Table 6 shows the participants' willingness to pay for different attributes of the AI chatbots. For answer accuracy, the price that participants are willing to pay for it increases overall as accuracy increases (70%: 5.924 CNY; 80%: 5.389 CNY; 90%: 6.721 CNY; 100%: 9.916 CNY); and for information accessibility, participants' willingness to pay rises overall as accessibility increases (hard to understand: 2.188 CNY; easier to understand: 7.646 CNY; very easy to understand: 9.567 CNY); for the utility of information, the higher the utility, the higher the price participants are willing to pay (less practical: 1.900 CNY; more practical: 6.167 CNY; very practical: 11.451 CNY); for the provision of information content, participants who provide knowledge about diseases are willing to pay a higher price (knowledge of maternity: 2.602 CNY; daily health care: 7.006 CNY; knowledge of diseases: 7.931 CNY).

Table 6: Participants' willingness to pay for AI chatbots (WTP[a])

Attributes and Levels	WTP (CNY[b])	95%CI
Response accuracy		
60%	REF	
70%	5.924	2.628-9.220

80%	5.389	2.034-8.744
90%	6.721	3.637-9.806
100%	9.916	6.843-12.292
Legibility		
Harder to understand	REF	
Difficult to understand.	2.188	-1.154-5.529
Easier to understand	7.646	4.883-10.409
Very easy to understand.	9.567	6.843-12.292
Whether or not to allow the collection of background information (only question-answering information)		
Yes	REF	
No	0.839	-1.017-2.695
Information Utility		
Very impractical.	REF	
Less practical	1.900	-1.278-5.078
More practical	6.167	2.701-9.633
Very practical	11.451	8.704-14.198
Provision of information content		
Gender knowledge	REF	
Knowledge of Maternity	2.602	-0.551-5.756
Daily health care	7.006	4.049-9.963
Knowledge of Diseases	7.931	4.975-10.886

[a]: willingness to pay

[b]: Chinese Yuan

Discussion

This study utilized a Discrete Choice Experiment (DCE) to explore women's preferences for artificial intelligence (AI) chatbots. The results revealed significant variations in user preferences across different chatbot attributes, particularly for information usefulness, response accuracy, readability, and content provision, whereas cost and data collection were of relatively lower importance.

The scalability, accessibility, ease of use, and rapid information dissemination of AI chatbots offer supplementary benefits to public health efforts[29], addressing issues such as capacity constraints, social distancing requirements, and misinformation[30]. Research suggests that people sometimes prefer interacting with AI chatbots over doctors in certain contexts, as these chatbots can respond more quickly, provide high-quality feedback[31], and offer empathetic interactions. The application of AI in medicine is expanding across various domains, including medical image analysis[32], drug interaction detection[33], high-risk patient identification[34], and medical record coding[35]. In healthcare education, tools like ChatGPT have been used to facilitate personalized learning, encourage critical thinking, and support problem-based learning[36]. The widespread use of AI has made chatbots popular for accessing health information and improving health literacy.

The study found that information utility is the most valued attribute of AI chatbots among participants. Participants tend to prefer chatbots that provide practical and relevant information, aligning with the findings of Kim et al. (2024) in their systematic review and meta-analysis on women's health[22], finding that AI chatbots positively impacted women's health by reducing anxiety and depression, promoting healthy behaviors, and offering health education. The review highlights that the provision of practical and relevant information is a key factor in improving health

outcomes for women. However, in our study, the P-value of the 'less practical' level did not reach significance, which means that the practicality of information only has a significant impact on user choices when it is very high or very low. Users expect the information provided by chatbots to have practical application value. Therefore, when designing chatbots, developers should focus on providing practical information that can meet users' actual needs and be immediately applied to their daily lives.

The results also indicated that response accuracy and readability significantly influence participants' preferences. Participants favor AI chatbots with high accuracy and comprehensible content, which aligns with the systematic review by Aggarwal et al. (2023)[37]. Aggarwal et al. included 15 empirical studies on AI chatbots that facilitate health behavior changes, including interventions like healthy lifestyles, smoking cessation, and medication adherence. While some studies demonstrated the efficiency of AI chatbots, mixed results were observed regarding their feasibility, acceptability, and usability[38,39]. This highlights the need for improving chatbot accuracy and readability to effectively promote behavior change. Accurate health information is key to building user trust[40], and this study further confirms that participants value accurate responses. Readability also plays a significant role in users' choices, particularly among those with low health literacy or unfamiliarity with complex medical terms, as easily understandable information enhances comprehension and adherence[41-43]. In our study, the P-values for the "difficult to understand" attribute were not statistically significant, suggesting that the readability of AI chatbot information may not substantially influence user preferences within certain thresholds. This indicates that users may prioritize levels of readability that are "easier to understand" or "very easy to understand." When information reaches a level deemed difficult to comprehend, users may be disinclined to engage further with the AI chatbot. These findings suggest that AI chatbots should optimize language and interface design based on accurate content to ensure comprehensibility and cater to diverse women's needs.

Regarding content preferences, participants showed a greater interest in disease knowledge, consistent with the systematic review by Alyasiri et al. (2024), which categorized AI applications across healthcare domains such as patient support and education[44]. AI chatbots can provide personalized health education and disease prevention information, offering content relevant to users' needs. The preference for pregnancy and childbirth knowledge is not significant, indicating that these types of information do not have a clear advantage in user selection preferences. This may be because women are more concerned with the prevention, diagnosis, and treatment of diseases themselves, and may not have a direct need for information related to pregnancy and childbirth as much as disease knowledge. While emphasizing the importance of providing information on disease knowledge, it is crucial not to overlook the demand for content in other areas. Women's health is a multifaceted field that includes knowledge related to gender, pregnancy, childbirth, daily healthcare, and various diseases. Therefore, when developing chatbots, it is essential to consider all these aspects comprehensively to address the diverse needs of users.

On the topic of service cost, participants prefer free AI chatbots. This preference is consistent with other systematic reviews[45], which noted that free or low-cost services are more easily accepted by users, particularly when economic pressures are high. In a fast-paced era, the accessibility of free AI chatbots allows users to seek answers anytime and anywhere, increasing their reliance and usage. The preference for free services spans all age groups and socioeconomic backgrounds, suggesting that financial factors strongly influence the choice for cost-free AI tools.

Finally, data collection did not significantly influence participants preferences. However, Kim et al. (2024) emphasized the importance of privacy protection in AI applications for women's health[22].

Although data collection can help improve service quality, ensuring data security remains crucial in healthcare settings. Users highly value the protection of personal health data, and limiting data collection to non-identifiable question-and-answer content can help alleviate concerns about data misuse or breaches.

While this study provides valuable insights, its limitations include a sample confined to hospitalized women, which restricts generalizability. Future research should encompass broader populations and incorporate other attributes such as response accuracy and information readability. Questionnaire assumptions may not fully reflect real-world situations, as preferences can vary by age, education, and other factors. Further research is needed to deepen our understanding of these preferences and improve chatbot design to meet diverse user needs effectively.

Conclusion

This discrete choice experiment highlights key preferences for AI chatbots in healthcare, indicating that users prioritize high response accuracy (100%), legibility, free services, and the practical utility of information, particularly in relation to disease knowledge. These findings suggest that the development of AI chatbots should emphasize clarity, accessibility at no cost, user privacy protection, and the relevance of health content. By incorporating these attributes, user engagement can be enhanced, the dissemination of health education can be improved, and the effectiveness of AI tools in promoting health literacy and disease management can be maximized.

List of abbreviations

NLP: natural language processing; DCE: Discrete Choice Experiment; WTP: Willingness to Pay; CNY: Chinese Yuan

Ethics approval and consent to participate
The study has been approved by the Ethics Committee of the Baoding No. 4 Central Hospital (approval number 2022013), and all study subjects have signed informed consent forms to ensure that they clearly understand the purpose and process of the study and that participate voluntarily.

Consent for publication
Not applicable.

Availability of data and materials
The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Competing interests
The authors declare that they have no competing interests.

Funding

Authors' contributions
YBW and HWM conceived and designed the study. JW, HWM, YJ and JBZ reviewed and revised the manuscript. YBW provided theoretical guidance and methodological expertise. HWM designed the questionnaire. TL and JHL are responsible for the study implementation. All authors provided revisions and approved the final manuscript.

Acknowledgments

We are grateful to the professionals and students who provided enormous support in the recruitment and data collection of this study, and the participants who took the time to cooperate with the survey.

References

[1] World Health Organization. (2024). *Maternal mortality*. Retrieved from https://www.who.int/news-room/fact-sheets/detail/maternal-mortality (World Health Organization (WHO)).

[2] Hibbard JH, Pope CR. Women's roles, interest in health and health behavior. women Health. 1987;12(2):67-84. doi: 10.1300/J013v12n02_05. PMID. 3424849.

[3] Yang X, Yang F, Fu W, et al.The current status of community mental health services in three northern areas of China.Family Medicine and Community Health 2014;2:doi: 10.15212/FMCH.2014.0104

[4] Lin F, Sun Q, Peng Z, et al.The innovations in China's primary health care reform: Development and characteristics of the community health services in Hangzhou.Family Medicine and Community Health. 2015;3:doi: 10.15212/FMCH.2015.0135

[5] Carter M, Burley V, Nykjaer C, Cade J.Adherence to a Smartphone Application for Weight Loss Compared to Website and Paper Diary: Pilot Randomized Controlled Trial.J Med Internet Res 2013;15(4):e32.URL: https://www.jmir.org/2013/4/e32.DOI: 10.2196/jmir.2283.

[6] Cheng Z, Guo S. Automatic Text Summarization for Public Health WeChat Official Accounts Platform Base on Improved TextRank. J Environ Public Health. 2022 Jul 13;2022:1166989. doi: 10.1155/2022/1166989. PMID: 35874881; PMCID: PMC9300291.

[7] Chen X, Zhou X, Li H, Li J, Jiang H. The value of WeChat application in chronic diseases management in China. Comput Methods Programs Biomed. 2020 Nov;196. 105710. doi: 10.1016/j.cmpb.2020.105710. Epub 2020 Aug 14. PMID: 32858284.

[8] Jiang Y, Liu F, Guo J, Sun P, Chen Z, Li J, Cai L, Zhao H, Gao P, Ding Z, Wu X. Evaluating an Intervention Program Using WeChat for Patients With Chronic Obstructive Pulmonary Disease: Randomized Controlled Trial. J Med Internet Res. 2020 Apr 21;22(4):e17089. doi: 10.2196/17089. PMID: 32314971; PMCID PMC7201319.

[9] Sun M, Yang L, Chen W, Luo H, Zheng K, Zhang Y, Lian T, Yang Y, Ni J. Current status of official WeChat accounts for public health education. j Public Health (Oxf). 2021 Sep 22;43(3):618-624. doi: 10.1093/pubmed/fdz163. PMID: 31974552.

[10] Safran Naimark J, Madar Z, Shahar DR. The impact of a Web-based app (eBalance) in promoting healthy lifestyles: randomized controlled trial. J Med Internet Res. 2015 Mar 2;17(3):e56. doi: 10.2196/jmir.3682. PMID: 25732936; PMCID: PMC4376125.

[11]Lapp B, Zakus D.Factors associated with the utilization of institutional and home birth services among women in Ethiopia: A scoping review.Family Medicine and Community Health 2016;4:doi: 10.15212/FMCH.2016.0125

[12]Kao CK, Liebovitz DM. Consumer Mobile Health Apps: Current State, Barriers, and Future Directions. PM R. 2017 May;9(5S):S106-S115. doi: 10.1016/j.pmrj.2017.02.018. PMID: 28527495.

[13] Manickam P, Mariappan SA, Murugesan SM, Hansda S, Kaushik A, Shinde R, Thipperudraswamy SP. Artificial Intelligence (AI) and Internet of Medical Things (IoMT) Assisted Biomedical Systems for Intelligent Healthcare. biosensors (Basel). 2022 Jul 25;12(8):562. doi: 10.3390/bios12080562. PMID: 35892459; PMCID: PMC9330886.

[14] Hendrix N, Hauber B, Lee CI, Bansal A, Veenstra DL. Artificial intelligence in breast cancer screening: primary care provider preferences. J Am Med Inform Assoc. 2021 Jun 12;28(6):1117-1124. doi: 10.1093/jamia/ocaa292. PMID: 33367670; PMCID: PMC8200265.

[15] Miller DD, Brown EW. Artificial Intelligence in Medical Practice: The Question to the Answer? Am J Med. 2018 Feb;131(2):129-133. doi: 10.1016/j.amjmed.2017.10.035. Epub 2017 Nov 7. PMID: 29126825.

[16] Wilson L, Marasoiu M. The Development and Use of Chatbots in Public Health: a Scoping Review. JMIR Hum Factors. 2022 Oct 5;9(4):e35882. doi: 10.2196/35882 . pmid: 36197708; pmcid: pmc9536768.

[17] Gilson A, Safranek CW, Huang T, Socrates V, Chi L, Taylor RA, Chartash D. How Does ChatGPT Perform on the United States Medical Licensing Examination? the Implications of Large Language Models for Medical Education and Knowledge Assessment. JMIR Med Educ. 2023 Feb 8;9:e45312. doi: 10.2196/45312. PMID. 36753318; PMCID: PMC9947764.

[18] Yeo YH, Samaan JS, Ng WH, Ting PS, Trivedi H, Vipani A, Ayoub W, Yang JD, Liran O, Spiegel B, Kuo A. Assessing the performance of ChatGPT in answering questions regarding cirrhosis and hepatocellular carcinoma. Clin Mol Hepatol. 2023 Jul;29(3):721-732. doi: 10.3350/cmh.2023.0089. Epub 2023 Mar 22 . pmid: 36946005; pmcid: pmc10366809.

[19] Liu J, Wang C, Liu S. Utility of ChatGPT in Clinical Practice. J Med Internet Res. 2023 Jun 28;25:e48568. doi: 10.2196/48568. PMID: 37379067; PMCID: PMC10365580.

[20] Milne-Ives M, de Cock C, Lim E, Shehadeh MH, de Pennington N, Mole G, Normando E, Meinert E. The Effectiveness of Artificial Intelligence Conversational Agents in Health Care: Systematic Review. J Med Internet Res. 2020 Oct 22;22(10):e20346. doi: 10.2196/20346. PMID: 33090118; PMCID: PMC7644372.

[21] Wilson L, Marasoiu M. The Development and Use of Chatbots in Public Health: a Scoping Review. JMIR Hum Factors. 2022 Oct 5;9(4):e35882. doi: 10.2196/35882 . pmid: 36197708; pmcid: pmc9536768.

[22] Kim HK. The Effects of Artificial Intelligence Chatbots on Women's Health: A Systematic Review and Meta-Analysis. Healthcare (Basel). 2024 Feb 23;12(5):534. doi: 10.3390/healthcare12050534. PMID: 38470645; PMCID: PMC10930454.

[23] Laba TL. Using Discrete Choice Experiment to elicit patient preferences for osteoporosis drug treatments: where to from here? Arthritis Res Ther. 2014. 16(2):106. doi: 10.1186/ar4501. Epub 2014 Mar 10. PMID: 25167089; PMCID: PMC4060568.

[24] von Wedel P, Hagist C. Physicians' preferences and willingness to pay for artificial intelligence-based assistance tools: a discrete choice experiment among german radiologists. BMC Health Serv Res. 2022 Mar 26;22(1):398. doi: 10.1186/s12913-022-07769-x. PMID: 35346183; PMCID. PMC8959781.

[25] Noben CYG, Stammen LA, Vaassen S, Haeren R, Stassen L, Mook WV, Essers B. Discrete choice experiment on educating value-based healthcare. postgrad Med J. 2021 Aug;97(1150):515-520. doi: 10.1136/postgradmedj-2019-137190. epub 2020 Aug 12. PMID: 32796111.

[26] Ryan M, Bate A, Eastmond CJ, Ludbrook A. Use of discrete choice experiments to elicit preferences. Qual Health Care. 2001 Sep;10 Suppl 1(Suppl 1):i55-60. doi: 10.1136/qhc.0100055... PMID: 11533440; PMCID: PMC1765744.

[27] Louviere JJ, Lancsar E. Choice experiments in health: the good, the bad, the ugly and toward a brighter future. Health Econ Policy Law. 2009 Oct;4(Pt 4). 527-46. doi: 10.1017/S1744133109990193. PMID: 19715635.

[28] von Wedel P, Hagist C. Physicians' preferences and willingness to pay for artificial intelligence-based assistance tools: a discrete choice experiment among german radiologists. BMC Health Serv Res. 2022 Mar 26;22(1):398. doi: 10.1186/s12913-022-07769-x. PMID: 35346183; PMCID. PMC8959781.

[29] Pearce C.Chronic disease in a digital health environment.Family Medicine and Community Health 2018;6:doi: 10.15212/FMCH.2017.0144

[30] Amiri P, Karahanna E. Chatbot use cases in the Covid-19 public health response. J Am Med Inform Assoc. 2022 Apr 13;29(5):1000-1010. doi: 10.1093/jamia/ ocac014. PMID: 35137107; PMCID: PMC8903403.

[31] Ayers JW, Poliak A, Dredze M, et al. Comparing Physician and Artificial Intelligence Chatbot Responses to Patient Questions Posted to a Public Social Media Forum. JAMA Intern Med. 2023;183(6):589-596. doi:10.1001/jamainternmed.2023.1838

[32] Meskó B, Topol EJ. The imperative for regulatory oversight of large language models (or generative AI) in healthcare. NPJ Digit Med. 2023 Jul 6;6(1):120. doi: 10.1038/s41746-023-00873-0. PMID: 37414860; PMCID: PMC10326069.

[33] Dunn P, Cianflone D. Artificial intelligence in cardiology: Exciting but handle with caution. Int J Cardiol. 2023 Oct 1;388:131117. doi: 10.1016/j.ijcard.2023.06.008. Epub 2023 Jun 9. PMID: 37302418.

[34] Zhong NN, Wang HQ, Huang XY, Li ZZ, Cao LM, Huo FY, Liu B, Bu LL. Enhancing head and neck tumor management with artificial intelligence: integration and perspectives. Semin Cancer Biol. 2023 Oct;95:52-74. doi: 10.1016/j.semcancer.2023.07.002. Epub 2023 Jul 18. PMID: 37473825.

[35] Li W, Zhang Y, Chen F. ChatGPT in Colorectal Surgery: A Promising Tool or a Passing Fad? Ann Biomed Eng. 2023 Sep;51(9):1892-1897. doi: 10.1007/s10439-023-03232-y. Epub 2023 May 10. PMID: 37162695.

[36] Sallam M. ChatGPT Utility in Healthcare Education, Research, and Practice: Systematic Review on the Promising Perspectives and Valid Concerns. Healthcare (Basel). 2023 Mar 19;11(6):887. doi: 10.3390/healthcare11060887. PMID: 36981544; PMCID: PMC10048148.

[37] Aggarwal A, Tam CC, Wu D, Li X, Qiao S. Artificial Intelligence-Based Chatbots for Promoting Health Behavioral Changes: Systematic Review. J Med Internet Res. 2023 Feb 24;25:e40789. doi: 10.2196/40789. PMID: 36826990; PMCID: PMC10007007.

[38] Caranfa JT, Bommakanti NK, Young BK, Zhao PY. Accuracy of Vitreoretinal Disease Information From an Artificial Intelligence Chatbot. JAMA Ophthalmol. 2023 Aug 3:e233314. doi: 10.1001/jamaophthalmol.2023.3314. Epub ahead of print. PMID: 37535363; PMCID: PMC10401388.

[39] Diamandis EP. Chatbot GPT can be grossly inaccurate. Clin Chem Lab Med. 2023 Jul 31. doi: 10.1515/cclm-2023-0765. Epub ahead of print. PMID: 37501646.

[40] Hirosawa T, Harada Y, Yokose M, Sakamoto T, Kawamura R, Shimizu T. Diagnostic Accuracy of Differential-Diagnosis Lists Generated by Generative Pretrained Transformer 3 Chatbot for Clinical Vignettes with Common Chief Complaints: a pilot study. Int J Environ Res Public Health. 2023 Feb 15;20(4). 3378. doi: 10.3390/ijerph20043378. PMID: 36834073; PMCID: PMC9967747.

[41] Moons P, Van Bulck L. Using ChatGPT and Google Bard to improve the readability of written patient information: a proof-of-concept. Eur J Cardiovasc Nurs. 2023 Aug 21:zvad087. doi: 10.1093/eurjcn/zvad087. Epub ahead of print. PMID: 37603843.

[42] Yurdakul OV, Kilicoglu MS, Bagcier F. Evaluating the reliability and readability of online information on osteoporosis. Arch Endocrinol Metab. 2021 Nov 1;65(1):85-92. doi: 10.20945/2359-3997000000311. Epub 2020 Nov 9. PMID: 33166440; PMCID: PMC10528702.

[43] Denecke K, Borycki EM, Kushniruk AW. What Can We Learn from Quality Requirements in ISO/TS 82304-2 for Evaluating Conversational Agents in Healthcare? Stud Health Technol Inform. 2022 Nov 3;299:245-250. doi: 10.3233/SHTI220992. PMID: 36325870.

[44] Younis HA, Eisa TAE, Nasser M, Sahib TM, Noor AA, Alyasiri OM, Salisu S, Hayder IM, Younis HA. A Systematic Review and Meta-Analysis of Artificial Intelligence Tools in Medicine and Healthcare: Applications, Considerations, Limitations, Motivation and Challenges. Diagnostics (Basel). 2024 Jan 4;14(1):109. doi: 10.3390/diagnostics14010109. PMID: 38201418; PMCID: PMC10802884.

[45] Grassini, E., Buzzi, M., Leporini, B. et al. A systematic review of chatbots in inclusive healthcare: insights from the last 5 years. Univ Access Inf Soc (2024). https://doi.org/10.1007/s10209-024-01118-x