

## Original Paper

# Medical Staff and Resident Preferences for Using Deep Learning in Eye Disease Screening: Discrete Choice Experiment

Senlin Lin<sup>1,2\*</sup>, MSc; Liping Li<sup>3\*</sup>, MPH; Haidong Zou<sup>1,2</sup>, MD; Yi Xu<sup>1,2</sup>, MD; Lina Lu<sup>1,2</sup>, MPH

<sup>1</sup>Shanghai Eye Disease Prevention and Treatment Center, Shanghai Eye Hospital, Shanghai, China

<sup>2</sup>Shanghai Key Laboratory of Ocular Fundus Diseases, Shanghai General Hospital, Shanghai Engineering Center for Visual Science and Photomedicine, Shanghai, China

<sup>3</sup>Shanghai Hongkou Center for Disease Control and Prevention, Shanghai, China

\*these authors contributed equally

**Corresponding Author:**

Lina Lu, MPH

Shanghai Eye Disease Prevention and Treatment Center, Shanghai Eye Hospital

No. 1440, Hongqiao Road

Shanghai, 200336

China

Phone: 86 02162539696

Email: [lulina781019@qq.com](mailto:lulina781019@qq.com)

## Abstract

**Background:** Deep learning–assisted eye disease diagnosis technology is increasingly applied in eye disease screening. However, no research has suggested the prerequisites for health care service providers and residents willing to use it.

**Objective:** The aim of this paper is to reveal the preferences of health care service providers and residents for using artificial intelligence (AI) in community-based eye disease screening, particularly their preference for accuracy.

**Methods:** Discrete choice experiments for health care providers and residents were conducted in Shanghai, China. In total, 34 medical institutions with adequate AI-assisted screening experience participated. A total of 39 medical staff and 318 residents were asked to answer the questionnaire and make a trade-off among alternative screening strategies with different attributes, including missed diagnosis rate, overdiagnosis rate, screening result feedback efficiency, level of ophthalmologist involvement, organizational form, cost, and screening result feedback form. Conditional logit models with the stepwise selection method were used to estimate the preferences.

**Results:** Medical staff preferred high accuracy: The specificity of deep learning models should be more than 90% (odds ratio [OR]=0.61 for 10% overdiagnosis;  $P<.001$ ), which was much higher than the Food and Drug Administration standards. However, accuracy was not the residents' preference. Rather, they preferred to have the doctors involved in the screening process. In addition, when compared with a fully manual diagnosis, AI technology was more favored by the medical staff (OR=2.08 for semiautomated AI model and OR=2.39 for fully automated AI model;  $P<.001$ ), while the residents were in disfavor of the AI technology without doctors' supervision (OR=0.24;  $P<.001$ ).

**Conclusions:** Deep learning model under doctors' supervision is strongly recommended, and the specificity of the model should be more than 90%. In addition, digital transformation should help medical staff move away from heavy and repetitive work and spend more time on communicating with residents.

(*J Med Internet Res* 2022;24(9):e40249) doi: [10.2196/40249](https://doi.org/10.2196/40249)

**KEYWORDS**

discrete choice experiment; preference; artificial intelligence; AI; vision health; screening

## Introduction

Vision loss, defined as either visual impairment or blindness, is becoming a vital aspect of public health [1], affecting economic, educational, and employment opportunities, reducing

the quality of life, and increasing the risk of death [1]. Therefore, according to the recent eye care competency framework by the World Health Organization, the continuum of eye care across all levels of the health system should be highlighted, particularly primary health care, to support universal health coverage [2].

High-quality eye disease prevention health care, such as effective screening, can help eliminate almost 57% of all blindness cases [3]. Nowadays, artificial intelligence (AI) is gradually adopted in eye disease screening and may assist in addressing the limited and difficult-to-sustain resources in screening capacity, personnel costs, and diagnosis expertise [4]. The accuracy of AI models greatly affects the cost-effectiveness of eye disease screening [5]. Unfortunately, though the US Food and Drug Administration (FDA) had set a mandatory level of accuracy with a sensitivity of more than 85% and a specificity of more than 82.5% [6], the accuracy of AI-assisted eye disease screening systems in the real world were far worse than that reported in the model development phase [7]. Therefore, it is essential to make clear the medical staff and resident requirements of the accuracy of AI models in the community-based eye disease screening in the real world. However, no related research has been conducted thus far.

To fill this evidence gap, we conducted discrete choice experiments (DCEs) for health care providers and residents in Shanghai, China, from August 2021 to January 2022. We aimed to reveal the preferences of medical staff and residents for using AI technology in community-based eye disease screening, particularly their preference for accuracy. The DCE technique, originating in mathematical psychology, has been introduced in health economics to elicit preferences for health and health care [8]. Additionally, the DCE technique is predictive of choices, mimicking real-world decisions in health care decision-making (correctly predicting >93% of choices) [9].

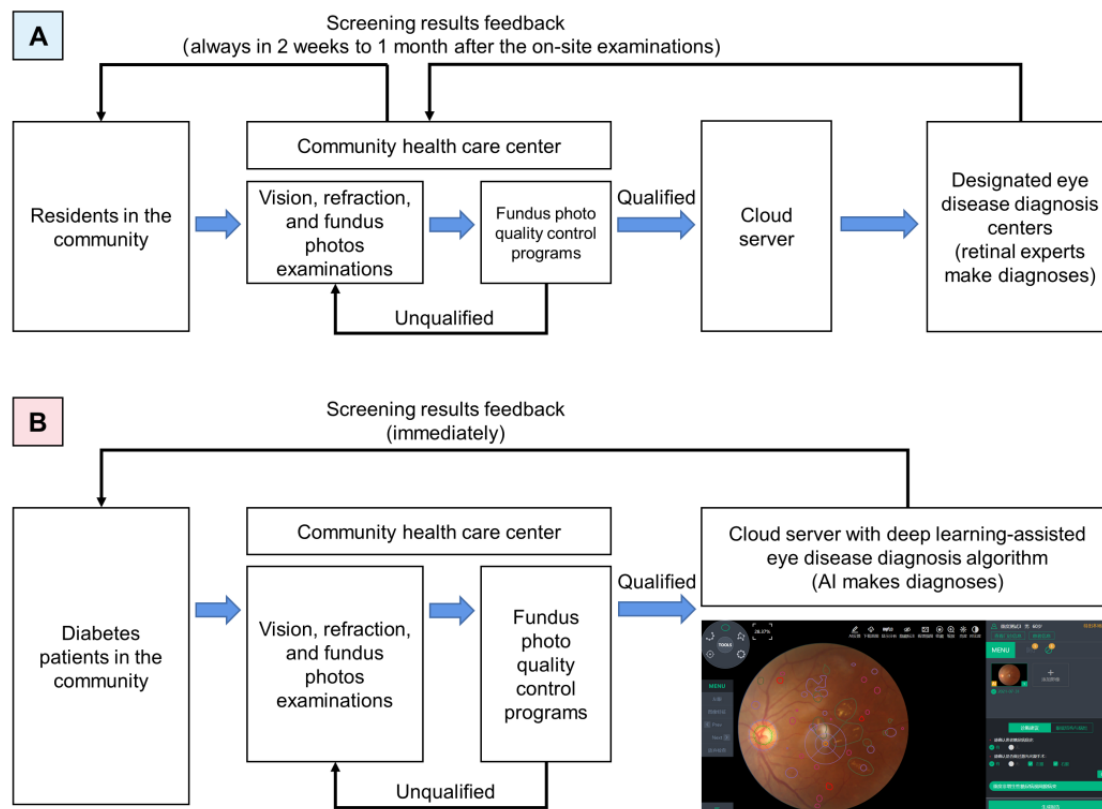
## Methods

### Study Setting

Shanghai, with a population of 24 million in 2019, is the economic, science, and technology innovation center in China. It is also one of the first cities in the world to adopt deep learning (DL) models to establish affordable and sustainable community-based eye disease screening systems. Since 2015, a teleophthalmology-based eye disease screening system covering all community health service centers has been developed in Shanghai. Residents can take free fundus photographs once a year by the trained general practitioners (GPs) in community health service centers. The fundus photos are then sent to the designated eye disease diagnosis centers through a dedicated information system. After the ophthalmologists in the diagnosis centers read the fundus photos and make diagnoses, the screening results are returned to the community health service centers. The GPs may inform residents of the screening results and provide medical advice.

In 2020, an AI-assisted eye disease screening system was established using DL model on cloud servers instead of ophthalmologists in the diagnosis centers making the screening diagnoses (Figure 1) [10-12]. The accuracy of DL models used for community-based eye disease screening has been reported widely [6,11,13,14]. Thus far, 56 community health service centers have shifted to the AI-assisted eye disease screening system. In 2021, these community health service centers screened over 40,000 residents with the help of the DL model and found over 7000 residents with suspected eye diseases.

**Figure 1.** Process of community-based eye disease screening in Shanghai. A: teleophthalmology-based eye disease screening system; B: deep learning–assisted eye disease screening system. The photograph in the lower right corner of process B is a sample of the operation interface of the deep learning–assisted eye disease diagnosis system. AI: artificial intelligence.



## Discrete Choice Experiments and Participants Inclusion

We conducted 2 DCEs to assess medical staff’s preferences (experiment 1) and residents’ preferences (experiment 2) for using the DL model in community-based eye disease screening. The main reason for using a DCE is that simply asking the respondents to rate the screening strategy attributes or choose their preferred item from a scale generally yields no more information than the fact that they want all the benefits and none of the indirect or direct costs [15]. Choosing between alternatives forces them to make a trade-off and choose, as in real life, between options that may increase utility (eg, improved diagnosis accuracy) and decrease utility (eg, screening cost of 40 CNY [US \$6.15] per resident instead of being free).

Based on previously published literature [4-7,16-18], 4 attributes were identified initially to describe the outline of the community-based eye disease screening, including the accuracy, screening result feedback efficiency, level of ophthalmologist’s involvement, and cost. It was worth stating that “screening result feedback efficiency” was included in the attributes because nearly instantaneous feedback might increase compliance [18]; moreover, “level of ophthalmologists involvement” was included because algorithmic aversion might exist [19]. To assess the appropriateness of these potential attributes and their levels, 5 experts on eye care were interviewed face-to-face in the Shanghai Eye Disease Control and Treatment Center. Based on

these interviews, the attribute *accuracy* was divided into the following 2 attributes: “missed diagnosis rate” and “overdiagnosis rate,” as they might have different impacts on the acceptability of eye disease screening. In addition, 2 new attributes were added: “organizational form” and “screening result feedback form,” as the adoption of the DL model had the potential to reform the screening programs. As a result, 7 attributes were used to describe the outline of the community-based eye disease screening, and each attribute was divided into 3-6 levels (Table 1). Three SAS (SAS Institute Inc) procedures—“%mktruns,” “%mktex,” and “%choiceff”—were used to develop the questionnaire [20]. The questionnaire consisted of the following two parts: the respondent’s basic information, such as sex and age, and a few choice sets, each of which contained 2 options with different screening attribute levels (Figure 2). The respondents were asked to choose the more favorable option in each choice set, and they were not allowed to choose both or neither in a set [21].

In Experiment 1, one municipal and 16 district-level eye disease control centers and over 250 community health service centers in Shanghai were enrolled. To receive rational rather than imaginary choices, the following two strict inclusion criteria were set: (1) they had over 5 years of experience in teleophthalmology-based eye disease screening and (2) they had over 1 year of experience in DL-assisted eye disease screening. A total of 34 institutions met the criteria, including 1 (3%) municipal, 16 (47%) district-level eye disease control

centers, and 17 (50%) community health service centers (Figure 3). All the 40 key persons in charge of community-based eye disease screening in these 34 institutions were invited and agreed to participate in the experiment. Due to the limited number of respondents, we had to ask each one to answer a relatively large number of questions. According to the rule of thumb, as proposed by Johnson and Orme [22], we divided the alternative screening strategies into 30 choice sets of 2 options to ensure that the sample size of 40 people met the statistical requirements. The experiment was conducted in the form of a self-administered questionnaire, with a trained investigator on standby to interpret the questionnaire. One respondent quit because of temporary work arrangements. Therefore, data from 39 medical staff were available in the final analysis.

In Experiment 2, we randomly selected 2 from the 17 community health service centers involved in Experiment 1 and conducted the residents' investigation when carrying out the AI-assisted community-based eye disease screening. All the residents who participated in the screening were invited to the experiment. Because the number of residents was relatively large, we divided the alternative screening strategies into 10 choice sets of 2 options to reduce the response burden for each respondent. According to the rule of thumb, as proposed by Johnson and Orme [22], the minimum of the required sample size was 125. A total of 318 residents were investigated (Figure 3). To help the residents understand the questionnaire, the experiment was conducted using face-to-face questioning by trained investigator.

**Table 1.** Attributes and levels in the discrete choice experiments.

Attributes	Levels					
	1	2	3	4	5	6
<b>Performance expectancy</b>						
Missed diagnosis rate (%)	None	5	10	15	20	— <sup>a</sup>
Overdiagnosis rate (%)	None	5	10	15	20	—
Screening result feedback efficiency	Immediately	In 2 weeks	In 1 month	—	—	—
<b>Effort expectancy</b>						
Level of ophthalmologist involvement	Fully automated <sup>b</sup> DL <sup>c</sup> model	Semiautomated <sup>d</sup> DL model	Fully manual diagnosis <sup>e</sup>	—	—	—
<b>Facilitating conditions</b>						
Organizational form	Centralized screening <sup>f</sup>	Residents' health self-examination cabin <sup>g</sup>	Opportunity screening in outpatient <sup>h</sup>	—	—	—
Cost	Free	40 CNY <sup>i</sup>	80 CNY	120 CNY	160 CNY	200 CNY
Screening result feedback form	Screening results <sup>j</sup>	Screening results and medical advice <sup>k</sup>	Screening results, medical advice, and oral explanation by GP <sup>l,m</sup>	—	—	—

<sup>a</sup>Not available.

<sup>b</sup>The screening results were provided entirely by the deep learning model, and the ophthalmologists were not involved in the diagnostic process.

<sup>c</sup>DL: deep learning.

<sup>d</sup>The deep learning model performed the initial screening of fundus photographs and then the ophthalmologists reviewed the results.

<sup>e</sup>The screening results were provided entirely by the ophthalmologists and the deep learning model was not involved in the diagnostic process.

<sup>f</sup>The community health service center informed the residents to undergo the screening at a uniform place and time.

<sup>g</sup>The equipment needed for screening was placed in a specific cabin in the community health service center, and residents could go to the cabin for self-examination at any time.

<sup>h</sup>Residents with chronic diseases and other risk factors would be recommended by general practitioners for eye disease screening during their outpatient follow-up.

<sup>i</sup>US 1\$=6.5 CNY.

<sup>j</sup>The report with only the screening results would be given to the residents without any recommendations or explanations.

<sup>k</sup>The report with the screening results and referral recommendations would be given to the residents without explanations.

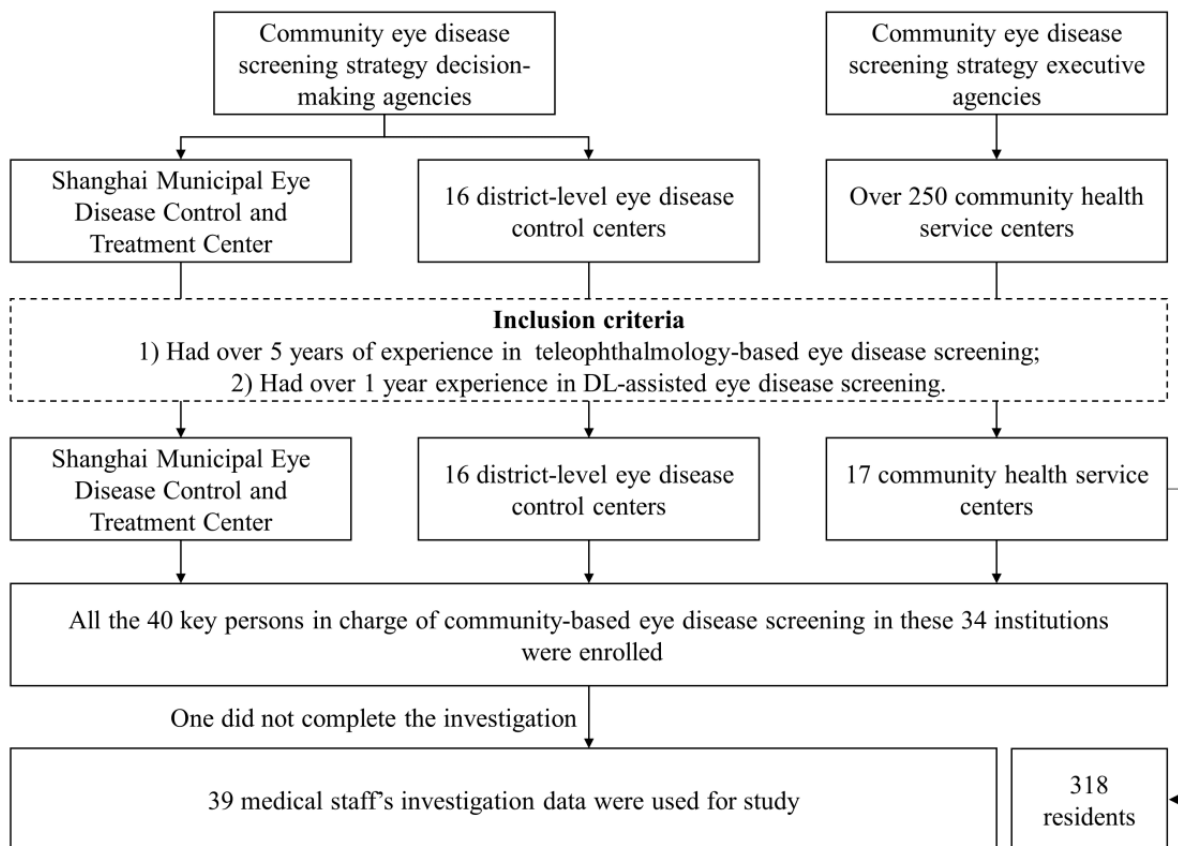
<sup>l</sup>Besides the report with the screening results and referral recommendations that would be given to the residents, a general practitioner would also explain the meaning of the report.

<sup>m</sup>GP: general practitioner.

**Figure 2.** Example of the choice sets applied. Both options include the same 7 attributes. Health care service providers and residents were asked to decide between options A and B (in 2021, 1 USD=6.5 CNY). AI: artificial intelligence; GP: general practitioner.

Which screening strategy would you choose?		
Set 1	A <input type="checkbox"/>	B <input type="checkbox"/>
<i>Diagnosis technology</i>	Fully AI diagnosis	Fully manual diagnosis
<i>Organizational form</i>	Opportunity screening in Outpatient	Residents' health self-examination cabin
<i>Missed diagnosis rate</i>	10%	None
<i>Overdiagnosis rate</i>	20%	10%
<i>Cost</i>	Free	120 CNY
<i>Screening result feedback efficiency</i>	In one month	Immediately
<i>Screening result feedback form</i>	1. Screening results 2. Referral recommendations 3. Oral explanation by GP	1. Screening results only

**Figure 3.** Medical staff and residents inclusion process. DL: deep learning.



## Statistical Analyses

Mean, median, and standard deviation were calculated for the quantitative variables. For categorical variables, the number in a specific category was calculated as a percentage. Pearson chi-square test for nominal variables and Mann-Whitney *U* test for continuous variables were used for statistical analysis. Conditional logit models with the stepwise selection method were used to explore the significant preferences for each attribute level, with the choice responses as the binary dependent variable and the difference in levels for each of the attributes as the independent variables [21]. Two models were used to estimate the medical staff's and residents' preference respectively, expressed as odds ratios (ORs) for each attribute level. SAS 9.4 (SAS Institute Inc) were used for statistical analysis. The level of significance was set at  $P < .05$ .

## Ethics Approval

All participants were adults. Written informed consent from all participants was obtained before enrollment. The study adhered to the principles of the Declaration of Helsinki on Ethics. This study was approved by the Shanghai General Hospital Ethics Committee (2022SQ272).

## Results

The medical staff's mean age was 39.67 (SD 6.98) years, and they had been responsible for eye disease screening for 6.73 (SD 5.76) years on average. The residents' mean age was 68.62 (SD 6.96) years; Of the 318 participants, 120 (37.74%) were

male and 198 (62.26%) were female. Detailed characteristics of the respondents are shown [Table 2](#).

[Table 3](#) presents the results of the conditional logit models, evaluating the influence of the tested attribute levels on medical staff's and residents' preferences. Among the 39 medical staff, the impact of selected attributes on preferences was statistically significant for 4 of the 7 attributes. Generally, medical staff prefer attribute levels with AI technology, lower overdiagnosis rates, lower screening costs, and higher screening result feedback efficiency. The results for the attribute "organizational form," "missed diagnosis rate," and "screening result feedback form" were inconclusive—none of the attribute levels were associated with statistically significant utility differences.

Further, we focused on the accuracy of the diagnosis. For the missed diagnosis rate, there were no significant differences of medical staff's preferences for a missed diagnosis rate between 0% and 20%. However, for the overdiagnosis rate, compared with no overdiagnosis, medical staff's preference for the 10% overdiagnosis rate significantly decreased (OR=0.61;  $P < .001$ ).

Among the 318 residents, the influence of selected attributes on preferences was statistically significant for 3 of the 7 attributes. Generally, residents were in disfavor of the attribute level with a fully automated DL model (OR=0.24;  $P < .001$ ), but they preferred attribute levels with lower screening costs and oral explanations by GP. The results for the attributes "organizational form," "missed diagnosis rate," "overdiagnosis rate," and "screening result feedback efficiency" were inconclusive. None of the attribute levels were associated with statistically significant utility differences.



**Table 2.** Characteristics of respondents.

Respondent and characteristics	Value
<b>Medical staff (n=39)</b>	
Age (years), mean (SD)	39.67 (6.98)
<b>Institution level, n (%)</b>	
Municipal eye disease control center	1 (2.56)
District-level eye disease control center	15 (38.46) <sup>a</sup>
Community health service center	23 (58.97)
<b>Position, n (%)</b>	
Institution leader	7 (17.95)
Department leader	22 (56.41)
Eye disease screening mainstay	10 (25.64)
Years in the current position, mean (SD)	6.73 (5.76)
<b>Resident (n=318)</b>	
Age (years), mean (SD)	68.62 (6.96)
<b>Sex, n (%)</b>	
Male	120 (37.74)
Female	198 (62.26)
<b>Education level, n (%)</b>	
Junior high school and below	216 (67.92)
Senior high school	72 (22.64)
Junior college	21 (6.6)
Undergraduate and above	9 (2.83)
<b>Eye disease, n (%)</b>	
Suspected	73 (22.96)
None	245 (77.04)

<sup>a</sup>One respondent from a district-level eye disease control center quit the experiment because of temporary work arrangements. Therefore, although 16 district-level eye disease control centers were included in our study, only 15 key persons from these institutions finished the questionnaire.

**Table 3.** Preferences for using deep learning in community-based eye disease screening.

Attribute and level	Medical staff <sup>a</sup> OR <sup>b</sup> (95% CI)	Residents OR (95% CI)
<b>Diagnostic technology</b>		
Semiautomated DL <sup>c</sup> model	2.08 (1.71, 2.52) <sup>d</sup>	0.89 (0.68, 1.15)
Fully automated DL model	2.39 (1.97, 2.90) <sup>d</sup>	0.24 (0.20, 0.29) <sup>d</sup>
Fully manual diagnosis	Reference	Reference
<b>Organizational form</b>		
Centralized screening	Reference	Reference
Residents' health self-examination cabin	Not significant	Not significant
Opportunity screening in outpatient <sup>e</sup>	Not significant	Not significant
<b>Missed diagnosis rate</b>		
None	Reference	Reference
5%	Not significant	Not significant
10%	Not significant	Not significant
15%	Not significant	Not significant
20%	Not significant	Not significant
<b>Overdiagnosis rate</b>		
None	Reference	Reference
5%	0.88 (0.68, 1.15)	Not significant
10%	0.61 (0.46, 0.81) <sup>d</sup>	Not significant
15%	0.63 (0.48, 0.83) <sup>f</sup>	Not significant
20%	0.51 (0.38, 0.68) <sup>d</sup>	Not significant
<b>Cost<sup>g</sup></b>		
Free	Reference	Reference
40 CNY	0.61 (0.46, 0.83) <sup>f</sup>	0.75 (0.56, 1.01)
80 CNY	0.47 (0.35, 0.64) <sup>d</sup>	0.56 (0.42, 0.74) <sup>d</sup>
120 CNY	0.39 (0.28, 0.54) <sup>d</sup>	0.82 (0.51, 1.31)
160 CNY	0.27 (0.19, 0.38) <sup>d</sup>	0.78 (0.46, 1.32)
200 CNY	0.21 (0.15, 0.29) <sup>d</sup>	0.57 (0.46, 0.71) <sup>d</sup>
<b>Screening result feedback form</b>		
Screening results	Not significant	0.52 (0.44, 0.61) <sup>d</sup>
Screening results and referral recommendations	Not significant	0.75 (0.65, 0.87) <sup>d</sup>
Screening results, referral recommendations, and oral explanation by GP <sup>h</sup>	Reference	Reference
<b>Screening result feedback efficiency</b>		
Immediately	Reference	Reference
In 2 weeks	0.68 (0.56, 0.82) <sup>d</sup>	Not significant
In 1 month	0.58 (0.48, 0.70) <sup>d</sup>	Not significant

<sup>a</sup>In each grid, an OR value over 1 means that the health care services providers were more inclined to this level, while the value less than 1 means that they disliked this level even more.

<sup>b</sup>OR: odds ratio.



<sup>c</sup>DL: deep learning.

<sup>d</sup> $P < .001$ .

<sup>e</sup>Residents with chronic diseases and other risk factors would be recommended by general practitioners for eye disease screening during their outpatient follow-up.

<sup>f</sup> $P = .001$ .

<sup>g</sup>In 2021, US \$1 = 6.5 CNY.

<sup>h</sup>GP: general practitioner.

## Discussion

### Principal Findings

To the best of our knowledge, this study is the first to quantitatively estimate both medical staff's and residents' preferences for using DL in community-based eye disease screening in the real world. Since one of the most important questions for achieving universal health coverage in a digital world is whether digital technologies help increase the acceptability of health care services [23,24], our study is significant for the transformation, application, and promotion of this new technology. It was based on the multicenter practices of AI-assisted eye disease screening from 34 medical institutions, where both medical staff and residents under investigation had real service experience of AI. We showed that when compared with a fully manual diagnosis, AI technology was more favored by the medical staff, even after adjusting for the impacts of diagnosis accuracy, cost, and efficiency. However, the residents were in disfavor of the AI technology without doctors' supervision. Furthermore, to meet the medical staff's preference, the accuracy of the AI-assisted eye disease screening technology should be much higher than the FDA's standards. On the contrary, accuracy was not a priority for the residents. They prefer to have the doctors involved in the screening process and leave the choice of accuracy to their general practitioners.

The adoption of DL model for community-based eye disease screening is necessary. Before the development of DL model, the screening relied on ophthalmologists heavily, regardless of conducting traditional face-to-face screening or a telemedicine system [25]. At this stage, continuous eye disease screening was not affordable in most of the countries [6] for two reasons. On the one hand, the limited human resources of the ophthalmologists resulted in extremely high screening costs [5]. On the other hand, the organization of the screening was challenging, requiring the coordination of ophthalmologists, community health centers, and residents at the same time [25]. As a result, in Shanghai, before the adoption of the DL model, each community only could provide screening service to approximately 300 residents per year. On the contrary, after the adoption of DL model, as the ophthalmologist resources were no longer the bottlenecks, the screening use volume dramatically increased to 800 residents per community per year.

Accuracy is regarded as one of the most important considerations in the adoption of DL model. When screening populations with a substantial disease, achieving both high sensitivity and specificity is critical in minimizing both false-positive and false-negative results [26]. The previous studies have shown that it is feasible to meet the mandatory level of accuracy as the primary endpoint with a sensitivity of

more than 85% and a specificity of more than 82.5%, which was recommended by the FDA [6,22,27,28]. However, when the DL models were applied in the real world, their accuracy greatly reduced [7]. Therefore, the question is, "what are the medical staff and residents' requirements of the accuracy of AI models in the real world?"

Our study attempted to answer this question from the perspective of medical staff's and residents' preferences in the real-world, community-based eye diseases screening. Although the ideal state is 100% accuracy, under the existing technical conditions, health care service providers must make a trade-off between higher sensitivity and specificity. Both outcomes are important—positive cases should be identified, but this should not come at the cost of overly sensitive screening systems [29].

We showed that if the overdiagnosis rate exceeded 10%, the preferences of the medical staff decreased significantly. Therefore, the specificity of the DL model should be controlled with over 90% accuracy. This does not mean that sensitivity is not important, but rather that the sensitivity standard of the FDA is sufficient. On the one hand, sensitivity is a patient safety criterion, because the primary goal of eye disease screening is to identify the people who are likely to have eye disease and require further evaluation by ophthalmologists [22]. One GP in our study claimed that "the missed diagnosis may harm residents' trust in eye disease screening and reduce their enthusiasm for screening," whereas trust acts as a critical element in medical care [30]. On the other hand, the overdiagnosis rate affects the number of residents who receive an unnecessary referral [22]. A higher overdiagnosis rate means more unnecessary specialist visits, which may lead to unnecessary psychological stress for suspected patients and add further referral costs [31]. Therefore, our results indicated that overdiagnosis would cause resentment from both decision-making and executive agencies.

However, though accuracy is critical for medical staff, results show that the residents do not regard it as a priority. They rather focus on whether the doctors are at the center of medical decision-making [32]. Humans are notoriously poor at comprehending probability and evaluating risk, especially when it pertains to their health or the health of a loved one [33]. In the AI era, although medical knowledge—which forms the basis of decision-making—will be as accessible to the patient as the doctor, most patients need a doctor to understand risk and to communicate this to them [33]. Patients look to the doctors for advice when facing uncertainty in their medical decisions [34]. A study of patient attitudes toward AI use has shown that patients felt their doctors should have the final say in their treatment plans to avoid experiencing the potential harm that might result from mistakes made by health care AI [32].

Therefore, AI tools should be used as decision support tools for human diagnosticians, but not in place of them [35].

When it comes to the other attributes, AI technology with a lower cost and higher feedback efficiency is logically preferable. Cost is an important issue in the adoption of AI-assisted eye disease diagnosis technology. Therefore, it is necessary to conduct health economics evaluation [36]. Fortunately, evidence has shown the cost of screening could be saved by using AI technology, which is mainly attributable to the substantial reduction in human assessment time and workforce without sacrificing screening performance [5].

Traditional ophthalmological diagnosis is heavily dependent on the interpretation of images, which is often subjective and qualitative [37]. Reading these images by trained personnel is neither sustainable nor an efficient use of expertise, and AI technology is essential in facilitating the capture, storage, and interpretation of photographs [17]. From the health system's perspective, the addition of the DL model to fundus photography provides an opportunity to improve this platform for detecting and monitoring retinal diseases on a large scale, and satisfactory results have been obtained [13]. In addition, AI algorithms may bridge the clinical gap [4]. The DL method used for discriminative tasks in ophthalmology, such as diagnosing diabetic retinopathy or age-related macular degeneration, could enhance existing data sets of common and rare ophthalmic diseases without concern for personally identifying information [38]. Other than helping address the limited screening capacity, the DL model may reduce workforce costs and relieve the burden placed on teleophthalmology health care staff [4,39]. The inadequacy of health resources and the vast medical burden

may be important reasons for the rapid acceptance the DL method by medical staff.

Regarding feedback efficiency, recent studies have shown that nearly instantaneous feedback may lead to increased patient compliance [5,18]. The most obvious context for the application of AI-assisted diagnosis technology is in primary eye care where the data to be analyzed are complex, the outcomes are simple and well-defined, and the number of people to process is large [18]. In this context, manual diagnosis requires extensive time and energy, whereas AI can work tirelessly and quickly.

### Limitations

The most obvious limitation of our study was that DCE was conducted only in Shanghai. However, as mentioned, Shanghai is one of the pioneers in eye care digital transformation. Therefore, our study is valuable for other regions of the world. The second limitation was that the residents in our experiment were mainly older adults. However, this was consistent with the population that participated in the community-based eye screening in Shanghai because young people mostly participated in physical examinations at their workplace.

### Conclusion

In conclusion, to meet the actual preferences of medical staff and residents for using AI in the community-based eye disease screening, the DL model under doctors' supervision is strongly recommended, and the specificity of the model should be more than 90%, which is higher than the FDA standard. In addition, digital transformation should help medical staff move away from heavy and repetitive work; however, it should not reduce their involvement in the health care service. Instead, medical staff should spend more time on communicating with residents.

### Acknowledgments

This study was funded by the Shanghai Public Health Three-Year Action Plan (No. GWV10.1-XK6), Science and Technology Commission of Shanghai Municipality (No. 20DZ1100200), Shanghai Hospital Development Center (SHDC12021613), Shanghai Eye Disease Control and Treatment Center (20LC01002), and Shanghai Municipal Health Commission (2022HP61).

We gratefully acknowledge the following investigators who also made contributions to the study: Yajun Peng, Tao Yu, Yao Yin, and Dan Qian.

### Availability of Data and Materials

The data sets used and analyzed during this study are available from the corresponding author on reasonable request.

### Conflicts of Interest

None declared.

### References

1. GBD 2019 Blindness Vision Impairment Collaborators, Vision Loss Expert Group of the Global Burden of Disease Study. Trends in prevalence of blindness and distance and near vision impairment over 30 years: an analysis for the Global Burden of Disease Study. *Lancet Glob Health* 2021 Feb;9(2):e130-e143 [FREE Full text] [doi: [10.1016/S2214-109X\(20\)30425-3](https://doi.org/10.1016/S2214-109X(20)30425-3)] [Medline: [33275950](https://pubmed.ncbi.nlm.nih.gov/33275950/)]
2. Eye care competency framework. World Health Organization. URL: <https://www.who.int/publications/i/item/9789240048416> [accessed 2022-09-09]
3. Cheng C, Wang N, Wong TY, Congdon N, He M, Wang YX, Vision Loss Expert Group of the Global Burden of Disease Study. Prevalence and causes of vision loss in East Asia in 2015: magnitude, temporal trends and projections. *Br J Ophthalmol* 2020 May 28;104(5):616-622. [doi: [10.1136/bjophthalmol-2018-313308](https://doi.org/10.1136/bjophthalmol-2018-313308)] [Medline: [31462416](https://pubmed.ncbi.nlm.nih.gov/31462416/)]

4. Ting DS, Lee AY, Wong TY. An ophthalmologist's guide to deciphering studies in artificial intelligence. *Ophthalmology* 2019 Nov;126(11):1475-1479 [FREE Full text] [doi: [10.1016/j.ophtha.2019.09.014](https://doi.org/10.1016/j.ophtha.2019.09.014)] [Medline: [31635697](https://pubmed.ncbi.nlm.nih.gov/31635697/)]
5. Xie Y, Nguyen QD, Hamzah H, Lim G, Bellemo V, Gunasekeran DV, et al. Artificial intelligence for teleophthalmology-based diabetic retinopathy screening in a national programme: an economic analysis modelling study. *The Lancet Digital Health* 2020 May;2(5):e240-e249. [doi: [10.1016/s2589-7500\(20\)30060-1](https://doi.org/10.1016/s2589-7500(20)30060-1)]
6. Schmidt-Erfurth U, Sadeghipour A, Gerendas BS, Waldstein SM, Bogunović H. Artificial intelligence in retina. *Prog Retin Eye Res* 2018 Nov;67:1-29 [FREE Full text] [doi: [10.1016/j.preteyeres.2018.07.004](https://doi.org/10.1016/j.preteyeres.2018.07.004)] [Medline: [30076935](https://pubmed.ncbi.nlm.nih.gov/30076935/)]
7. Lee A, Yanagihara R, Lee C, Blazes M, Jung H, Chee Y, et al. Multicenter, head-to-head, real-world validation study of seven automated artificial intelligence diabetic retinopathy screening systems. *Diabetes Care* 2021 May;44(5):1168-1175 [FREE Full text] [doi: [10.2337/dc20-1877](https://doi.org/10.2337/dc20-1877)] [Medline: [33402366](https://pubmed.ncbi.nlm.nih.gov/33402366/)]
8. Ryan M. Discrete choice experiments in health care. *BMJ* 2004 Feb 14;328(7436):360-361 [FREE Full text] [doi: [10.1136/bmj.328.7436.360](https://doi.org/10.1136/bmj.328.7436.360)] [Medline: [14962852](https://pubmed.ncbi.nlm.nih.gov/14962852/)]
9. de Bekker-Grob EW, Swait JD, Kassahun HT, Bliemer MC, Jonker MF, Veldwijk J, et al. Are healthcare choices predictable? The impact of discrete choice experiment designs and models. *Value Health* 2019 Sep;22(9):1050-1062 [FREE Full text] [doi: [10.1016/j.jval.2019.04.1924](https://doi.org/10.1016/j.jval.2019.04.1924)] [Medline: [31511182](https://pubmed.ncbi.nlm.nih.gov/31511182/)]
10. Xu Y, Wang Y, Liu B, Tang L, Lv L, Ke X, et al. The diagnostic accuracy of an intelligent and automated fundus disease image assessment system with lesion quantitative function (SmartEye) in diabetic patients. *BMC Ophthalmol* 2019 Aug 14;19(1):184 [FREE Full text] [doi: [10.1186/s12886-019-1196-9](https://doi.org/10.1186/s12886-019-1196-9)] [Medline: [31412800](https://pubmed.ncbi.nlm.nih.gov/31412800/)]
11. Dai L, Wu L, Li H, Cai C, Wu Q, Kong H, et al. A deep learning system for detecting diabetic retinopathy across the disease spectrum. *Nat Commun* 2021 May 28;12(1):3242 [FREE Full text] [doi: [10.1038/s41467-021-23458-5](https://doi.org/10.1038/s41467-021-23458-5)] [Medline: [34050158](https://pubmed.ncbi.nlm.nih.gov/34050158/)]
12. Li F, Wang Y, Xu T, Dong L, Yan L, Jiang M, et al. Deep learning-based automated detection for diabetic retinopathy and diabetic macular oedema in retinal fundus photographs. *Eye (Lond)* 2022 Jul 01;36(7):1433-1441. [doi: [10.1038/s41433-021-01552-8](https://doi.org/10.1038/s41433-021-01552-8)] [Medline: [34211137](https://pubmed.ncbi.nlm.nih.gov/34211137/)]
13. Lin D, Xiong J, Liu C, Zhao L, Li Z, Yu S, et al. Application of Comprehensive Artificial intelligence Retinal Expert (CARE) system: a national real-world evidence study. *The Lancet Digital Health* 2021 Aug;3(8):e486-e495. [doi: [10.1016/s2589-7500\(21\)00086-8](https://doi.org/10.1016/s2589-7500(21)00086-8)]
14. Cen L, Ji J, Lin J, Ju S, Lin H, Li T, et al. Automatic detection of 39 fundus diseases and conditions in retinal photographs using deep neural networks. *Nat Commun* 2021 Aug 10;12(1):4828 [FREE Full text] [doi: [10.1038/s41467-021-25138-w](https://doi.org/10.1038/s41467-021-25138-w)] [Medline: [34376678](https://pubmed.ncbi.nlm.nih.gov/34376678/)]
15. Reed Johnson F, Lancsar E, Marshall D, Kilambi V, Mühlbacher A, Regier DA, et al. Constructing experimental designs for discrete-choice experiments: report of the ISPOR Conjoint Analysis Experimental Design Good Research Practices Task Force. *Value Health* 2013;16(1):3-13 [FREE Full text] [doi: [10.1016/j.jval.2012.08.2223](https://doi.org/10.1016/j.jval.2012.08.2223)] [Medline: [23337210](https://pubmed.ncbi.nlm.nih.gov/23337210/)]
16. Holden RJ, Karsh B. The technology acceptance model: its past and its future in health care. *J Biomed Inform* 2010 Feb;43(1):159-172 [FREE Full text] [doi: [10.1016/j.jbi.2009.07.002](https://doi.org/10.1016/j.jbi.2009.07.002)] [Medline: [19615467](https://pubmed.ncbi.nlm.nih.gov/19615467/)]
17. Li JO, Liu H, Ting DS, Jeon S, Chan RP, Kim JE, et al. Digital technology, tele-medicine and artificial intelligence in ophthalmology: A global perspective. *Prog Retin Eye Res* 2021 May;82:100900 [FREE Full text] [doi: [10.1016/j.preteyeres.2020.100900](https://doi.org/10.1016/j.preteyeres.2020.100900)] [Medline: [32898686](https://pubmed.ncbi.nlm.nih.gov/32898686/)]
18. Lee A, Taylor P, Kalpathy-Cramer J, Tufail A. Machine learning has arrived!. *Ophthalmology* 2017 Dec;124(12):1726-1728. [doi: [10.1016/j.ophtha.2017.08.046](https://doi.org/10.1016/j.ophtha.2017.08.046)] [Medline: [29157423](https://pubmed.ncbi.nlm.nih.gov/29157423/)]
19. Dietvorst BJ, Simmons JP, Massey C. Algorithm aversion: people erroneously avoid algorithms after seeing them err. *J Exp Psychol Gen* 2015 Feb;144(1):114-126. [doi: [10.1037/xge0000033](https://doi.org/10.1037/xge0000033)] [Medline: [25401381](https://pubmed.ncbi.nlm.nih.gov/25401381/)]
20. Kuhfeld WF. Marketing research methods in SAS: experimental design, choice, conjoint, and graphical techniques. SAS Institute Inc. URL: <https://support.sas.com/techsup/technote/mr2010.pdf> [accessed 2022-09-09]
21. Sculpher M, Bryan S, Fry P, de Winter P, Payne H, Emberton M. Patients' preferences for the management of non-metastatic prostate cancer: discrete choice experiment. *BMJ* 2004 Feb 14;328(7436):382 [FREE Full text] [doi: [10.1136/bmj.37972.497234.44](https://doi.org/10.1136/bmj.37972.497234.44)] [Medline: [14751919](https://pubmed.ncbi.nlm.nih.gov/14751919/)]
22. Abramoff MD, Lavin PT, Birch M, Shah N, Folk JC. Pivotal trial of an autonomous AI-based diagnostic system for detection of diabetic retinopathy in primary care offices. *NPJ Digit Med* 2018 Aug 28;1(1):39 [FREE Full text] [doi: [10.1038/s41746-018-0040-6](https://doi.org/10.1038/s41746-018-0040-6)] [Medline: [31304320](https://pubmed.ncbi.nlm.nih.gov/31304320/)]
23. Kickbusch I, Piselli D, Agrawal A, Balicer R, Banner O, Adelhardt M, et al. The Lancet and Financial Times Commission on governing health futures 2030: growing up in a digital world. *The Lancet* 2021 Nov;398(10312):1727-1776. [doi: [10.1016/s0140-6736\(21\)01824-9](https://doi.org/10.1016/s0140-6736(21)01824-9)]
24. Schwalbe N, Wahl B. Artificial intelligence and the future of global health. *The Lancet* 2020 May;395(10236):1579-1586. [doi: [10.1016/s0140-6736\(20\)30226-9](https://doi.org/10.1016/s0140-6736(20)30226-9)]
25. Li R, Yang Z, Zhang Y, Bai W, Du Y, Sun R, et al. Cost-effectiveness and cost-utility of traditional and telemedicine combined population-based age-related macular degeneration and diabetic retinopathy screening in rural and urban China. *Lancet Reg Health West Pac* 2022 Jun;23:100435 [FREE Full text] [doi: [10.1016/j.janwpc.2022.100435](https://doi.org/10.1016/j.janwpc.2022.100435)] [Medline: [35355615](https://pubmed.ncbi.nlm.nih.gov/35355615/)]

26. Gulshan V, Peng L, Coram M, Stumpe MC, Wu D, Narayanaswamy A, et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *JAMA* 2016 Dec 13;316(22):2402-2410. [doi: [10.1001/jama.2016.17216](https://doi.org/10.1001/jama.2016.17216)] [Medline: [27898976](https://pubmed.ncbi.nlm.nih.gov/27898976/)]
27. Bhaskaranand M, Ramachandra C, Bhat S, Cuadros J, Nittala MG, Sadda SR, et al. The value of automated diabetic retinopathy screening with the EyeArt system: A study of more than 100,000 consecutive encounters from people with diabetes. *Diabetes Technol Ther* 2019 Nov;21(11):635-643 [FREE Full text] [doi: [10.1089/dia.2019.0164](https://doi.org/10.1089/dia.2019.0164)] [Medline: [31335200](https://pubmed.ncbi.nlm.nih.gov/31335200/)]
28. Ming S, Xie K, Lei X, Yang Y, Zhao Z, Li S, et al. Evaluation of a novel artificial intelligence-based screening system for diabetic retinopathy in community of China: a real-world study. *Int Ophthalmol* 2021 Apr 03;41(4):1291-1299. [doi: [10.1007/s10792-020-01685-x](https://doi.org/10.1007/s10792-020-01685-x)] [Medline: [33389425](https://pubmed.ncbi.nlm.nih.gov/33389425/)]
29. Tufail A, Kapetanakis VV, Salas-Vega S, Egan C, Rudisill C, Owen CG, et al. An observational study to assess if automated diabetic retinopathy image assessment software can replace one or more steps of manual imaging grading and to determine their cost-effectiveness. *Health Technol Assess* 2016 Dec;20(92):1-72 [FREE Full text] [doi: [10.3310/hta20920](https://doi.org/10.3310/hta20920)] [Medline: [27981917](https://pubmed.ncbi.nlm.nih.gov/27981917/)]
30. Owsley C, McGwin G, Scilley K, Girkin CA, Phillips JM, Searcey K. Perceived barriers to care and attitudes about vision and eye care: focus groups with older African Americans and eye care providers. *Invest Ophthalmol Vis Sci* 2006 Jul 01;47(7):2797-2802. [doi: [10.1167/iovs.06-0107](https://doi.org/10.1167/iovs.06-0107)] [Medline: [16799016](https://pubmed.ncbi.nlm.nih.gov/16799016/)]
31. Wong RL, Tsang C, Wong DS, McGhee S, Lam C, Lian J, et al. Are we making good use of our public resources? The false-positive rate of screening by fundus photography for diabetic macular oedema. *Hong Kong Med J* 2017 Aug 7;23(4):356-364 [FREE Full text] [doi: [10.12809/hkmj166078](https://doi.org/10.12809/hkmj166078)] [Medline: [28684650](https://pubmed.ncbi.nlm.nih.gov/28684650/)]
32. Richardson JP, Smith C, Curtis S, Watson S, Zhu X, Barry B, et al. Patient apprehensions about the use of artificial intelligence in healthcare. *NPJ Digit Med* 2021 Sep 21;4(1):140 [FREE Full text] [doi: [10.1038/s41746-021-00509-1](https://doi.org/10.1038/s41746-021-00509-1)] [Medline: [34548621](https://pubmed.ncbi.nlm.nih.gov/34548621/)]
33. Liu X, Keane PA, Denniston AK. Time to regenerate: the doctor in the age of artificial intelligence. *J R Soc Med* 2018 Apr 12;111(4):113-116 [FREE Full text] [doi: [10.1177/0141076818762648](https://doi.org/10.1177/0141076818762648)] [Medline: [29648509](https://pubmed.ncbi.nlm.nih.gov/29648509/)]
34. Dietvorst BJ, Bharti S. People reject algorithms in uncertain decision domains because they have diminishing sensitivity to forecasting error. *Psychol Sci* 2020 Oct 11;31(10):1302-1314. [doi: [10.1177/0956797620948841](https://doi.org/10.1177/0956797620948841)] [Medline: [32916083](https://pubmed.ncbi.nlm.nih.gov/32916083/)]
35. Sarwar S, Dent A, Faust K, Richer M, Djuric U, Van Ommeren R, et al. Physician perspectives on integration of artificial intelligence into diagnostic pathology. *NPJ Digit Med* 2019 Apr 26;2(1):28 [FREE Full text] [doi: [10.1038/s41746-019-0106-0](https://doi.org/10.1038/s41746-019-0106-0)] [Medline: [31304375](https://pubmed.ncbi.nlm.nih.gov/31304375/)]
36. Ruamviboonsuk P, Chantra S, Seresirikachorn K, Ruamviboonsuk V, Sangroongruangsri S. Economic evaluations of artificial intelligence in ophthalmology. *Asia Pac J Ophthalmol (Phila)* 2021 Jul 13;10(3):307-316. [doi: [10.1097/APO.0000000000000403](https://doi.org/10.1097/APO.0000000000000403)] [Medline: [34261102](https://pubmed.ncbi.nlm.nih.gov/34261102/)]
37. Coyner AS, Campbell JP, Chiang MF. Demystifying the jargon: the bridge between ophthalmology and artificial intelligence. *Ophthalmol Retina* 2019 Apr;3(4):291-293 [FREE Full text] [doi: [10.1016/j.oret.2018.12.008](https://doi.org/10.1016/j.oret.2018.12.008)] [Medline: [31014678](https://pubmed.ncbi.nlm.nih.gov/31014678/)]
38. Burlina PM, Joshi N, Pacheco KD, Liu TYA, Bressler NM. Assessment of deep generative models for high-resolution synthetic retinal image generation of age-related macular degeneration. *JAMA Ophthalmol* 2019 Mar 01;137(3):258-264 [FREE Full text] [doi: [10.1001/jamaophthalmol.2018.6156](https://doi.org/10.1001/jamaophthalmol.2018.6156)] [Medline: [30629091](https://pubmed.ncbi.nlm.nih.gov/30629091/)]
39. Domalpally A, Channa R. Real-world validation of artificial intelligence algorithms for ophthalmic imaging. *Lancet Digit Health* 2021 Aug;3(8):e463-e464 [FREE Full text] [doi: [10.1016/S2589-7500\(21\)00140-0](https://doi.org/10.1016/S2589-7500(21)00140-0)] [Medline: [34325850](https://pubmed.ncbi.nlm.nih.gov/34325850/)]

## Abbreviations

- AI:** artificial intelligence
- DCE:** discrete choice experiment
- DL:** deep learning
- FDA:** Food and Drug Administration
- GP:** general practitioner
- OR:** odds ratio

*Edited by R Kukafka; submitted 12.06.22; peer-reviewed by Z Hou, W Li; comments to author 13.07.22; revised version received 08.08.22; accepted 02.09.22; published 20.09.22*

*Please cite as:*

*Lin S, Li L, Zou H, Xu Y, Lu L*

*Medical Staff and Resident Preferences for Using Deep Learning in Eye Disease Screening: Discrete Choice Experiment  
J Med Internet Res 2022;24(9):e40249*

*URL: <https://www.jmir.org/2022/9/e40249>*

*doi: [10.2196/40249](https://doi.org/10.2196/40249)*

*PMID:*

©Senlin Lin, Liping Li, Haidong Zou, Yi Xu, Lina Lu. Originally published in the Journal of Medical Internet Research (<https://www.jmir.org>), 20.09.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://www.jmir.org/>, as well as this copyright and license information must be included.