How users make judgements about the quality of online health information: a cross-sectional survey study

Wenjing Pian1,2, Laibao Lin1, Baiyang Li3, Chunxiu Qin4 and Huizhong Lin5*

Abstract

Background: People increasingly use the Internet to seek health information. However, the overall quality of online health information remains low. This situation is exacerbated by the unprecedented “infodemic”, which has had negative consequences for patients. Therefore, it is important to understand how users make judgements about health information by applying different judgement criteria.

Objective: The objective of this study is to determine how patients apply different criteria in their judgement of the quality of online health information during the pandemic. In particular, we investigate whether there is consistency between the likelihood of using a particular judgement criterion and its perceived importance among different groups of users.

Methods: A cross-sectional survey was conducted in one of the leading hospitals in a coastal province of China with a population of forty million. Combined-strategy sampling was used to balance the randomness and the practicality of the recruiting process. A total of 1063 patients were recruited for this study. Chi-square and Kruskal–Wallis analyses were used to analyse the survey data.

Results: In general, patients make quality judgement of health information more frequently based on whether it is familiar, aesthetic, and with expertise. In comparison, they put more weights on whether health information is secure, trustworthy, and with expertise when determining its quality. Criteria that were considered more important were not always those with a higher likelihood of being used. Patients may not use particular criteria, such as familiarity, identification, and readability, more frequently than others even if they consider them to be more important than other do and vice versa. Surprisingly, patients with a primary school degree put more weight on whether health information is comprehensive than those with higher degrees do in determining its quality. However, they are less likely to use this guideline in practice.

Conclusions: To the best of our knowledge, this is the first study to investigate the consistency between the likelihood of using certain quality judgement criteria and their perceived importance among patients grouped by different demographic variables and eHealth literacy levels. The findings highlight how to improve online health information services and provide fine-grained customization of information for users.

Keywords: Health information, Information quality, Judgement criteria, Social media, eHealth literacy, Infodemic

© The Author(s) 2022. Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/. The Creative Commons Public Domain Dedication waiver (http://creativecommons.org/publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated in a credit line to the data.
users searched for information via the internet [1], most of whom searched for health information. More than 70% of U.S. adults were reported to perform such searches [2], as did European adults [3] and Chinese citizens [4]. They use the information to facilitate diagnosis, manage conditions, consult with their doctors and make informed medical decisions for themselves and others [5–9]. However, its overall quality remains low [10, 11]. This situation is exacerbated by the COVID-19 pandemic and the unprecedented "infodemic", which is characterized by a flood of rumours and conspiracy theories through various online platforms [12]. During the pandemic, low-quality information resulted in negative consequences for patients due to anxiety about deteriorating conditions, delayed treatments [13] and even death in extreme cases [14]. The quality of information has been defined in different ways. It has been defined as "fitness for use" [15], which is difficult to apply to the judgement of health information due to its vagueness. It has also been defined as meeting or exceeding consumer expectations [16] with a focus on the value of particular information to consumers and the satisfaction of users’ needs. Personal judgement of its quality is a process that involves the subjective evaluation of information based on the user’s own information needs and characteristics [17]. Hence, users may make judgements for different reasons or based on certain criteria (e.g., whether it is objective, comprehensive, and accurate). Therefore, it is important to understand how users use these criteria to make judgements and what factors affect the way they use these criteria.

The objective of this study is to determine how patients apply these different criteria in their judgement of the quality of online health information during the pandemic in terms of how frequently they apply these criteria and how important they consider them. In particular, we investigate whether there is consistency between the likelihood of using a particular criterion and its perceived importance among groups of users with different demographics (age, gender, and educational level) and levels of health literacy.

Models/frameworks of information quality judgement and their criteria

Different conceptualizations of information quality have been proposed in previous literature. The process-oriented model views information quality as a product of measurement where accuracy is guaranteed during the measurement process [18], while the system-oriented model defines information quality as various steps of information collection, retrieval, and display [19]. In comparison, the user-oriented model conceptualizes information quality as users’ subjective perception of whether information satisfies their information needs [20], and information quality judgement is the process by which users apply quality criteria to determine the overall quality of certain information [20]. This study adopts the user-oriented approach to investigate how users make subjective judgements on the quality of health information.

Several user-oriented models were found in previous literature. Bovee, Srivastava, and Mak (2003) proposed an intuitive model of information quality, proposing a three-layer model with integrity, accessibility, interpretability, and relevance in the first layer. Integrity included four criteria: accuracy, completeness, consistency, and existence. Datedness, usefulness, and other criteria were included as the second layer of the relevance criterion. Age and volatility were included within datedness as the third layer of the criteria [21]. Kahn, Strong, and Wang (2002) developed a two-by-two conceptual model that defined two dimensions of information quality: one dimension for product quality and service quality and the other for conforming to specifications and meeting or exceeding consumer expectations. They assigned multiple criteria to these quadrants, such as freedom from errors, conciseness, completeness, and consistency within the quadrant "product quality * conforms to specifications" and timeless and security within the quadrant "service quality * conforms to specifications" [17]. Stvilia et al. (2007) developed a three-category framework, including intrinsic, relational, and reputational categories. Within each category, several criteria were included (e.g., currency and complexity within the intrinsic category and accuracy and naturalness in the relational category) [22].

Sellitto and Burgess (2005) developed a weighted framework consisting of criteria with different weights [5]. Criteria "reputable" was assigned the highest weight, followed by "no advertising" and "creation". Stvilia, Mon, and Yi (2009) extended Stvilia's model (2007) and suggested a model of consumer health information quality with five constructs: accuracy, completeness, authority, usefulness, and accessibility, each of which contained several subcriteria (e.g., credibility and reliability in the construct of accuracy) [23]. Subsequently, Kim and Oh (2009) developed a fine-grained list of criteria for information quality judgement, including clarity, rationality, novelty, and quickness [24], which were not included in the list of Stvilia et al. Sun et al. (2015) further extended the list by identifying additional criteria: solution feasibility, certification, politeness and humour [25]. Other lists of judgement criteria have also been found in the works of other scholars (e.g., see Choi & Shah, 2016 [26]; Emamjome et al., 2013 [27]). They share most of their criteria with the aforementioned lists.
The issue with these studies is that terms were used inconsistently within different lists of criteria. For instance, the consistency criterion from Stvilia’s study was replaced by the term readability in Zhu et al.’s study with the same definition [28], which was further changed to representation interpretability in Ge et al.’s study [29]. Besides, currency was interchanged with timeliness and quickness in different studies [30]. A summary of the aforementioned criteria is presented in Table 1. Recently, Sun et al. (2019) conducted a systematic review of quality judgement of online health information and summarize all the criteria in previous literature into a list of twenty-five criteria [11]. This comprehensive list of criteria was used in this study.

Factors influencing the judgement of health information quality

Several factors have been found to influence the judgement of health information quality. Benotsch, Kalichman, and Weinhardt (2004) conducted a survey of HIV patients to determine how patients evaluated the quality of online health information for certain health websites [18]. Educational and income levels were found to be significantly related to the overall assessment of online health information quality in addition to health literacy (including knowledge) and other psychological factors. Another national survey of Americans’ health information seeking indicated that educational level affected users’ ability to navigate within the online environment to seek health information [31], which affected their judgement of online health information.

Another set of studies focused on the factors that are related to particular criteria for the quality of health information, i.e., trustworthiness and credibility. Age, sex, and education were found to be significantly related to these criteria [32]. Three reviews summarized a list of possible influencing factors found in empirical studies: gender, education, health status, income, age, health literacy, race, and health beliefs. However, health status has been found to be significantly related to trustworthiness and credibility in certain studies but not in others [33–37]. In addition, controversial results on health beliefs have been found (e.g., Atkinson et al. found that people with poor health status trusted online health information more, while Cotten et al. observed the opposite [33–35, 37]). Furthermore, health beliefs have been found to affect the intention to use health information rather than the direct judgement of quality [40]. Therefore, these two factors were left for further investigation, while gender, age, education level, and health literacy were retained in this study. In addition, the health literacy scale was replaced with the eHealth literary scale, which is more appropriate for the online environment. Race was not used in this study since the participants were all Chinese.

Although quite a few studies have focused on the overall judgement of health information quality or examined some of its dimensions/criteria while ignoring other criteria, no previous studies have investigated both how likely users are to use certain quality judgement criteria and how important they consider these criteria. In other words, it remains unclear whether the likelihood of using certain criteria is consistent with the perceived importance of these criteria among users with different demographic characteristics. This study investigated how these demographic variables and eHealth literacy affect these patterns and consistency.

Table 1 Models of information quality judgment and their criteria

| Authors              | Publication date | List of judgement criteria for information quality                                                                 |
|----------------------|------------------|-------------------------------------------------------------------------------------------------------------------|
| Bovee et al [21]     | 2003             | Integrity, accessibility, Interpretable, and Relevance                                                            |
| Kahn et al [17]      | 2002             | Accessibility, appropriate amount of information, believability, completeness, concise representation, consistent   |
|                      |                  | representation, ease of manipulation, free of error, interpretability, objectivity, relevancy, reputation, security, |
|                      |                  | timeliness, understandability, and value-added                                                                   |
| Sellitto & Brugess [5]| 2005             | Authority and currency, accuracy, objectivity, and privacy                                                        |
| Stvilia et al [22]   | 2007, 2008       | Accuracy/validity, cohesiveness, complexity, semantic consistency, structural consistency, currency, informativeness, |
|                      |                  | naturalness, precision, relevance, accessibility, security, and verifiability                                     |
| Stvilia et al [23]   | 2009             | Authoritative, complementarity, privacy, attribution, justifiability, transparency, financial disclosure, and       |
|                      |                  | advertising policy                                                                                               |
| Emamjome et al [27]  | 2013             | Completeness, originality, novelty, accuracy, content quality, verifiability, reliability, amount of data,        |
|                      |                  | relevancy, creditability, user feedback, timeliness, understandability, value-added, conciseness, consistency,     |
|                      |                  | and accessibility                                                                                               |
| Weiskopf & Wen       | 2013             | Completeness, correctness, concordance, plausibility, and currency                                               |
| Sun et al [25]       | 2015             | Accuracy, specificity, objectivity, completeness, relevance, language expression, valuable words, novelty,         |
|                      |                  | understandability, profession, originality, external links, external certification, quickness, interactivity,      |
|                      |                  | effectiveness, and solution feasibility                                                                         |
| Choi & Shah [26]     | 2016             | Responsiveness, alternatively, completeness, emotional support, verifiability, and trustworthy                    |
Research methods

Research setting and sampling procedure
The research participants were recruited from one of the leading hospitals in a coastal province of China with a population of forty million for the following reasons. First, the hospital is listed in the top 100 hospitals of China¹ and is the most comprehensive tertiary hospital in that province, with more than two million visits annually. Second, the hospital provides health services to patients from each region of the province, which increases the representativeness of the sample.

Combined-strategy sampling was used to balance the randomness and the practicality of the recruiting process [41]. Given the large number of departments within the hospital, randomization stratified for inpatient and outpatient divisions was undertaken to select the target departments where the patients were recruited. Within each department, a systematic sampling procedure was performed with a randomized seeding number assigned to recruit patients. A paper-based survey was administered, and the data were collected from September to October 2021. The participants (who were 18 years old or above) were approached by nurses in person and briefed with the objective of this study. They were asked to sign an informed consent form before the survey began. In addition, they were able to withdraw from the survey at any time, and anonymity was ensured. Surveys were conducted following confirmation of informed consent which was signed prior to the survey questions. This consent process was approved by the Institutional Review Board (IRB) of Fujian Medical University Union Hospital (IRB No. 2022KY004). All methods were performed in accordance with the relevant guidelines and regulations (Declaration of Helsinki).

To determine the sample size, two guidelines were considered. According to the COSMIn (Consensus-based Standards for the selection of health Measurement Instruments) criteria, the ideal sample size for a survey should be no less than ten times the number of survey items [42]: In addition, a power analysis was performed to calculate the minimum sample size for a study with variance analysis of between-group differences. The minimum sample size was 231, with a predefined effect size of 0.40 (which is considered large for six groups), error margin of 0.01, and power of 0.99 for a priori analysis. Therefore, a sample of one thousand participants was considered sufficient for this study (the study included forty survey questions).

¹ https://www.cn-healthcare.com/articlewm/20220721/content-1404153.html

Measure and data analysis
The demographics of the participants were collected and categorized as follows. In terms of age, the participants were categorized into five groups from 18 years old to 60 years and above (18–29; 30–39; 40–49; 50–59; 60 and above). Educational levels were also categorized into five groups from junior high school or below to Ph.D. degree (junior high school or below; high school; undergraduate; master’s; Ph.D.). When gender was considered, two groups (female; male) were used to categorize the participants (no participant indicated gender as unclear or transgender).

In addition to the demographic variables, two sets of questions were used in this study. The first set was items to assess the intention of using a particular criterion in the judgement of online health information and the perceived importance of that criterion, adopted from Sun et al’s review article [11]. For each criterion, the participant was asked to indicate whether he or she used it in evaluating the quality of online health information. If yes, the participants proceeded to indicate how important the criterion was using a five-point Likert scale. The second set was the eHealth Literacy scale (eHEALS) to assess consumers’ ability to engage in eHealth [43]. It was used to replace the traditional health literacy scales since the context of this study was online health information seeking and evaluation. This scale has been validated by various empirical studies and shows good psychometric properties. The distribution of the eHealth literacy scores among the participants was skewed ($S = -0.446$), with a median total score of 30 for the eHEALS scale (range: 8–40). Since there was no consensus on the cut-off score of the eHEALS scale to categorize people into high and low eHealth literacy, we used the median score to divide the participants into these two user groups.

All the measures used in the study were forwards and backwards translated following the COSMIN guidelines. Each type of translation involved two bilingual translators whose mother tongue was the original language: one was the domain expert, and the other was naive about the domain. The backwards translation was compared with the original questions, and discrepancies were resolved by discussion among the authors and translators.

Two kinds of statistical analyses were adopted in this study. To determine whether a particular criterion was used by a patient, chi-square analysis was conducted to compare the group differences in the frequency of criterion use since it was a binary variable. For the perceived importance of each criterion, nonparametric analysis (Kruskal–Wallis test and 1-way ANOVA for post hoc test) was used to compare the group differences (the normal distribution assumption was not guaranteed). It should be noted that in the analysis of importance,
participants who indicated that they did not use a particular criterion were excluded since the measurement of the five-point Likert scale assumed that the participants used the criterion regardless of how important they considered it (even if it was considered unimportant, the value was 1 in the scale). Participants who did not use the criterion would introduce bias into the analysis interpretation of the results. The scores of the eHealth literacy scale items were summed to indicate participants’ level of literacy when they conducted online health information seeking and evaluation [44]. The data analysis was conducted using SPSS 26 software.

Results
Descriptive results of participants’ demographics
A total of 1500 patients were invited to participate in this study, of whom 1063 accepted the invitation and finished the survey (acceptance rate: 70.9%). Among them, nearly two-thirds were under forty years old, while less than 10% were over sixty. In terms of educational level, more than half of the participants held a college degree, and more than one-quarter of them had received higher degrees (master’s and Ph.D.). In comparison, less than 20% held a primary or high school degree. Details are shown in Table 2.

Regarding the eHealth literacy level of the participants, the distribution was skewed ($S = -0.446$) with a median total score of 30 for the eHEALS scale (range: 8–40). This is similar to recent survey results of eHealth literacy from other countries using the same scale [45, 46]. Since there was no consensus on the cut-off score of the eHEALS scale to categorize people into high and low eHealth literacy, we used the median score to divide the participants into these two user groups.

The frequency of the use of judgement criteria on health information quality and the perceived importance among different user groups
The statistical results of the frequency of use for the judgement criteria for health information quality among different user groups and the perceived importance are presented according to the demographic variables and eHealth literacy levels mentioned above. The details of the post hoc analysis can be found in the appendix.

eHealth literacy
Patients with high eHealth literacy were more inclined to use the criteria of transparency and usefulness than those with low eHealth literacy. However, there were no differences in the perceived importance of these two criteria between these two groups of patients. In contrast, for the criteria of familiarity, accessibility, aesthetics, comprehensiveness, and practicality, no significant differences were found in their frequencies of use between the two groups of participants, although they perceived them to have significantly different importance. The nonparametric statistical analysis (K-W) showed that patients with high eHealth literacy considered these criteria to be more important than did those with low eHealth literacy. For the criteria of identification, interactivity and balance, patients with high eHealth literacy considered these criteria to be more important and therefore were more inclined to use them than those with low eHealth literacy. No differences in the frequency of use and perceived importance were found for the remaining criteria.
between patients with high and low eHealth literacy. Details are shown in Table 3.

**Age**

For the criteria of familiarity, identification, aesthetics, and anonymity, no significant differences were found in the likelihood of using these criteria in judging the quality of online health information among different age groups. However, these criteria were considered to have different levels of importance among these groups. The nonparametric statistical analysis (K-W) showed that for the criteria of familiarity, identification, and aesthetics, patients from 18 to 29 years old considered them less important than did those from 30 to 39, 40 to 49, and 50–59 years old. In comparison, the difference between any other two age groups was not significant. In terms of the criterion of anonymity, only patients from 18 to 29 years old considered it less important than those from 40 to 49 years old did. Surprisingly, patients over 60 years old were least likely to use this criterion. Patients from 30 to 39 years old were found to have the second-lowest likelihood of adopting this criterion. In addition, patients from 18 to 29 years old were less likely than those aged 40 to 59 years old to apply this criterion to the quality judgment of online health information. For the other criteria, no significant differences were found in either the likelihood of use or the perceived importance among patients in different age groups. Details are shown in Table 4.

**Educational level**

A similar situation was found among patients with different educational levels. For the criteria of expertise, familiarity, aesthetics, and security, patients with different educational levels were found to have no significant differences in their inclination to use these criteria. However, they assigned different importance to these criteria. The expertise criterion was considered more important by patients with a master’s degree than by patients with...
either a bachelor’s or a high school degree. Similarly, the security criterion was considered more important by patients with a master’s degree than by those with either a bachelor’s or a high school degree. In contrast, familiarity was considered more important by patients with either a bachelor’s or high school degree than by those with a master’s degree. Furthermore, patients with high school degrees considered familiarity even more important than those with a bachelor’s degree. For the criteria of objectivity, identification, and accuracy, no significant differences in their perceived importance were found among patients with different educational levels. However, their likelihood of being used differed significantly among these patient groups. For the objectivity criterion, patients with primary school degrees were more inclined to use this criterion than those with higher degrees when making a quality judgement about online health information. Surprisingly, patients with a Ph.D. degree were more inclined to ignore this criterion than those with either a bachelor’s or a master’s degree. It should be noted that patients with a primary school degree were less likely to use this criterion for quality judgement than patients with any other educational level. In sharp contrast, they considered this criterion to be more important than those with a bachelor’s degree. Details are shown in Table 5.

### Gender

For the criteria of trustworthiness, relevance, believability, readability, practicality, completeness, anonymity, and security, female patients considered them to be more important than male patients. However, female patients were not more inclined to use these criteria than male patients in judging the quality of online health information. No significant differences in the likelihood of using the remaining criteria and their perceived importance were found between these two groups of patients. Details are shown in Table 6.

---

**Table 4** Chi-square and K-W tests for patients in different age groups

|                                | Chi-Square test (for frequency of use) | Kruskal–Wallis test (for perceived importance) |
|--------------------------------|----------------------------------------|-----------------------------------------------|
|                                | Value df Asymp.Sig                      | Value df Asymp.Sig |
|                                |                                        | $\chi^2$ df Asymp.Sig |
| Trustworthiness                | 1.971 4 0.741                           | 4.336 4 0.362 |
| Expertise                      | 2.173 4 0.704                           | 5.327 4 0.255 |
| Objectivity                    | 1.881 4 0.758                           | 1.028 4 0.906 |
| Transparency                   | 1.093 4 0.895                           | 2.645 4 0.619 |
| Popularity                     | 4.299 4 0.367                           | 2.931 4 0.57 |
| Understandability              | 15.218 4 0.004$^b$                     | 7.74 4 0.102 |
| Relevance                      | 4.246 4 0.374                           | 6.033 4 0.197 |
| Familiarity                    | 4.092 4 0.394                           | 13.377 4 0.01$^a$ |
| Accessibility                  | 0.125 4 0.998                           | 7.387 4 0.117 |
| Identification                 | 1.370 4 0.849                           | 14.91 4 0.005$^b$ |
| Believability                  | 3.369 4 0.498                           | 4.798 4 0.309 |
| Accuracy                       | 0.801 4 0.938                           | 4.237 4 0.375 |
| Readability                    | 4.525 4 0.340                           | 4.027 4 0.402 |
| Currency                       | 3.549 4 0.470                           | 3.395 4 0.494 |
| Navigability                   | 1.609 4 0.807                           | 6.025 4 0.197 |
| Aesthetics                     | 2.173 4 0.704                           | 33.824 4 0.000$^c$ |
| Interactivity                  | 2.250 4 0.690                           | 2.075 4 0.722 |
| Comprehensiveness              | 2.603 4 0.626                           | 2.827 4 0.587 |
| Practicality                   | 0.969 4 0.914                           | 1.232 4 0.873 |
| Completeness                   | 4.299 4 0.367                           | 4.503 4 0.342 |
| Usefulness                     | 2.196 4 0.700                           | 4.132 4 0.388 |
| Balanced                       | 7.756 4 0.101                           | 4.249 4 0.373 |
| Anonymity                      | 3.623 4 0.459                           | 12.744 4 0.013$^a$ |
| Security                       | 2.840 4 0.585                           | 3.957 4 0.412 |
| Learntability                  | 4.655 4 0.325                           | 1.147 4 0.887 |

$^a$ indicates the significance level < 0.05  
$^b$ < 0.01  
$^c$ < 0.001
Discussion

To our knowledge, this is the first study to investigate the consistency between the likelihood of using particular quality judgement criteria and their perceived importance among different groups of patients when making quality judgements about online health information. It was found that for particular criteria such as familiarity, identification, and readability, patients in one demographic group may not use them more frequently than other groups even if they consider these criteria more important than the other groups do, while patients in particular groups may use these criteria more frequently even if they do not consider them more important than other groups do. Furthermore, it is surprising that patients with a primary school degree considered the comprehensiveness criterion to be more important than those with a bachelor’s degree but were less likely to use it in practice, which is counterintuitive to common sense. In the next sections, the results are interpreted and discussed with both theoretical and practical implications.

Theoretical implications

The inconsistency between the likelihood of use and perceived importance indicates that patients may use different sets of quality judgement criteria in different ways. The criteria of popularity, currency, and navigability are those in which no significant difference was found in either the likelihood of use or perceived importance among the patients grouped by all the demographic variables and eHealth literacy. This suggests that these criteria may be independent of patients’ demographic and eHealth literacy backgrounds. Patients share similar attitudes regarding their importance and how to use them. The underlying reason could be that applying these criteria to the quality judgement of online health information does not involve patients’ knowledge, experience, and other cognitive features, which are influenced by

Table 5 Chi-square and K-W tests for patients with different educational levels

| Criteria              | Chi-Square test (for frequency of use) | Kruskal–Wallis test (for perceived importance) |
|-----------------------|----------------------------------------|-----------------------------------------------|
|                       | Value df Asymp.Sig                      | χ² df Asymp.Sig                               |
| Trustworthiness       | 4.042 4 0.400                           | 9.303 4 0.054                                 |
| Expertise             | 5.032 4 0.284                           | 18.212 4 0.001                               |
| Objectivity           | 15.192 4 0.004b                         | 6.544 4 0.162                                 |
| Transparency          | 1.949 4 0.745                           | 4.601 4 0.331                                 |
| Popularity            | 4.432 4 0.351                           | 8.07 4 0.089                                  |
| Understandability     | 7.272 4 0.122                           | 1.466 4 0.833                                 |
| Relevance             | 6.661 4 0.155                           | 3.647 4 0.456                                 |
| Familiarity           | 5.032 4 0.284                           | 12.408 4 0.015a                              |
| Accessibility         | 3.337 4 0.503                           | 2.555 4 0.635                                 |
| Identification        | 11.073 4 0.026a                         | 6.133 4 0.189                                 |
| Believability         | 2.777 4 0.596                           | 3.863 4 0.425                                 |
| Accuracy              | 10.340 4 0.035a                         | 5.299 4 0.258                                 |
| Readability           | 5.753 4 0.218                           | 3.194 4 0.526                                 |
| Currency              | 3.095 4 0.542                           | 5.937 4 0.204                                 |
| Navigability          | 5.753 4 0.218                           | 1.809 4 0.771                                 |
| Aesthetics            | 6.449 4 0.168                           | 14.454 4 0.006b                              |
| Interactivity         | 8.492 4 0.075                           | 5.685 4 0.224                                 |
| Comprehensiveness     | 22.505 4 0.000c                         | 11.543 4 0.021b                              |
| Practicality          | 5.940 4 0.204                           | 2.978 4 0.562                                 |
| Completeness          | 4.432 4 0.351                           | 0.44 4 0.979                                  |
| Usefulness            | 3.331 4 0.504                           | 1.765 4 0.779                                 |
| Balanced              | 6.389 4 0.172                           | 1.321 4 0.858                                 |
| Anonymity             | 2.289 4 0.683                           | 0.724 4 0.948                                 |
| Security              | 2.538 4 0.638                           | 12.353 4 0.015a                              |
| Learnability          | 11.227 4 0.024a                         | 2.058 4 0.725                                 |

* indicates the significance level < 0.05
b < 0.01
< 0.001
their age, education, health literacy, and gender. Another possible reason could be that there are universal prototypes for popularity, currency, and navigability among all patients and even the entire population. Therefore, demographic variables may not affect whether these criteria are used and how important they are perceived to be. Unfortunately, no previous studies were found in this area, and further investigation is required to test these possible explanations.

In contrast, the perceived importance of criteria does not always guarantee the inclination to use these criteria. For example, patients with a master’s degree assigned more weight to the expertise criterion than those with either a high school or bachelor’s degree. However, in practice, they did not show a preference for this particular criterion over the other two groups of patients. The reason could be that the more education patients receive, the more attention they give to features that are closely related to the information content. Hence, the perception of expertise and security conveyed by the information is more attractive to well-educated patients. In comparison, for the criterion of aesthetics, the opposite was observed. The same reasoning could be applied to the perceived importance of the criterion of familiarity among different age groups. As patients get older, they may have fewer cognitive resources and may rely on more superficial features of online health information, such as the heuristic of familiarity, to assess its overall quality.

It should be noted that for the criterion of comprehensiveness, although patients with a primary school degree considered it more important than those with a bachelor’s degree, they were actually more likely to ignore it when making a quality assessment of online health information. It is common sense that if one group of people considers particular criteria more important than another group, they tend to use these criteria more frequently. This counterintuitive finding may be due to the irrational reasoning and decision-making of these patients during a crisis (i.e., the COVID-19 pandemic and infodemic). During the infodemic, patients are likely

### Table 6 Chi-square and K-W tests for patients in different genders

| Criteria               | Chi-Square test (for frequency of use) | Kruskal–Wallis test (for perceived importance) |
|-----------------------|----------------------------------------|-----------------------------------------------|
|                       | Value | df | Asymp. Sig | χ² | df | Asymp. Sig |
| Trustworthiness       | 0.956 | 1  | 0.328      | 9.308 | 1  | 0.002b |
| Expertise             | 0.072 | 1  | 0.788      | 2.745 | 1  | 0.098 |
| Objectivity           | 1.035 | 1  | 0.309      | 1.928 | 1  | 0.165 |
| Transparency          | 0.052 | 1  | 0.820      | 3.403 | 1  | 0.065 |
| Popularity            | 0.437 | 1  | 0.509      | 0.044 | 1  | 0.834 |
| Understandability     | 0.128 | 1  | 0.721      | 0.296 | 1  | 0.586 |
| Relevance             | 1.015 | 1  | 0.314      | 4.485 | 1  | 0.034b |
| Familiarity           | 0.072 | 1  | 0.788      | 1.492 | 1  | 0.222 |
| Accessibility         | 0.109 | 1  | 0.742      | 3.039 | 1  | 0.081 |
| Identification        | 0.145 | 1  | 0.703      | 1.239 | 1  | 0.266 |
| Believability         | 1.092 | 1  | 0.296      | 14.725 | 1  | 0.000c |
| Accuracy              | 0.000 | 1  | 0.988      | 2.101 | 1  | 0.147 |
| Readability           | 0.005 | 1  | 0.946      | 4.767 | 1  | 0.029a |
| Credibility           | 0.658 | 1  | 0.417      | 3.626 | 1  | 0.057 |
| Navigability          | 0.437 | 1  | 0.509      | 2.179 | 1  | 0.14 |
| Aesthetics            | 0.327 | 1  | 0.567      | 0.297 | 1  | 0.586 |
| Interactivity         | 0.377 | 1  | 0.539      | 1.235 | 1  | 0.266 |
| Comprehensiveness     | 1.777 | 1  | 0.183      | 0.333 | 1  | 0.564 |
| Practicality          | 0.635 | 1  | 0.426      | 7.886 | 1  | 0.005b |
| Completeness          | 0.635 | 1  | 0.426      | 4.418 | 1  | 0.036a |
| Usefulness            | 0.025 | 1  | 0.875      | 0.006 | 1  | 0.937 |
| Balanced              | 0.915 | 1  | 0.339      | 0.606 | 1  | 0.436 |
| Anonymity             | 0.046 | 1  | 0.830      | 5.423 | 1  | 0.024a |
| Security              | 0.008 | 1  | 0.931      | 3.982 | 1  | 0.046a |

* indicates the significance level < 0.05
b < 0.01
c < 0.001
to be influenced by the risk situation and emotion [47, 48], which further biases their rational thinking process and affects their quality judgement. Therefore, they may not naturally rely on the criteria they consider important [49], which contributes to the inconsistency.

Practical implications
The findings of this study provide guidelines on how online health information should be presented to facilitate users’ quality judgement, which further helps them use information appropriately.

First, online health information services require fine-grained customization to facilitate the evaluation of users with different demographic characteristics. As patients can be grouped into fine-grained user groups by combining these demographic variables, this study provides practical guidelines on prioritizing detailed user groups during the customization of online health information services. There are different priority sequences for these user groups to facilitate their quality judgement.

Second, the health literacy of patients should be given more attention when attempting to motivate patients to find higher-quality health information. As one of the main drivers of health behaviour change [50], this may influence patients’ willingness to follow certain online suggestions and perform particular types of protective behaviours, even if these suggestions are of low quality. Therefore, it is necessary for scholars to develop more efficient interventions to increase patients’ health literacy.

Research limitations and directions for future research
Several limitations deserve attention when interpreting the results of this study. First, only one of the leading hospitals of the coastal province with a population of forty million was chosen as the research context for this study. Hence, patients from other hospitals in this coastal province were not included in this study, nor were patients from different provinces of the country. However, this hospital was listed in the top 100 hospitals of China and was the top hospital with more than two million visits annually. Second, the number of patients who did not use health information was perceived as more important by patients with a primary school degree than by those with a bachelor’s degree, although it was less likely to be used. Possible reasons for these findings include the different nature of these criteria and the existence of stable prototypes for particular criteria by certain patient groups does not guarantee a higher likelihood of using these criteria in practice and vice versa. The criterion of comprehensiveness was perceived as more important by patients with a primary school degree than by those with a bachelor’s degree, although it was less likely to be used. Possible reasons for these findings include the different nature of these criteria and the existence of stable prototypes for a certain set of criteria. The findings highlight how to improve online health information services and provide fine-grained customization of information for users to make judgements easier and faster.

Conclusion
To the best of our knowledge, this is the first study to investigate the consistency between the likelihood of using certain quality judgement criteria and their perceived importance among patients grouped by different demographic variables and eHealth literacy levels. It was found that the patterns are not always consistent among patients of different ages, genders, educational levels and eHealth literacy levels. The perceived greater importance of particular criteria by certain patient groups does not guarantee a higher likelihood of using these criteria in practice and vice versa. The criterion of comprehensiveness was perceived as more important by patients with a primary school degree than by those with a bachelor’s degree, although it was less likely to be used. Possible reasons for these findings include the different nature of these criteria and the existence of stable prototypes for a certain set of criteria. The findings highlight how to improve online health information services and provide fine-grained customization of information for users to make judgements easier and faster.

Supplementary Information
The online version contains supplementary material available at https://doi.org/10.1186/s12889-022-14418-9.

Acknowledgements
The authors would like to thank the Editors in Chief, the handling Editor, and the anonymous reviewers for their very insightful and constructive comments. The authors would also like to thank Dr. Gaohui Cao for his suggestions on this manuscript, and Prof. Feicheng Ma and Gang Li for their financial support.

Authors’ contributions
Wenjing Pian and Laibao Lin wrote the main manuscript text. Baiyang Li and Chunxiu Qin were responsible for data analysis. Huizhong Lin played a guiding role in the work and was responsible for ensuring that the descriptions were accurate.
are accurate. All authors reviewed the manuscript. All authors have read and approved this manuscript.

**Funding**

Data collection and preliminary analysis were sponsored by National Natural Science Foundation of China (No: 71904028). Scientific and Technological Innovation 2030-New Generation of Artificial Intelligence Major Project (No:2020AAA0108505), and National Natural Science Foundation of China (No. 71921002). The funders had no role in study design, data collection and analysis, decision to publish, or reparation of the manuscript.

**Availability of data and materials**

The datasets generated and/or analysed during the current study are not publicly available due ethical permissions not allowing data sharing, but they are available from the corresponding author on reasonable request.

**Declarations**

**Ethics approval and consent to participate**

This study was approved by the Institutional Review Board (IRB) of Fujian Medical University Union Hospital. Surveys were conducted following confirmation of informed consent which was recorded verbally prior to the survey questions. This consent process was approved by the Institutional Review Board (IRB) of Fujian Medical University Union Hospital.

**Consent for publication**

Not applicable.

**Competing interests**

The authors declare no competing interests.

**Authors details**

1. School of Economics and Management, Fuzhou University, Fuzhou city, China. 2. Center for Studies of Information Resources, Wuhan University, Wuhan city, China. 3. Laboratory of Data Intelligence and Interdisciplinary Innovation, Nanjing University, Nanjing city, China. 4. Department of Information Management, School of Economics and Management, XIDIAN University, Xi’an city, China. 5. Department of Cardiology, Fujian Medical University Union Hospital, Fuzhou city, China.

**Received:** 23 March 2022  **Accepted:** 21 October 2022  **Published online:** 01 November 2022

**References**

1. Kemp S. Digital in 2018: World’s internet users pass the 4 billion mark (2018) [Internet]. 2022 [cited 2022 Mar 6]. Available from: https://weare

2. Pew Research Center. Health online (2013). 2022 [Cited 2022 Mar 6]. Avail-

3. Emamjome F, Rabaa'i A, Gable G, Bandara W. Information quality in online Q&A services? J Inf Sci. 2016;43(3):424–35. https://doi.org/10.5465/amr.1994.9412271805.

4. Reeves CA, Bednar DA. Defining quality: alternatives and implications. Acad Manage Rev. 1994;19(3):419–45. https://doi.org/10.5465/amr.1994.9412271805.

5. Kim S, Oh S. Users’ relevance criteria for evaluating answers in a social Q&A site. J Am Soc Inf Sci Tec. 2009;60(4):716–27. https://doi.org/10.1002/asi.20652.

6. Bovee M, Srivastava RP, Mak B. A conceptual framework and belief-function approach to assessing overall information quality. In: J Intell Syst. 2003;18(1):51–74. https://doi.org/10.1007/BF00121074.

7. Stvilija B, Gasser L. Analyzing the quality of answers in social Q&A sites. Proceedings of the 2015 Pacific Asia Conference on Information Systems. Singapore; 2015. https://aisel.aisnet.org/paci2015/15.

8. Chou E, Shah C. Asking for more than an answer: What do asks expect in online Q&A services? J Inf Sci. 2016;43(3):424–35. https://doi.org/10.1177/0165551516645330.

9. Emajome F, Raba’i A, Gable G, Bandara W. Information quality in social media: a conceptual model. In: Proceedings of the 17th Pacific Asia Conference on Information Systems (PACIS). Jeju Island, 2013. https://aisel.aisnet.org/paci2013/72.

10. Zhi B, Bemhard D, Gurevych I. A multi-dimensional model for assessing the quality of answers in social Q&A sites. Darmstadt, Germany: UKP Lab, Technische Universität Darmstadt; 2009.
29. Ge M, Helfert M, Jannach D. Information quality assessment: Validating measurement dimensions and processes. Proceedings of the 2011 European Conference on Information Systems. Vol. 75. Queensland; 2011. https://aisel.aisnet.org/ecnv2011/75.

30. Fu H, Oh S. Quality assessment of answers with user-identified criteria and data-driven features in social Q&A. Inform Process Manag. 2019;56(1):14–28. https://doi.org/10.1016/j.ipm.2018.08.007.

31. Arora NK, Hesse BW, Rimer BK, Viswanath K, Clayman ML, Coyle RT. Frustrated and confused: the American public rates its cancer-related information-seeking experiences. J Gen Intern Med. 2008;23(3):223–8. https://doi.org/10.1007/s11606-007-0406-y (Epub 2007 Oct 6. PMID: 17922166; PMCID: PMC2394651).

32. Hesse BW, Nelson DE, Kreps GL, Croyle RT, Arora NK, Rimer BK, Viswanath K. Trust and sources of health information: the impact of the Internet and its implications for health care providers: findings from the first Health Information National Trends Survey. Arch Intern Med. 2005;165(22):2618–24. https://doi.org/10.1001/archinte.165.22.2618 (PMID: 16344419).

33. Atkinson NL, Saperstein SL, Pleis J. Using the internet for health-related activities: findings from a national probability sample. J Med Internet Res 2009;11(1):e4 [FREE Full text] [https://doi.org/10.2196/jmir.1035] [Medline: 19275980].

34. Hou J, Shim M. The role of provider-patient communication and trust in online sources in Internet use for health-related activities. J Health Commun. 2010;15(Suppl 3):186–99. https://doi.org/10.1080/10810730.2010.522691 [Medline: 21154099].

35. Rice RE. Influences, usage, and outcomes of Internet health information searching: multivariate results from the Pew surveys. Int J Med Inform. 2006;75(1):34–49. https://doi.org/10.1016/j.ijmedinf.2005.07.032 (Medline: 16125433).

36. Koch-Weser S, Bradshaw YS, Guaitani L, Gallagher SS. The Internet as a health information source: findings from the 2007 Health Information National Trends Survey and implications for health communication. J Health Commun. 2010;15(Suppl 3):279–93. https://doi.org/10.1080/10810730.2010.522700 [Medline: 21154099].

37. Ye Y. Correlates of consumer trust in online health information: findings from the Health Information National Trends Survey. J Health Commun. 2011;16(1):34–49. https://doi.org/10.1080/10810730.2010.529491 [Medline: 21086209].

38. Cotten SR, Gupta SS. Characteristics of online and offline health information seekers and factors that discriminate between them. Soc Sci Med. 2004;59(9):1795–806. https://doi.org/10.1016/j.socscimed.2004.02.020 [Medline: 15312915].

39. Miller LM, Bell RA. Online health information seeking: the influence of age, information trustworthiness, and search challenges. J Aging Health. 2012;24(3):525–41. https://doi.org/10.1177/0898264311428167 [Medline: 22187092].

40. Risker DC. The health belief model and consumer information searches: toward an integrated model. Health Mark Q. 1996;13(3):13–26. https://doi.org/10.1080/10448439609525063 (PMID: 10158485).

41. Gravetter FJ, Forzano LAB. Research methods for the behavioral sciences. New York: Cengage Learning. 2018.

42. Mokkink LB, Prinsen CAC, Patrick DL, Alonso J, Bouter LM, De Vet, et al. COSMIN Study Design checklist for Patient-reported outcome measures. J Med Internet Res. 2006;8(4):e27. https://doi.org/10.2196/jmir.8.4.27 PMID: 17213046; PMCID: PMC1794004.

43. Norman CD, Skinner HA. eHEALS: The eHealth Literacy Scale. J Med Internet Res. 2006;8(4):e27. https://doi.org/10.2196/jmir.8.4.e27 PMID: 17213046; PMCID: PMC1794004.

44. Van der Vaart R, van Deursen AJ, Drossaert CH, Taal E, van Dijk JA, van de Laar MA. Does the eHealth Literacy Scale (eHEALS) measure what it intends to measure? Validation of a Dutch version of the eHEALS in two adult populations. J Med Internet Res. 2011;13(4):e86. https://doi.org/10.2196/jmir.1344.0 PMID: 21701338; PMCID: PMC3222202.

45. Baek JH, Soares GH, da Rosa GC, Mialhe FL, Bazevic MGH, Michel-Crosato E. Network analysis and psychometric properties of the Brazilian version of the eHealth Literacy Scale in a dental clinic setting. Int J Med Inform. 2021;153:104532. https://doi.org/10.1016/j.ijmedin.2021.104532 (Epub 2021 Jul 17 PMID: 342098425).

46. Kim H, Yang E, Ryu H, Kim HJ, Jang SJ, Chang SJ. Psychometric comparisons of measures of eHealth literacy using a sample of Korean older adults. Int J Older People Nurs. 2021;16(3):e12369. https://doi.org/10.1111/ipon.12369 (Epub 2021 Feb 1. PMID: 33527701).

47. Xiaofei Z, Guo X, Ho SY, Lai KH, Vogel D. Effects of emotional attachment on mobile health-monitoring service usage: An affect transfer perspective. Inform Manage-Amster. 2021;58(3):103512. https://doi.org/10.1016/j.im.2020.103512.

48. Meng F, Zhang X, Liu L, Ren C. Converting readers to patients? From free to paid knowledge-sharing in online health communities. Inform Process Manag. 2021;58(3):102490. https://doi.org/10.1016/j.ipm.2021.102490.

49. Qin Q, Ke Q, Du JT, Xie Y. How Users’ Gaze Behavior Is Related to Their Quality Evaluation of a Health Website Based on HONcode Principles? Data Inf Manag. 2021;5(1):75–85. https://doi.org/10.2478/dim-2020-0045.

50. Li S, Jiang Q, Zhang P. Factors influencing the health behavior during public health emergency: A case study on norovirus outbreak in a university. Data Inf Manag. 2020;5(1):27–39. https://doi.org/10.2478/dim-2020-0022.

51. Pian W, Song S, Zhang Y. Consumer health information needs: A systematic review of measures. Inform Process Manag. 2020;57(2):102077. https://doi.org/10.1016/j.ipm.2019.102077.

52. Chi J, Pian W, Zhang S. Consumer health information needs: A systematic review of instrument development. Inform Process Manag. 2020;57(6):102376. https://doi.org/10.1016/j.ipm.2020.102376.

53. Jiang Q, Zhang Y, Ran W. Chatbot as an emergency exist: Mediated empathy for resilience via human-AI interaction during the COVID-19 pandemic. Inform Process Manag. 2021;56(1):14–28. https://doi.org/10.1016/j.ipm.2020.103074.

Publisher’s Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.