

UDK: 316.774:614

**Wenshuai Qin**

**PhD, Faculty of Economics and Tourism, Bukhara State University**

Email: wenshuai\_qin@163.com

ORCID: 0009-0005-9783-1357

## **INTERNET USE, ONLINE HEALTH INFORMATION BEHAVIORS, AND SELF-RATED HEALTH IN CHINA: EVIDENCE FROM THE CHINESE GENERAL SOCIAL SURVEY**

**Annotatsiya.** Ushbu tadqiqot Internetdan foydalanish va o'zini baholagan salomatlik o'rtasidagi bog'liqlikni hamda ushbu munosabatda onlayn salomatlik axboroti bilan bog'liq xatti-harakatlarning rolini o'rganadi. Tadqiqotda 2021-yilgi Chinese General Social Survey (CGSS 2021) ma'lumotlaridan foydalanilib, tartibli logistik regressiya va bootstrap mediatsiya tahlili qo'llanildi. Natijalar shuni ko'rsatadiki, Internetdan foydalanish nazorat o'zgaruvchilari kiritilmagan modelda o'zini baholagan salomatlik bilan ijobiy bog'liq. Biroq demografik, ijtimoiy-iqtisodiy va onlayn salomatlik axboroti xatti-harakatlari bilan bog'liq o'zgaruvchilar kiritilgandan so'ng ushbu bevosita bog'liqlik zaiflashadi. Keyingi tahlillar Internetdan foydalanish onlayn rejimda sog'liqqa oid ma'lumotlarni izlash xulq-atvori va onlayn sog'liqni saqlashga oid axborotni qayta ishlash xulq-atvori bilan ijobiy bog'liqligini ko'rsatdi. Ammo to'liq sozlangan modelda onlayn sog'liqni saqlashga oid axborotni qayta ishlash xulq-atvori o'zini baholagan salomatlik bilan ijobiy va statistik jihatdan ahamiyatli bog'liq bo'lsa, onlayn rejimda sog'liqqa oid ma'lumotlarni izlash xulq-atvori statistik jihatdan ahamiyatli emas. Bootstrap mediatsiya tahlili onlayn sog'liqni saqlashga oid axborotni qayta ishlash xulq-atvori orqali o'tuvchi bilvosita yo'l statistik jihatdan ahamiyatli ekanini, onlayn rejimda sog'liqqa oid ma'lumotlarni izlash xulq-atvori orqali o'tuvchi yo'l esa ahamiyatli emasligini ko'rsatdi. Ushbu natijalar raqamli kirish imkoniyati, salomatlik axborotini izlash va axborotni qayta ishlash o'rtasidagi farqni ta'kidlaydi. Tadqiqot natijalari Internetdan foydalanishning salomatlikka oid qiymati nafaqat onlayn axborotga kirish imkoniyatiga, balki shaxslarning bunday axborotni tushunish, baholash va amalda qo'llash qobiliyatiga ham bog'liq ekanini ko'rsatadi. Shu sababli, onlayn salomatlik axborotining sifati va foydalanish qulayligini oshirish hamda raqamli salomatlik savodxonligini kuchaytirish raqamli salomatlik rivoji uchun muhim bo'lishi mumkin.

**Kalit so'zlar:** Internetdan foydalanish; o'zini baholagan salomatlik; onlayn salomatlik axborotini izlash xatti-harakati; onlayn salomatlik axborotini qayta ishlash xatti-harakati; CGSS

**Аннотация.** В данном исследовании рассматривается связь между использованием Интернета и самооценкой здоровья, а также изучается роль поведения, связанного с онлайн-информацией о здоровье, в данной взаимосвязи. На основе данных Chinese General Social Survey 2021 года (CGSS 2021) в исследовании применяются порядковая логистическая регрессия и bootstrap-анализ медиаторных эффектов. Результаты показывают, что использование Интернета положительно связано с самооценкой здоровья в модели без контрольных переменных, однако эта прямая связь ослабевает после включения демографических, социально-экономических переменных и переменных, связанных с поведением в отношении онлайн-информации о здоровье. Дальнейший анализ показывает, что использование Интернета положительно связано как с Поведение, связанное с поиском медицинской информации в интернете, так и с поведением, связанное с обработкой медицинской информации в интернете. Однако в полностью скорректированной модели поведение, связанное с обработкой медицинской информации в интернете положительно и статистически значимо связано с самооценкой здоровья, тогда как поведение, связанное с поиском медицинской информации в интернете не является статистически значимым. Bootstrap-анализ медиаторных эффектов также показывает, что косвенный путь через поведение, связанное с обработкой медицинской информации в

интернете является статистически значимым, тогда как путь через поведение, связанное с поиском медицинской информации в интернете не является значимым. Эти результаты подчеркивают различие между цифровым доступом, поиском информации о здоровье и обработкой информации. Они показывают, что ценность использования Интернета для здоровья зависит не только от доступа к онлайн-информации, но и от способности людей понимать, оценивать и применять такую информацию. Поэтому повышение качества и удобства использования онлайн-информации о здоровье, а также укрепление цифровой грамотности в области здоровья могут иметь важное значение для развития цифрового здравоохранения.

**Ключевые слова:** использование Интернета; самооценка здоровья; поиск онлайн-информации о здоровье; обработка онлайн-информации о здоровье; CGSS

**Abstract.** This study examines the association between Internet use and self-rated health and investigates the role of online health information behaviors in this relationship. Using data from the 2021 Chinese General Social Survey (CGSS 2021), this study employs ordered logistic regression and bootstrap mediation analysis. The results show that Internet use is positively associated with self-rated health in the unadjusted model, but this direct association becomes weaker after demographic, socioeconomic, and online health information behavior variables are included. Further analysis shows that Internet use is positively associated with both Online Health Information Seeking Behavior (OHISB) and Online Health Information Processing Behavior (OHIPB). However, in the fully adjusted model, OHIPB is positively and significantly associated with self-rated health, whereas OHISB is not. The bootstrap mediation analysis further indicates that the indirect pathway through OHIPB is statistically significant, while the pathway through OHISB is not. These findings highlight the distinction between digital access, health information seeking, and information processing. They suggest that the health value of Internet use depends not only on access to online information, but also on individuals' ability to understand, evaluate, and apply such information. Improving the quality and usability of online health information and strengthening digital health literacy may therefore be important for digital health development.

**Keywords:** Internet use; self-rated health; online health information seeking behavior; online health information processing behavior; CGSS

### 1. Introduction

The growing use of digital technologies has transformed the way individuals obtain and use health information across societies. Internet users can now access information on disease prevention, symptoms, treatment options, healthy lifestyles, and medical services with substantially lower temporal and geographical barriers. At the same time, the expansion of online health information has created a new problem: greater information availability does not necessarily lead to better health outcomes. Online health information varies considerably in accuracy, credibility, completeness, and usability, while users differ in their capacity to identify reliable sources and translate information into health-related decisions. Earlier research had already identified both the opportunities and risks associated with Internet-based health information, and subsequent studies have continued to emphasize problems of information credibility, trust, and quality evaluation<sup>[1–6]</sup>. The central question in digital health research has therefore gradually shifted from whether health information is available online to how Internet engagement is associated with individuals' ability to obtain, interpret, and use health information in ways that may be related to health outcomes.

Internet use may be associated with health through several channels. Online resources can increase exposure to health knowledge, facilitate access to medical information, support self-management, and provide additional sources for health-related decision-making. However, Internet access alone does not ensure effective engagement with health information. Norman and Skinner<sup>[7]</sup> emphasized that digital health engagement requires not only access to information but also the abilities to locate, understand, evaluate, and apply it. This distinction has become increasingly relevant as online environments have expanded in scale and complexity. Kim et al.

[8], in a systematic review and meta-analysis, found a moderate positive association between eHealth literacy and health-related behaviors, but also noted that the studies included in their review were cross-sectional. The available evidence therefore supports an association between digital information capabilities and health-related behavior while remaining insufficient for strong causal inference. Internet use and health outcomes are thus more appropriately understood through the specific information behaviors that occur between digital engagement and health-related decisions rather than as a simple technological effect [7,8].

One of the most extensively studied forms of such behavior is Online Health Information Seeking Behavior (OHISB). Health information seeking refers to purposeful efforts to obtain information in response to health concerns, health management needs, or decision-making requirements [9]. As Internet access has expanded, online channels have become increasingly important sources of health information, and health information seeking has developed into a distinct field of digital health research [10,11]. Existing evidence shows that OHISB varies considerably across individuals and social groups. Mirzaei et al. [12], based on a systematic review of 344 studies, identified age, education, gender, health condition, and financial income among the most recurrent predictors of health information-seeking behavior. Studies focusing on older populations similarly show substantial variation in information needs, preferred sources, search practices, and barriers to online health information acquisition [13]. These findings indicate that Internet use and online health information seeking should not be treated as equivalent behaviors: access to digital resources creates opportunities for information seeking, but the extent to which those opportunities are actively used depends on individual needs, resources, and behavioral choices [9,11–13].

The acquisition of health information is only one stage of digital health engagement. After obtaining information, individuals may need to assess its credibility, compare different sources, interpret its relevance, and decide whether and how it should be incorporated into health-related decisions. Research on online health information evaluation demonstrates that these activities constitute a complex process. Diviani et al. [14] found systematic differences in the ability of users with different health literacy levels to evaluate online health information, while Sbaffi and Rowley [4] showed that trust and credibility judgments are central to the selection and use of web-based health information. Sun et al. [5] identified 25 evaluation criteria and 165 indicators used by consumers when assessing online health information, illustrating the complexity and contextual nature of information evaluation. Zhang and Kim [6], in a meta-analysis of 147 studies, further found that evaluations of web-based health information quality were more strongly associated with cognitive appraisal and behavioral intention than with actual behavior. These findings suggest that searching for information and effectively processing it are related but distinct activities. In this study, the evaluation and utilization of online health information are conceptualized as Online Health Information Processing Behavior (OHIPB), distinguishing the post-acquisition processing of health information from the preceding act of information seeking [4–6,14].

The distinction between information seeking and information processing is also consistent with behavioral approaches to health decision-making. The Health Belief Model (HBM) proposes that health-related behaviors are shaped by perceived susceptibility, perceived severity, perceived benefits, perceived barriers, and cues to action [15]. Online health information can function as an informational cue, while individual judgments regarding the usefulness and credibility of such information may shape subsequent behavioral responses. Empirical research has applied the HBM to online health information seeking and found that perceived risk and perceived benefits are associated with individuals' willingness to seek health information online. Zhao et al. [16], for example, found significant relationships between HBM-related perceptions and OHIS among individuals with chronic conditions. Within this framework, Internet use increases opportunities for exposure to health information, OHISB reflects active information acquisition, and OHIPB reflects subsequent evaluation and utilization. These stages need not exhibit identical relationships with health outcomes, because access, seeking, evaluation, and application involve

different behavioral and cognitive demands <sup>[7,15,16]</sup>.

China has experienced rapid expansion of Internet use alongside major changes in the ways residents obtain information and interact with digital services. According to the 55th *Statistical Report on Internet Development in China*, the number of Internet users increased from 620,000 in 1997 to 1.108 billion by the end of 2024, with the Internet penetration rate reaching 78.6% <sup>[17]</sup>. Research based on Chinese survey data has increasingly examined the relationship between Internet use and health. Wang et al. <sup>[18]</sup>, using CGSS data from 2012 and 2015, reported associations between Internet use and several dimensions of health among older adults, although the relationship with self-rated health was not consistently significant. Lyu and Sun <sup>[19]</sup>, using the China Family Panel Studies, found that Internet use was associated with self-rated health among older adults and examined social capital as a potential pathway. Liu et al. <sup>[20]</sup>, using CGSS 2017, similarly identified an association between Internet use and self-rated health among older adults and examined the role of social support. These studies demonstrate that Internet-health relationships have already received empirical attention in China, while much of the existing evidence has focused on older populations and on social pathways such as social capital and social support <sup>[18–20]</sup>.

More recent studies have brought health information behavior into this literature. Liu et al. <sup>[21]</sup>, based on a survey of 10,000 Chinese Internet users, reported significant relationships among online health information seeking, eHealth literacy, perceived information quality, and health-related behaviors, while Yang and Qi <sup>[22]</sup>, using CGSS 2021, identified an association between Internet use and self-rated health and examined heterogeneity by family economic status and gender. Existing research has therefore established important links between Internet use and health, as well as between online health information seeking and health-related behavior. However, these research streams have often been examined separately. Moreover, studies of online health information behavior frequently focus on older adults, patients, students, or existing Internet users, while relatively less attention has been given to distinguishing information acquisition from subsequent information evaluation and utilization within a general-population framework.

Building on this literature, the present study uses data from the 2021 Chinese General Social Survey (CGSS 2021) to examine the relationship among Internet use, online health information behaviors, and self-rated health among Chinese residents. Specifically, this study distinguishes Online Health Information Seeking Behavior (OHISB) from Online Health Information Processing Behavior (OHIPB) and incorporates Internet use, information seeking, information processing, and self-rated health into a unified analytical framework. By doing so, the study examines whether different stages of digital health engagement are associated with self-rated health in different ways. This study contributes to the digital health literature by moving beyond a simple access-based perspective and by highlighting the importance of information processing and application in understanding the health implications of Internet use. Given the cross-sectional nature of the data, the findings are interpreted as associations and statistical mediation patterns rather than causal effects.

## 2. Research methods

**2.1 Theoretical framework** This study adopts the Health Belief Model (HBM) as the theoretical framework for understanding the relationships among Internet use, online health information behaviors, and self-rated health. The HBM suggests that health-related behaviors are shaped by individuals' perceptions of health risks, benefits, barriers, and cues to action. In the digital context, Internet use may increase individuals' exposure to health-related information and provide additional cues for health-related decision-making.

Based on this framework, this study distinguishes two forms of online health information behavior: Online Health Information Seeking Behavior (OHISB) and Online Health Information Processing Behavior (OHIPB). OHISB refers to the acquisition of health-related information online, while OHIPB refers to the subsequent understanding, evaluation, judgment, and application of such information. These two behaviors represent different stages of digital health

engagement and may play different roles in shaping self-rated health.

Accordingly, this study examines the association between Internet use and self-rated health and further investigates whether OHISB and OHIPB serve as mechanism pathways linking Internet use to self-rated health. Specifically, the analysis first examines the relationships between Internet use and the two online health information behaviors, and then evaluates whether these behaviors are associated with self-rated health after controlling for demographic and socioeconomic characteristics, including gender, age, education, income, marital status, ethnicity, religious belief, and household registration status.

## 2.2 Data

This study uses data from the 2021 wave of the Chinese General Social Survey (CGSS), a nationally representative survey covering demographic, socioeconomic, digital-use, and health-related information among Chinese residents.

The variables used in this study are drawn from different CGSS questionnaire modules. The Internet use variable is taken from the core questionnaire module and therefore has a relatively large number of valid observations. In contrast, self-rated health and online health information behavior variables are mainly drawn from the D/ISSP health module, which was administered to a smaller subsample in the released CGSS 2021 dataset. Therefore, the valid sample size for these variables is smaller than that for the general Internet use variable.

After data cleaning and variable construction using Stata MP 18.0, invalid responses and missing values were excluded on a model-specific basis. Because different model specifications include different combinations of variables, the analytical sample size varies across models. The D/ISSP health module includes approximately 2,690 respondents, and the fully specified regression model includes 1,477 complete observations. Thus, the reduction in sample size mainly reflects the modular questionnaire design of CGSS 2021 and further item-level missingness, rather than arbitrary sample deletion.

## 2.3 Variable description

Table 1 presents the definitions and coding schemes of the variables used in this study. The dependent variable is self-rated health, measured by the item asking respondents: "Overall, how would you rate your current physical and psychological health?" The responses are reverse-coded so that higher values indicate better self-rated health. Self-rated health has been widely used as a subjective indicator of overall health status and has been shown to be associated with subsequent health outcomes<sup>[23]</sup>.

The key independent variable is Internet use, measured by the item asking whether respondents had used the Internet through computers, mobile phones, smart wearable devices, or other means during the past six months. Respondents who reported using the Internet are coded as 1, while those who did not are coded as 0.

Two mechanism variables are included in this study: Online Health Information Seeking Behavior (OHISB) and Online Health Information Processing Behavior (OHIPB). OHISB is measured by items on whether and how often respondents searched online for health or medical information, including information on healthy lifestyles, anxiety or stress, and vaccination. OHIPB is measured by items on whether online health information affects health behaviors, helps respondents understand doctors' advice, judge whether medical consultation is needed, and confirm whether doctors' advice is appropriate. Higher values indicate stronger engagement in the corresponding behavior. The internal consistency of the two mechanism variables is acceptable, with Cronbach's  $\alpha$  values of 0.856 for OHISB and 0.723 for OHIPB.

Control variables include gender, age, educational attainment, marital status, household income, ethnicity, religious belief, and household registration status.

Table 1.

| Variable definitions and coding (CGSS 2021) |                                                       |                                                                                                                                                                                                                                                                              |
|---------------------------------------------|-------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Variable Type                               | Variable Name                                         | Variable Description and Coding                                                                                                                                                                                                                                              |
| Dependent variable                          | Self-rated health                                     | “Overall, how would you rate your current physical and psychological health?” Reverse-coded as: <b>1 = poor, 5 = very good</b> .                                                                                                                                             |
| Independent variable                        | Internet use                                          | “In the past six months, have you used the Internet through computers, mobile phones, smart wearable devices, or other means?” Coded as: <b>1 = yes, 0 = no</b> .                                                                                                            |
| Mechanism variable                          | Online Health Information Seeking Behavior (OHISB)    | “Did you search online for health or medical information, including information on healthy lifestyles, anxiety or stress, and vaccination?” Higher values indicate more frequent and broader seeking behavior.                                                               |
| Mechanism variable                          | Online Health Information Processing Behavior (OHIPB) | “Does online health information affect your health behavior, help you understand doctors’ advice, judge whether medical consultation is needed, and confirm whether doctors’ advice is appropriate?” Higher values indicate stronger information processing and application. |
|                                             | Gender                                                | 1 = male, 0 = female.                                                                                                                                                                                                                                                        |
|                                             | Age                                                   | Calculated from birth year.                                                                                                                                                                                                                                                  |
|                                             | Educational attainment                                | 1 = primary school or below; 2 = junior high school; 3 = high school/technical secondary school; 4 = college or bachelor’s degree; 5 = graduate education or above.                                                                                                          |
| Control variable                            | Marital status                                        | 1 = married, 0 = otherwise.                                                                                                                                                                                                                                                  |
|                                             | Income                                                | Annual household income, recoded as 1 = ≤30,000 CNY; 2 = 30,001–100,000 CNY; 3 = >100,000 CNY.                                                                                                                                                                               |
|                                             | Ethnicity                                             | 1 = Han, 0 = ethnic minorities.                                                                                                                                                                                                                                              |
|                                             | Religious belief                                      | 1 = yes, 0 = no.                                                                                                                                                                                                                                                             |
|                                             | Household registration status                         | 1 = rural registration, 0 = non-rural registration.                                                                                                                                                                                                                          |

Notes: OHISB refers to Online Health Information Seeking Behavior, and OHIPB refers to Online Health Information Processing Behavior. Both are mechanism variables constructed from relevant CGSS 2021 items. Higher values indicate stronger engagement in the corresponding behavior.

#### 2.4 Statistical analysis methods

This study conducts empirical analyses using Stata MP 18.0. First, descriptive statistical analysis is performed to summarize the sample characteristics and the distribution of the study variables. Second, variance inflation factor (VIF) tests are conducted to examine potential multicollinearity among the explanatory, mechanism, and control variables.

Because the dependent variable is measured on an ordinal scale, ordered logistic regression models are employed to examine the association between Internet use and self-rated health. The baseline models gradually introduce control variables and online health information behaviors to assess how the association between Internet use and self-rated health changes across model specifications. In addition, the proportional-odds assumption is examined using a generalized

ordered logit model with the autofit procedure.

Mechanism and mediation analyses are then conducted to examine whether online health information behaviors explain the association between Internet use and self-rated health. Specifically, OHISB and OHIPB are first examined as outcome variables to assess their associations with Internet use. They are then incorporated into the self-rated health model. To further test the indirect pathways, bootstrap mediation analysis with 5,000 replications is performed. The indirect effects through OHISB and OHIPB are evaluated using bias-corrected 95% confidence intervals. Finally, robustness checks are conducted using alternative measurements of the dependent variable and the key explanatory variable. Specifically, physical self-rated health is used as an alternative dependent variable, and Internet use frequency is used as an alternative measure of Internet use.

The empirical models are specified as follows:

Baseline model:

$$Y_i^* = \alpha_0 + \alpha_1 InternetUse_i + \alpha_2 Controls_i + \varepsilon_i \quad (1)$$

where  $Y_i^*$  represents the latent level of self-rated health,  $InternetUse_i$

indicates Internet use,  $Controls_i$  represents the vector of control variables, and

$\varepsilon_i$  is the error term.

Mechanism models:

$$OHISB_i = \beta_0 + \beta_1 InternetUse_i + \beta_2 Controls_i + \mu_i \quad (2)$$

$$OHIPB_i = \gamma_0 + \gamma_1 InternetUse_i + \gamma_2 Controls_i + v_i \quad (3)$$

Full mediation model:

$$Y_i^* = \delta_0 + \delta_1 InternetUse_i + \delta_2 OHISB_i + \delta_3 OHIPB_i + \delta_4 Controls_i + \eta_i \quad (4)$$

where  $OHISB_i$  and  $OHIPB_i$  represent online health information seeking behavior and online health information processing behavior, respectively. The indirect effects are

calculated as  $\beta_1 \times \delta_2$  for the OHISB pathway and  $\gamma_1 \times \delta_3$  for the OHIPB pathway. A mediation effect is considered statistically significant when the bias-corrected 95% confidence interval does not include zero.

### 3. Data analysis

#### 3.1 Sample information

Table 2 presents the descriptive statistics of the variables included in this study. The mean value of self-rated health is 2.91 (SD = 1.16), indicating that respondents' overall self-rated health is around the middle level. Regarding Internet use, approximately 69.8% of respondents reported using the Internet. For online health information behaviors, the mean value of OHISB is 1.93 (SD = 0.92), while the mean value of OHIPB is 3.33 (SD = 0.67). These results indicate that respondents vary in both online health information seeking and online health information processing behaviors. Since OHISB and OHIPB capture different aspects of online health information engagement, their mean values should be interpreted according to their respective measurement meanings rather than directly compared.

Among the control variables, respondents have an average age of 51.64 years, and approximately 45.2% are male. The sample also shows variation in educational attainment,

household income, ethnicity, religious belief, and household registration status. Rural household registration accounts for 59.4% of the sample. These descriptive characteristics provide the basis for the subsequent regression analyses examining the associations among Internet use, online health information behaviors, and self-rated health.

**Table 2.****Descriptive statistics for study variables**

| Variable                     | N     | Mean    | Std. Dev. | Min | Max |
|------------------------------|-------|---------|-----------|-----|-----|
| Self-rated health            | 2,688 | 2.9115  | 1.1641    | 1   | 5   |
| Internet use                 | 8,138 | 0.6976  | 0.4593    | 0   | 1   |
| OHISB                        | 2,563 | 1.9279  | 0.9186    | 1   | 5   |
| OHIPB                        | 1,834 | 3.3261  | 0.6706    | 1   | 5   |
| Gender (1 = male)            | 8,148 | 0.4515  | 0.4977    | 0   | 1   |
| Age                          | 8,148 | 51.6437 | 17.5739   | 18  | 99  |
| Educational attainment       | 8,127 | 2.2820  | 1.1511    | 1   | 5   |
| Marital status (1 = married) | 8,148 | 0.7046  | 0.4563    | 0   | 1   |
| Income category              | 6,420 | 1.8880  | 0.7399    | 1   | 3   |
| Ethnicity (1 = Han)          | 8,148 | 0.9264  | 0.2612    | 0   | 1   |
| Religious belief (1 = yes)   | 8,148 | 0.0750  | 0.2634    | 0   | 1   |
| Rural household registration | 8,148 | 0.5943  | 0.4911    | 0   | 1   |

Notes: Self-rated health ranges from 1 = poor to 5 = very good. Internet use is coded as 1 = yes and 0 = no. OHISB and OHIPB are mechanism variables constructed from relevant CGSS 2021 items. Higher values indicate stronger engagement in the corresponding behavior. Missing responses are excluded.

**3.2 Data quality verification**

Table 3 presents the results of the multicollinearity diagnostics for the explanatory, mechanism, and control variables.

**Table 3.****Results of Multicollinearity Diagnostics**

| Variable                      | VIF          | 1/VIF |
|-------------------------------|--------------|-------|
| Internet use                  | 1.427        | 0.701 |
| OHISB                         | 1.422        | 0.703 |
| OHIPB                         | 1.128        | 0.887 |
| Gender                        | 1.024        | 0.977 |
| Age                           | 1.666        | 0.600 |
| Educational attainment        | 1.769        | 0.565 |
| Marital status                | 1.124        | 0.889 |
| Income                        | 1.362        | 0.734 |
| Ethnicity                     | 1.073        | 0.932 |
| Religious belief              | 1.063        | 0.941 |
| Household registration status | 1.305        | 0.766 |
| <b>Mean VIF</b>               | <b>1.306</b> |       |

Notes: VIF = variance inflation factor. The results indicate no serious multicollinearity among the

explanatory variables.

The VIF values for all variables are below 2, with a mean VIF of 1.306, indicating that multicollinearity is unlikely to be a serious concern in the regression models.

Specifically, the VIF values for Internet use, OHISB, and OHIPB are 1.427, 1.422, and 1.128, respectively. These values suggest that the key explanatory variables can be included jointly in the subsequent regression analyses.

In addition, the proportional-odds assumption was examined using a generalized ordered logit model with the autofit procedure. The results show that the parallel-lines restriction can be imposed on the key explanatory variables, including Internet use, OHISB, and OHIPB. The Wald test for the final model is not statistically significant,  $\chi^2(27) = 16.06$ ,  $p = 0.9518$ , indicating that the final specification does not violate the proportional-odds assumption.

### 3.3 Baseline Regression Results

Table 4 presents the baseline regression results examining the association between Internet use, online health information behaviors, and self-rated health. Ordered logistic regression models are estimated with variables introduced progressively across the four specifications. Because ordered logit coefficients are reported in log-odds units, the results are interpreted mainly in terms of direction, statistical significance, and selected odds ratios for key variables.

**Table 4.**

| Baseline regression results   |                     |                      |                      |                      |
|-------------------------------|---------------------|----------------------|----------------------|----------------------|
| Variables                     | Model 1             | Model 2              | Model 3              | Model 4              |
| Internet use                  | 0.844***<br>(0.080) | 0.167<br>(0.114)     | 0.217*<br>(0.127)    | 0.103<br>(0.187)     |
| OHISB                         | —                   | —                    | -0.016<br>(0.056)    | -0.100<br>(0.065)    |
| OHIPB                         | —                   | —                    | —                    | 0.201***<br>(0.077)  |
| Gender                        | —                   | 0.134*<br>(0.080)    | 0.162**<br>(0.082)   | 0.100<br>(0.096)     |
| Age                           | —                   | -0.016***<br>(0.003) | -0.016***<br>(0.003) | -0.016***<br>(0.004) |
| Educational attainment        | —                   | 0.103**<br>(0.048)   | 0.124**<br>(0.049)   | 0.136**<br>(0.056)   |
| Marital status                | —                   | 0.027<br>(0.094)     | 0.052<br>(0.096)     | 0.062<br>(0.114)     |
| Income                        | —                   | 0.387***<br>(0.065)  | 0.369***<br>(0.067)  | 0.319***<br>(0.078)  |
| Ethnicity                     | —                   | 0.236<br>(0.159)     | 0.217<br>(0.161)     | 0.089<br>(0.184)     |
| Religious belief              | —                   | 0.033<br>(0.155)     | -0.016<br>(0.161)    | -0.236<br>(0.195)    |
| Household registration status | —                   | 0.351***<br>(0.092)  | 0.360***<br>(0.094)  | 0.382***<br>(0.108)  |
| N                             | 2,687               | 2,128                | 2,042                | 1,477                |
| Log likelihood                | -3921.459           | -3031.164            | -2900.246            | -2071.454            |
| Pseudo R <sup>2</sup>         | 0.014               | 0.032                | 0.032                | 0.026                |

Notes: The dependent variable is self-rated health. Ordered logistic regression models are estimated. Standard errors are reported in parentheses. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

Model 1 examines the association between Internet use and self-rated health without control variables. Internet use is positively and significantly associated with self-rated health ( $\beta = 0.844$ ,  $p < 0.01$ ). The corresponding odds ratio is approximately 2.33, suggesting that Internet users have higher odds of reporting better self-rated health than non-users before controls are

introduced. After demographic and socioeconomic characteristics are introduced in Model 2, the coefficient of Internet use remains positive but becomes statistically non-significant ( $\beta = 0.167$ ), indicating that the direct association between Internet use and self-rated health is partly explained by individual characteristics.

Model 3 further includes OHISB. The coefficient of Internet use is positive and weakly significant ( $\beta = 0.217$ ,  $p < 0.10$ ), while OHISB is not significantly associated with self-rated health. In Model 4, which additionally includes OHIPB, Internet use remains positive but statistically non-significant. OHISB is also not significant, whereas OHIPB is positively and significantly associated with self-rated health ( $\beta = 0.201$ ,  $p < 0.01$ ). This coefficient corresponds to an odds ratio of approximately 1.22, indicating that a one-unit increase in OHIPB is associated with higher odds of reporting better self-rated health, holding other variables constant.

Among the control variables, age is consistently negatively associated with self-rated health. Educational attainment, household income, and household registration status show positive and statistically significant associations in the adjusted models. Overall, the results indicate that the direct association between Internet use and self-rated health becomes weaker after controlling for individual characteristics and online health information behaviors. In the full model, OHIPB, rather than OHISB, is the mechanism variable most clearly associated with better self-rated health. This finding suggests that the health-related value of Internet use may depend more on individuals' ability to process and apply online health information than on information seeking alone.

### 3.4 Mechanism and Mediation Analysis

Table 5 reports the results of the bootstrap mediation analysis. The analysis examines two mechanism pathways linking Internet use to self-rated health: the pathway through online health information seeking behavior and the pathway through online health information processing behavior.

The indirect effect through OHISB is not statistically significant, as the bias-corrected 95% confidence interval includes zero. In contrast, the indirect effect through OHIPB is positive and statistically significant, with a bias-corrected 95% confidence interval that does not include zero. This result indicates that OHIPB constitutes a significant mechanism linking Internet use to self-rated health, whereas the OHISB pathway is not supported.

The direct effect of Internet use on self-rated health is not statistically significant after the two mechanism variables are included. The total effect is also not statistically significant. These results suggest that Internet use is more closely associated with self-rated health through the processing and application of online health information than through online health information seeking alone.

**Table 5.**

| Mechanism analysis results               |             |                     |                       |                 |
|------------------------------------------|-------------|---------------------|-----------------------|-----------------|
| Path / Effect                            | Coefficient | Bootstrap Std. Err. | Bias-corrected 95% CI | Result          |
| Internet use → OHISB → Self-rated health | -0.0876     | 0.0604              | [-0.2069, 0.0296]     | Not significant |
| Internet use → OHIPB → Self-rated health | 0.0637      | 0.0295              | [0.0171, 0.1336]      | Significant     |
| Total indirect effect                    | -0.0239     | 0.0603              | [-0.1372, 0.0977]     | Not significant |
| Direct effect                            | 0.1033      | 0.2187              | [-0.3190, 0.5490]     | Not significant |

Notes: The mediation effects were estimated using 5,000 bootstrap replications. Bias-corrected 95% confidence intervals are reported. A mediation effect is considered statistically significant when the confidence interval does not include zero.

### 3.5 Robustness Checks

Table 6 presents the robustness checks using alternative measures of the dependent variable and the key explanatory variable. Model 1 uses physical self-rated health as an alternative dependent variable, while Model 2 replaces binary Internet use with Internet use frequency.

In Model 1, the coefficient of Internet use remains positive but is not statistically significant. OHISB is close to zero and statistically non-significant, while OHIPB remains positive but is also not statistically significant. This result indicates that the direction of the association between Internet use, online health information processing, and health remains generally consistent, although the statistical evidence is weaker when physical self-rated health is used as the outcome.

In Model 2, Internet use frequency is positively and significantly associated with self-rated health ( $\beta = 0.099$ ,  $p < 0.05$ ). OHIPB remains positively and significantly associated with self-rated health ( $\beta = 0.199$ ,  $p < 0.01$ ), while OHISB shows a weak negative association. These results are broadly consistent with the baseline findings, particularly regarding the importance of online health information processing behavior.

Overall, the robustness checks provide partial support for the main findings. The results suggest that the role of Internet use depends on how Internet engagement is measured, while OHIPB remains more consistently associated with self-rated health than OHISB. Therefore, the findings should be interpreted as emphasizing the importance of online health information processing and application rather than Internet access or information seeking alone.

**Table 6.**

| Variables                     | Robustness checks                   |                                 |
|-------------------------------|-------------------------------------|---------------------------------|
|                               | Model 1: Physical self-rated health | Model 2: Internet use frequency |
| Internet use                  | 0.207<br>(0.185)                    | —                               |
| Internet use frequency        | —                                   | 0.099**<br>(0.046)              |
| OHISB                         | -0.002<br>(0.065)                   | -0.123*<br>(0.065)              |
| OHIPB                         | 0.117<br>(0.078)                    | 0.199***<br>(0.077)             |
| Gender                        | 0.121<br>(0.097)                    | 0.099<br>(0.096)                |
| Age                           | -0.030***<br>(0.004)                | -0.014***<br>(0.004)            |
| Educational attainment        | 0.138**<br>(0.057)                  | 0.119**<br>(0.057)              |
| Marital status                | 0.074<br>(0.116)                    | 0.039<br>(0.114)                |
| Income                        | 0.357***<br>(0.079)                 | 0.309***<br>(0.078)             |
| Ethnicity                     | 0.208<br>(0.189)                    | 0.064<br>(0.184)                |
| Religious belief              | -0.061<br>(0.192)                   | -0.252<br>(0.195)               |
| Household registration status | 0.269**<br>(0.110)                  | 0.392***<br>(0.108)             |
| N                             | 1,476                               | 1,477                           |
| Log likelihood                | -1905.070                           | -2069.269                       |
| Pseudo R <sup>2</sup>         | 0.052                               | 0.027                           |

Notes: Model 1 uses physical self-rated health as an alternative dependent variable. Model 2 replaces Internet use with Internet use frequency. Ordered logistic regression models are estimated. Standard errors are reported in parentheses. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

#### 4. Discussion, Conclusion and Implications

Using data from the 2021 Chinese General Social Survey (CGSS), this study examined the relationship among Internet use, online health information behaviors, and self-rated health. Rather than treating Internet use only as a general indicator of digital access, this study distinguished between Online Health Information Seeking Behavior (OHISB) and Online Health Information Processing Behavior (OHIPB). This distinction allows the study to examine whether different stages of digital health engagement are associated with self-rated health in different ways.

The empirical results show that Internet use is positively and significantly associated with self-rated health in the model without control variables. However, this association becomes statistically non-significant after demographic and socioeconomic characteristics are controlled. After OHISB and OHIPB are incorporated into the model, Internet use remains positive but non-significant. Among the two online health information behavior variables, OHISB does not show a significant positive association with self-rated health, whereas OHIPB is positively and significantly associated with self-rated health in the fully adjusted model. The bootstrap mediation analysis further shows that the indirect effect through OHISB is not statistically significant, while the indirect effect through OHIPB is positive and significant.

These findings suggest that the health-related value of Internet use depends not only on access to online information, but also on how individuals process and apply such information. Simply searching for health information online may not be sufficient to improve health perceptions. Instead, the key lies in whether individuals can understand, evaluate, judge, and use online health information in health-related decision-making. This finding supports the view that digital health engagement should be understood as a multi-stage process, including access, information seeking, information processing, and practical application.

Previous studies in the Chinese context have shown that Internet use is associated with health outcomes, particularly among older adults, and have often emphasized social pathways such as social capital and social support. For example, Wang et al.<sup>[18]</sup>, Lyu and Sun<sup>[19]</sup>, and Liu et al.<sup>[20]</sup> mainly examined the relationship between Internet use and health from the perspective of older populations and social resources. More recent research has also used CGSS 2021 data to examine Internet use, family economic status, gender differences, and self-rated health<sup>[22]</sup>. Building on this literature, the present study focuses on the general population and shifts attention to online health information behaviors. By distinguishing online health information seeking from online health information processing, this study shows that the relationship between Internet use and self-rated health is more clearly reflected in the processing and application of online health information than in information seeking alone.

This study contributes to the digital health literature in two main ways. First, it clarifies the conceptual difference between online health information seeking and online health information processing. The results show that these two behaviors are not equivalent: information seeking alone is not significantly associated with self-rated health, whereas information processing and application play a more important role. Second, the study provides empirical evidence from a general-population sample in China, suggesting that the health implications of Internet use should not be understood only from the perspective of digital access, but also from the perspective of digital health literacy and information-processing capacity.

For countries advancing digital health transformation, the findings suggest that the effectiveness of digital health initiatives depends not only on expanding access to digital services, but also on how individuals understand and use the information provided through these services. This implication is relevant to contexts such as Uzbekistan, where healthcare digitalization and e-health development have become part of ongoing health system reforms. As digital health platforms continue to develop, greater attention should be paid to the quality, credibility, clarity,

and usability of online health information. Strengthening digital health literacy may help ensure that digital health resources are not only accessible, but also effectively used in everyday health-related decision-making.

Overall, the findings indicate that the potential health value of Internet use depends less on Internet access itself and more on whether individuals can effectively understand, evaluate, and apply online health information. Digital health policies should therefore move beyond expanding access alone and pay greater attention to the quality of information environments and users' capacity to make informed health-related decisions.

### 5. Limitations

This study has several limitations. First, the cross-sectional nature of CGSS 2021 limits causal inference, and the observed relationships should therefore be interpreted as associations rather than causal effects. Although the bootstrap mediation analysis identifies a significant indirect pathway through OHIPB, the cross-sectional design does not allow strong causal interpretation of this pathway. Second, OHISB and OHIPB are constructed from self-reported CGSS items and are not fully validated psychometric scales. Self-rated health is also a subjective measure and may be influenced by individual perceptions, reporting tendencies, and social contexts. Third, valid responses vary across variables, particularly for the online health information measures, resulting in smaller samples in the fully specified models. Although missing responses are excluded in the main analysis, potential selection bias cannot be completely ruled out. Future research could use longitudinal data, validated measurement instruments, and alternative approaches to missing data to further examine the relationships among Internet use, online health information behaviors, and self-rated health.

### Acknowledgment

The authors would like to express their sincere gratitude to the Chinese General Social Survey (CGSS) project for providing access to the valuable data used in this study. The authors also appreciate the constructive comments and suggestions from colleagues and reviewers, which contributed to improving the quality of this manuscript. Any remaining errors and omissions are solely the responsibility of the authors.

### Data Availability Statement

This study utilized data from the China General Social Survey (CGSS) 2021, which are publicly available from the official website of the CGSS at <http://cgss.ruc.edu.cn>. The authors did not have any special access privileges; other researchers can obtain these data in the same manner.

### Declaration of Interest

All authors disclosed no relevant relationships.

### References

1. Cline, R. J. W., & Haynes, K. M. (2001). Consumer health information seeking on the Internet: The state of the art. *Health Education Research*, 16(6), 671–692. doi:10.1093/her/16.6.671.
2. Eysenbach, G., Powell, J., Kuss, O., & Sa, E.-R. (2002). Empirical studies assessing the quality of health information for consumers on the World Wide Web: A systematic review. *JAMA*, 287(20), 2691–2700. doi:10.1001/jama.287.20.2691.
3. Murray, E., Lo, B., Pollack, L., Donelan, K., Catania, J., Lee, K., Zapert, K., & Turner, R. (2003). The impact of health information on the Internet on health care and the physician-patient relationship: National U.S. survey among 1,050 U.S. physicians. *Journal of Medical Internet Research*, 5(3), e17. doi:10.2196/jmir.5.3. e17.
4. Sbaffi, L., & Rowley, J. (2017). Trust and credibility in web-based health information: A review and agenda for future research. *Journal of Medical Internet Research*, 19(6), e218. doi:10.2196/jmir.7579.
5. Sun, Y., Zhang, Y., Gwizdka, J., & Trace, C. B. (2019). Consumer evaluation of the quality of online health information: Systematic literature review of relevant criteria and indicators.

*Journal of Medical Internet Research*, 21(5), e12522. doi:10.2196/12522.

6. Zhang, Y., & Kim, Y. (2022). Consumers' evaluation of web-based health information quality: Meta-analysis. *Journal of Medical Internet Research*, 24(4), e36463. doi:10.2196/36463.

7. Norman, C. D., & Skinner, H. A. (2006). eHealth literacy: Essential skills for consumer health in a networked world. *Journal of Medical Internet Research*, 8(2), e9. doi:10.2196/jmir.8.2.e9.

8. Kim, K., Shin, S., Kim, S., & Lee, E. (2023). The relation between eHealth literacy and health-related behaviors: Systematic review and meta-analysis. *Journal of Medical Internet Research*, 25, e40778. doi:10.2196/40778.

9. Lambert, S. D., & Loiselle, C. G. (2007). Health information seeking behavior. *Qualitative Health Research*, 17(8), 1006–1019. doi:10.1177/1049732307305199.

10. Zimmerman, M. S., & Shaw, G. (2020). Health information seeking behaviour: A concept analysis. *Health Information & Libraries Journal*, 37(3), 173–191. doi:10.1111/hir.12287.

11. Jia, X., Pang, Y., & Liu, L. S. (2021). Online health information seeking behavior: A systematic review. *Healthcare*, 9(12), 1740. doi:10.3390/healthcare9121740.

12. Mirzaei, A., Aslani, P., Luca, E. J., & Schneider, C. R. (2021). Predictors of health information-seeking behavior: Systematic literature review and network analysis. *Journal of Medical Internet Research*, 23(7), e21680. doi:10.2196/21680.

13. Zhao, Y. C., Zhao, M., & Song, S. (2022). Online health information seeking behaviors among older adults: Systematic scoping review. *Journal of Medical Internet Research*, 24(2), e34790. doi:10.2196/34790.

14. Diviani, N., van den Putte, B., Giani, S., & van Weert, J. C. M. (2015). Low health literacy and evaluation of online health information: A systematic review of the literature. *Journal of Medical Internet Research*, 17(5), e112. doi:10.2196/jmir.4018.

15. Janz, N. K., & Becker, M. H. (1984). The Health Belief Model: A decade later. *Health Education Quarterly*, 11(1), 1–47. doi:10.1177/109019818401100101.

16. Zhao, Y. C., Zhao, M., & Song, S. (2022). Online health information seeking among patients with chronic conditions: Integrating the Health Belief Model and social support theory. *Journal of Medical Internet Research*, 24(11), e42447. doi:10.2196/42447.

17. China Internet Network Information Center. (2025). *The 55th statistical report on China's Internet development*. China Internet Network Information Center.

18. Wang, J., Liang, C., & Li, K. (2020). Impact of Internet use on elderly health: Empirical study based on Chinese General Social Survey (CGSS) data. *Healthcare*, 8(4), 482. doi:10.3390/healthcare8040482.

19. Lyu, S., & Sun, J. (2021). Internet use and self-rated health among Chinese older adults: The mediating role of social capital. *Geriatrics & Gerontology International*, 21(1), 34–38. doi:10.1111/ggi.14090.

20. Liu, N., He, Y., & Li, Z. (2022). The relationship between Internet use and self-rated health among older adults in China: The mediating role of social support. *International Journal of Environmental Research and Public Health*, 19(22), 14785. doi:10.3390/ijerph192214785.

21. Liu, D., Yang, S., Cheng, C. Y., Cai, L., & Su, J. (2024). Online health information seeking, eHealth literacy, and health behaviors among Chinese Internet users: Cross-sectional survey study. *Journal of Medical Internet Research*, 26, e54135. doi:10.2196/54135.

22. Yang, X., & Qi, M. (2025). Residents' Internet use, family economic status, and self-rated health: Gender-based differences. *BMC Public Health*, 25, 223. doi:10.1186/s12889-025-21371-w.

23. Idler, E. L., & Benyamini, Y. (1997). Self-rated health and mortality: A review of twenty-seven community studies. *Journal of Health and Social Behavior*, 38(1), 21–37. doi:10.2307/2955359.