Natural Resources for
Human Health

eISSN: 2583-1194

Original Research

Received 22 May 2026
Revised 26 June 2026
Accepted 18 July 2026

Natural Resources for Human Health 2026, Vol. 6 (8s): 523–545

KEYWORDS:
health information-seeking behaviour; digital health; artificial intelligence; misinformation; trust; digital health literacy; Santal community; young adults; West Bengal; India

Health Information-Seeking Behaviour in the Digital Age: Trust, Misinformation and AI-Based Health Information among Young Adults in the Santal Community of West Bengal

[1]Dr. Navin Kumar., [2]Mr. Debamoy Pal, [3]Dr Joydeb Patra, [4]Dr. Anuj Kumar Singh, [5]Bulbul Roy Chowdhury (Ghosh)

[1]Assistant Professor Department of Law, Brainware University, Kolkata (W B)
Email iD : navingudia25@gmail.com
[2]Assistant Professor, School of Law, Brainware University, Barasat.
E-mail ID : debamoy.pal9831@gmail.com
ORCID ID: 0009-0005-5395-0071
[3]Assistant Professor School of Law, Brainware University
E-mail iD: joydebajoy@gmail.com,
ORCID iD: 0000-0001-7797-6309
[4]Assistant Professor, Institute of Legal Study & Research,GLA University, Mathura
Email iD- anujksinghexam@gmail.com
[5]Assistant Professor Brainware University
blrc.law@brainwareuniversity.ac.in
ORCID ID- 0009-0001-5889-7768
Corresponding Author: Dr Joydeb Patra, E-mail iD: joydebajoy@gmail.com

ABSTRACT:

Background: The rapid expansion of smartphones, social media, search engines and artificial intelligence (AI)-based conversational systems has transformed how young adults obtain, evaluate and use health information. Although digital platforms can improve the accessibility and immediacy of health information, they may also expose users to misinformation, unreliable sources and AI-generated information of uncertain accuracy. These challenges may be particularly significant for socially and geographically disadvantaged communities, where access to reliable health-information services may be constrained by socioeconomic, linguistic and infrastructural factors.

Objective: This study aims to empirically examine health information-seeking behaviour among young adults in the Santal community of West Bengal, with particular emphasis on digital health-information sources, trust in information sources, exposure to health misinformation, use of AI-based health information and information-verification behaviour.

Methods: A community-based cross-sectional quantitative study will be conducted among Santal young adults aged 18–30 years in selected districts of West Bengal. A structured questionnaire will be administered to approximately 400 respondents using a multistage sampling strategy. The questionnaire will measure digital health-information-seeking behaviour, trust in health-information sources, exposure and susceptibility to health misinformation, AI-based health-information use, digital health literacy and verification behaviour. Descriptive statistics, reliability analysis, exploratory factor analysis, confirmatory factor analysis and structural equation modelling will be

employed. Demographic and digital-access variables will be examined as potential moderators.

Expected Contribution: The study is expected to identify the factors that shape digital health-information seeking among young Santal adults and clarify the relationship between trust, misinformation exposure, AI-based information use and verification behaviour. The findings may contribute to the development of culturally and linguistically appropriate digital health-information services and information-literacy interventions for tribal communities.

1. INTRODUCTION

Digital technologies have fundamentally changed the way individuals search for, access and interpret health information. Search engines, social-media platforms, health websites, mobile applications and AI-based conversational systems now provide users with immediate access to a vast range of health-related information. Young adults constitute an important population in this transformation because of their extensive engagement with smartphones, internet services and social-media platforms. Recent research indicates that young adults increasingly use search engines and social media for health-related information, although their ability to evaluate the credibility of such information varies considerably. A 2026 study of Indian youths reported extensive online health-information use but very low levels of active credibility verification, highlighting the importance of digital health literacy and critical evaluation skills.

The emergence of generative AI has added a new dimension to this information environment. AI-based systems such as conversational large language models can provide apparently personalised explanations, symptom-related information, health suggestions and summaries. Recent evidence indicates that LLM-based chatbots are already being used for health information, particularly among younger users and individuals with more positive attitudes towards AI. At the same time, questions remain concerning accuracy, credibility, transparency and users' willingness to verify AI-generated information. Health misinformation represents another major challenge. Digital platforms can facilitate the rapid circulation of misleading information concerning diseases, medicines, nutrition, vaccination, reproductive health and treatment. The problem is not simply the availability of false information; it also concerns the processes by which individuals assess the trustworthiness of information and act on it.

These issues become particularly important when examined within socially and geographically disadvantaged communities. The Santal community is one of the major Scheduled Tribes in West Bengal. Census records identify Santal as a Scheduled Tribe category in West Bengal. Previous research in West Bengal has demonstrated that tribal health behaviour is shaped by socioeconomic conditions, geographical accessibility, cultural practices and changing patterns of health-care utilisation.

However, the transition from traditional health-care seeking to digital health-information seeking among young Santal adults remains insufficiently investigated. In particular, there is limited empirical evidence concerning whether young members of the Santal community trust digital health information, whether they encounter misinformation, whether they use AI-based systems and whether they verify information before acting upon it. This study therefore proposes an empirical investigation of the intersection between digital health-information seeking, trust, misinformation and AI-based health information among young adults in the Santal community of West Bengal.

2. RESEARCH PROBLEM

The rapid growth of smartphones, social media, search engines and artificial intelligence (AI) has made health information more accessible than ever. However, easy access to information does not necessarily mean access to accurate or reliable health knowledge. Young adults may encounter contradictory, incomplete or misleading information while searching for health-related advice online.

This issue may be particularly significant among young adults in tribal communities such as the Santal community of West Bengal. First, unequal access to reliable internet infrastructure may limit their ability to obtain authoritative health information. Second, differences in education and digital health literacy may affect their capacity to identify credible sources.

Third, language barriers may make professionally produced health information difficult to understand. Fourth, social-media platforms may become important sources of information, although their content is not always scientifically verified.

Trust also plays a significant role. Young adults may rely on information provided by family members, community networks, health professionals, social-media users or digital platforms. At the same time, repeated exposure to health misinformation may make it difficult to distinguish evidence-based information from rumours or false claims. Limited verification practices can further increase this vulnerability.

The growing use of AI-generated health information introduces another important dimension. AI tools can provide rapid and understandable responses, but users may not always possess the skills necessary to evaluate their accuracy. Therefore, this study investigates how trust, misinformation exposure, digital health literacy and AI-based health-information use influence health information-seeking and verification behaviour among young Santal adults in West Bengal.

3. RESEARCH GAP

Existing research provides evidence that young adults increasingly seek health information through digital platforms. A recent scoping review identified search engines and social media as important channels for young adults' online health-information seeking.

Research in India has also examined online health-information seeking and credibility assessment among young people. However, there remains a need for community-specific research that examines how these processes operate within tribal populations.

Furthermore, recent research on AI-based health information has examined general populations and digitally active users, but there is limited evidence concerning AI-based health-information seeking among young adults belonging to indigenous/tribal communities in India.

The proposed study addresses four interconnected gaps:

Gap 1: Limited evidence on digital health-information seeking among young Santal adults.

Gap 2: Limited empirical understanding of trust in different health-information sources within this population.

Gap 3: Insufficient evidence concerning exposure to and responses to digital health misinformation.

Gap 4: Very limited empirical research examining the emerging role of AI-based health information among young tribal populations.

The study therefore moves beyond the conventional question of whether people seek health information to investigate where they seek it, whom they trust, what misinformation they encounter, whether they use AI and whether they verify information before acting upon it.

4. THEORETICAL FRAMEWORK

The **Comprehensive Model of Information Seeking (CMIS)** was originally developed to explain health-information-seeking behaviour by connecting individual characteristics, perceived health-related needs and characteristics of information sources. Importantly for your research, Basnyat et al. (2018) specifically adapted the CMIS to online health-information seeking in India, demonstrating that the framework can be applied beyond traditional health-information environments.

The contemporary digital environment makes such an extension particularly relevant. A 2026 Indian study of 196 young people found that 87.8% searched for health information online, with smartphones being the principal device (83.2%). However, only 5.1% actively verified the credibility of the information, while 68.9% did not check sources. Higher education and better digital skills were associated with greater online health-information seeking. This provides direct empirical justification for including digital health literacy and verification behaviour in your model.

The role of misinformation is also supported by the **World Health Organisation (WHO)**. WHO describes an "infodemic" as an environment containing excessive information, including false or misleading information, which can create confusion, encourage risky behaviour and undermine trust in health authorities. WHO further emphasises improving digital and health literacy as a means of building resilience to misinformation. The increasing importance of social media provides another empirical justification. In 2024, WHO and TikTok announced a collaboration to promote science-based health information, explicitly recognising that social-media platforms influence health-related behaviours while also exposing users to misinformation.

5. RESEARCH OBJECTIVES

General Objective

To examine health information-seeking behaviour in the digital age among young adults in the Santal community of West Bengal.

Specific Objectives

1. To identify the major digital and non-digital sources used by young Santal adults for obtaining health information.
2. To examine the level of trust placed in different health-information sources.
3. To assess the extent of exposure to health misinformation.
4. To examine the use of AI-based systems for obtaining health information.
5. To assess digital health literacy among respondents.
6. To examine the relationship between trust and health information-seeking behaviour.
7. To determine whether misinformation exposure influences information-verification behaviour.
8. To examine whether digital health literacy influences the ability to verify online health information.
9. To investigate whether trust in AI-based health information influences AI-based health-information seeking.
10. To develop an empirical model explaining digital health-information seeking among young Santal adults.

6. RESEARCH QUESTIONS

More recent Indian evidence strengthens the rationale: a 2026 study among 196 Indian youths found that 87.8% searched for health information online, but only 5.1% actively verified credibility**,** while 68.9% did not check sources. Education and digital skills were associated with online health-information seeking and credibility verification.

RQ1. What are the major sources of health information used by young adults in the Santal community?

This question establishes the **information environment** within which young Santal adults obtain health-related knowledge. The study should identify the relative use of traditional and digital sources, including health professionals, family members, friends, community members, government health websites, hospitals, search engines, YouTube, Facebook/Instagram/WhatsApp and AI-based tools. Respondents may also be asked about the frequency and purpose of using each source. This question is primarily **descriptive** and will be answered through frequencies, percentages, means and cross-tabulations.

RQ2. Which health-information sources are perceived as most trustworthy?

Accessibility does not necessarily correspond to credibility. This question therefore examines **source trust**, an important information-carrier characteristic within the modified CMIS. Respondents will rate their perceived trustworthiness of different sources. The analysis can compare trust in health professionals, government sources, family/community networks, social media, search engines and AI systems. This distinction is important because WHO identifies misinformation and declining trust as important consequences of an increasingly complex information environment.

RQ3. To what extent are young Santal adults exposed to digital health misinformation?

This question measures respondents' exposure to false, misleading, exaggerated or contradictory health information encountered through digital platforms. It should not assume that every encountered claim is objectively false; instead, the questionnaire can initially measure perceived exposure and uncertainty. Examples may include misinformation concerning medicines, diseases, nutrition, vaccination, reproductive health and home remedies. WHO notes that digitalisation can accelerate the circulation of misleading health information and create confusion and potentially harmful health behaviours?

RQ4. What proportion of respondents use AI-based systems for obtaining health information?

This question introduces the **AI dimension** of the study. Respondents should first be asked whether they have ever used an AI-based system—such as ChatGPT, Gemini, Copilot or another conversational AI tool—for general health information. Among users, the study can examine frequency, purpose and type of information sought.

Importantly, the study should distinguish AI-based health-information seeking from AI-based diagnosis or treatment. This makes the research ethically safer and conceptually clearer.

RQ5. Does trust in digital health information significantly influence health-information-seeking behaviour?

This is the first major **explanatory/relational research question**. It moves beyond identifying information sources and tests whether respondents who perceive digital health information as trustworthy are more likely to engage in digital health-information seeking.

The proposed relationship is:

Trust in Digital Health Information → Health Information-Seeking Behaviour

This question is strongly supported by the modified CMIS. The Indian CMIS study found that information-carrier characteristics related to **trust**, together with perceived usefulness and relevance, had direct effects on online health-information-seeking behaviour. It can therefore be tested using correlation and multiple regression, or preferably **PLS-SEM/CB-SEM** if the complete conceptual model is being tested.

RQ6. Does digital health literacy influence health-information verification behaviour?

This question examines whether respondents who possess greater ability to find, understand, evaluate and critically assess digital health information are more likely to verify information before accepting or using it.

The proposed relationship is:

Digital Health Literacy → Verification Behaviour

This is particularly important given the 2026 Indian evidence showing a substantial gap between searching for online health information and checking its credibility. A recent meta-analysis also specifically identifies eHealth literacy as an important factor in online health-information seeking.

RQ7. Does misinformation exposure influence respondents' tendency to verify health information?

This question examines whether exposure to misleading or contradictory health information changes respondents' verification behaviour**.**

The proposed relationship is:

Misinformation Exposure → Verification Behaviour

Two possible outcomes are theoretically plausible. Greater misinformation exposure may encourage individuals to become more cautious and verify information more frequently. Alternatively, repeated exposure may create confusion, information fatigue or normalisation of questionable information, potentially reducing critical evaluation. Therefore, this should be treated as an empirical question rather than assuming a particular direction.

This is consistent with WHO's emphasis on strengthening community resilience and digital/health literacy to reduce the harmful effects of misinformation.

RQ8. Does trust in AI-based health information influence AI-based health-information-seeking behaviour?

This question focuses specifically on the emerging relationship between trust in AI and AI-mediated health-information seeking**.**

The proposed relationship is:

Trust in AI Health Information → AI-Based Health-Information Seeking

Respondents who perceive AI-generated information as useful, understandable and credible may be more likely to use AI systems for health-related information. However, this relationship should be tested rather than assumed because AI-generated health information can also raise concerns about accuracy, reliability and verification.

How the 8 RQs fit together

Your research questions can ultimately be organised into **three analytical levels**:

Level	Research Questions	Purpose
Information environment	RQ1–RQ4	Identify sources, trust, misinformation and AI use
Information behaviour	RQ5–RQ7	Test relationships between trust, literacy, misinformation and verification
AI-mediated information behaviour	RQ8	Test the specific role of AI trust

7. HYPOTHESES

The hypotheses of the present study are developed from the **Comprehensive Model of Information Seeking (CMIS)**, adapted to the contemporary digital health-information environment. The model assumes that individual characteristics, information-source characteristics, perceived usefulness and information credibility influence information-seeking behaviour. In the present study, the framework is extended by incorporating digital health literacy, misinformation exposure, AI-based health-information use and verification behaviour**.**

The proposed hypotheses examine how digital health literacy, trust, perceived credibility, misinformation exposure, AI-based health-information use, and verification behaviour influence digital health-information-seeking among young adults in the Santal community of West Bengal. H1–H4 propose that digital health literacy, source trust, perceived credibility, and digital health literacy positively influence information-seeking and verification behaviour. H5 examines the relationship between misinformation exposure and verification behaviour, with a proposed negative direction. H6 and H7 suggest that trust and perceived usefulness of AI-based health information positively influence AI use, while H8 proposes that AI use increases overall digital health-information seeking. H9 proposes that verification behaviour mediates the relationship between digital health literacy and responsible health-information use. Finally, H10 examines whether digital health literacy moderates the relationship between trust and information-seeking behaviour. These hypotheses can be empirically tested using **PLS-SEM or CB-SEM**, including bootstrapping for mediation and interaction analysis for moderation.

Summary of the Hypothesised Model

Hypothesis	Independent Variable	Relationship	Dependent Variable	Role
H1	Digital Health Literacy	+	Health Information Seeking	Direct
H2	Trust in Digital Sources	+	Health Information Seeking	Direct
H3	Information Credibility	+	Health Information Seeking	Direct
H4	Digital Health Literacy	+	Verification Behaviour	Direct
H5	Misinformation Exposure	−	Verification Behaviour	Direct
H6	AI Information Trust	+	AI Information Use	Direct
H7	AI Perceived Usefulness	+	AI Information Use	Direct

H8	AI Information Use	+	Digital Health Information Seeking	Direct
H9	Digital Health Literacy → Verification	→	Responsible Information Use	**Mediation**
H10	Trust × Digital Health Literacy	→	Health Information Seeking	**Moderation**

8. RESEARCH DESIGN

8.1 Research Approach

The study adopted a quantitative, cross-sectional and empirical research design to examine digital health-information-seeking behaviour among young adults in the Santal community of West Bengal. A quantitative approach was adopted because the study sought to measure and statistically test the relationships among digital health literacy, trust in digital health-information sources, perceived information credibility, misinformation exposure, AI-based health-information use, verification behaviour and health-information-seeking behaviour.

The study employed a cross-sectional design because data were collected from respondents during a single defined period, rather than through longitudinal observation. This design enabled the researchers to assess the prevalence and patterns of digital health-information practices and examine statistical associations among the principal constructs.

The research design was explicitly hypothesis-driven. The ten hypotheses (H1–H10) were empirically tested using appropriate statistical techniques. Descriptive statistics were initially used to examine respondents' demographic characteristics, digital access and health-information sources. Reliability and validity analyses were subsequently conducted to assess the measurement instruments. Correlation, multiple regression and Structural Equation Modelling (SEM) were then employed to examine the proposed relationships.

For H1–H8, the direct relationships among the principal variables were examined. Bootstrapping-based mediation analysis was employed to test H9, while an interaction-effect analysis was conducted to examine the moderating effect proposed in H10. This analytical approach enabled the study to move beyond descriptive findings and empirically evaluate the proposed explanatory model.

8.2 Study Population

The target population comprised young adults aged 18–30 years belonging to the Santal community and residing in selected Santal-dominated areas of West Bengal, India. The age range was deliberately specified because the study focused on young adults who were likely to have substantial interaction with smartphones, internet-based information sources, social-media platforms and emerging AI technologies. Restricting the study population to adults also facilitated the informed-consent process.

The study included respondents who had access to a smartphone, computer or another internet-enabled device and who had encountered or sought health-related information through digital sources. These sources included search engines, social-media platforms, websites, mobile applications, online health services and AI-based conversational systems. Respondents were recruited from selected Santal-dominated communities in West Bengal. Potential study areas included selected communities in Jhargram, Paschim Medinipur, Bankura, Purulia and Birbhum, subject to the availability of an appropriate sampling frame and fieldwork feasibility.

The operational population was therefore defined as:

"Santal adults aged 18–30 years residing in the selected study areas of West Bengal who had access to digital technologies and had encountered or sought health-related information through digital media."

This operational definition ensured that the study population corresponded directly with the principal research constructs, namely digital health-information seeking, digital health literacy, misinformation exposure, AI-based health-information use and verification behaviour, thereby enabling the proposed hypotheses to be empirically tested.

9. STUDY AREA

The study **was conducted** in selected Santal-dominated areas of West Bengal, with field sites located in Jhargram, Paschim Medinipur, Bankura, Purulia and Birbhum. The districts were selected based on Santal population concentration, accessibility, availability of eligible respondents and the feasibility of fieldwork, rather than convenience alone. Previous research on Santal communities in West Bengal highlighted the importance of community-based and bottom-up approaches to understanding rural health and healthcare-seeking behaviour. Accordingly, the selection of Santal-dominated communities strengthened the contextual relevance, geographical coverage and empirical validity of the study.

10. SAMPLING DESIGN

The study employed a multistage sampling technique to obtain an adequately distributed sample of young Santal adults from selected areas of West Bengal. In the **first stage**, 2–3 districts with substantial Santal populations were purposively selected based on population concentration and their relevance to the research objectives. In the **second stage**, administrative blocks containing substantial Santal settlements **were identified**. In the **third stage**, Santal-dominated villages or communities **were selected** using a sampling frame prepared with support from local administrative records and community-level sources. In the **fourth stage**, eligible respondents aged 18–30 years **were selected** from the identified communities through systematic or simple random sampling wherever an appropriate respondent list was available.

For the **CFA and PLS-SEM analyses**, a target sample of approximately 400–500 respondents were considered appropriate, with approximately 450 respondents targeted for the study. The final sample size was determined through an a priori statistical power analysis**,** considering the number of predictors, anticipated effect size, significance level and desired statistical power. Accordingly, the study did not rely solely on the traditional "10-times rule" for determining SEM sample adequacy.

11. INCLUSION CRITERIA

The study included respondents who identified themselves as belonging to the Santal community and were aged between 18 and 30 years. Participants were required to reside in the selected study areas of West Bengal and to have access to and experience using a **mobile** phone, smartphone, computer or other internet-enabled device. Respondents were also required to have searched for or encountered health-related information through digital media during the previous six months, ensuring that their responses reflected relatively recent information-seeking experiences. Finally, only those individuals who provided informed consent and voluntarily agreed to participate in the study were included in the final sample.

12. EXCLUSION CRITERIA

Respondents who did not satisfy the eligibility requirements were excluded from the study. Individuals below 18 years of age were excluded because the study focused exclusively on adult participants. Persons who were unable or unwilling to provide informed consent were also excluded. Temporary visitors to the selected Santal communities were excluded to ensure that the sample represented individuals residing within the study areas. In addition, individuals who had never accessed digital information sources or had no experience of encountering or seeking health-related information through digital media were excluded, as their inclusion would not have been relevant to the study's focus on digital health-information-seeking behaviour.

13. RESEARCH INSTRUMENT

A structured questionnaire should contain approximately **35–45 substantive items**, in addition to demographic questions.

A five-point Likert scale can be used:

1 = Strongly Disagree
2 = Disagree
3 = Neutral
4 = Agree
5 = Strongly Agree

The questionnaire should be translated into the language(s) necessary for field administration and back-translated to ensure conceptual equivalence.

eISSN: 2583-1194

14. VARIABLES AND MEASUREMENT MODEL

Construct	Code	Suggested Items
Digital Health Literacy	DHL	5
Trust in Digital Health Sources	TDH	5
Perceived Information Credibility	PIC	4
Health Misinformation Exposure	HME	5
AI Health Information Use	AHI	5
Trust in AI Health Information	TAI	4
Perceived Usefulness of AI	PUA	4
Health Information-Seeking Behaviour	HISB	5
Verification Behaviour	IVB	4

This gives approximately **41 measurement items**.

15. SAMPLE QUESTIONNAIRE ITEMS

A. Digital Health Literacy (DHL)

DHL1: I can find health information online when I need it.

DHL2: I can identify reliable websites for health information.

DHL3: I can compare health information obtained from different online sources.

DHL4: I can distinguish between professional health information and personal opinions online.

DHL5: I feel confident evaluating the quality of online health information.

B. Trust in Digital Health Sources (TDH)

TDH1: I generally trust health information provided by government health websites.

TDH2: I trust health information provided by qualified health professionals online.

TDH3: I trust information provided by established hospitals and health institutions.

TDH4: I consider health information shared by people I know to be trustworthy.

TDH5: I generally trust health information that I find through social media.

C. Perceived Information Credibility (PIC)

PIC1: Online health information usually provides sufficient evidence.

PIC2: I can identify whether an online health claim is supported by reliable evidence.

PIC3: Information from multiple credible sources increases my confidence in a health claim.

PIC4: I consider the source of information before accepting a health-related claim.

D. Health Misinformation Exposure (HME)

HME1: I frequently encounter health claims online that appear doubtful or misleading.

HME2: I have encountered false or exaggerated health claims on social media.

HME3: I have received health-related messages whose accuracy was uncertain.

HME4: I sometimes find contradictory information about the same health issue online.

HME5: I have encountered health information online that was later found to be incorrect.

E. AI-Based Health Information Use (AHI)

First ask:

Have you ever used an AI-based tool/chatbot such as ChatGPT, Gemini, Copilot or another AI system to obtain health-related information?

Response:

- Yes
- No
- Not sure

For those answering "Yes":

AHI1: I use AI-based tools to obtain information about health conditions.

AHI2: I use AI-based tools to understand medical terminology.

AHI3: I use AI-based tools to obtain general information about symptoms.

AHI4: I use AI-based tools to explore possible health-related questions before consulting a professional.

AHI5: AI tools have become an additional source of health information for me.

Important: The paper should distinguish **information seeking** from **medical diagnosis or treatment advice**. The study should not encourage respondents to rely on AI for clinical decisions.

16. Trust in AI Health Information

TAI1: I consider AI-generated health information useful.

TAI2: I believe AI systems can provide understandable health information.

TAI3: I feel comfortable using AI systems for general health information.

TAI4: I consider AI-generated health information sufficiently reliable for preliminary information seeking.

17. Perceived Usefulness of AI

PUA1: AI helps me find health information quickly.

PUA2: AI makes complex health information easier to understand.

PUA3: AI can help me identify additional questions to ask a health professional.

PUA4: AI reduces the time required to search for general health information.

18. Health Information-Seeking Behaviour

HISB1: I regularly search online when I have a health-related question.

HISB2: I use digital platforms to learn about diseases and health conditions.

HISB3: I search for information about preventive health practices online.

HISB4: I search online before visiting a health professional.

HISB5: I use more than one digital source when looking for health information.

19. Information Verification Behaviour

IVB1: I check the source before accepting health information found online.

IVB2: I compare health information across different sources.

IVB3: I consult a qualified health professional when online information concerns an important health decision.

IVB4: I check whether health claims are supported by evidence.

20. DEMOGRAPHIC VARIABLES

The study **collected** relevant demographic, socioeconomic and digital-access information from the respondents to understand the characteristics of the study population and examine their possible association with digital health-information-seeking behaviour. The variables collected included age, gender, educational qualification, occupation, monthly household income, marital status and geographical location. Information relating to respondents' digital access and usage was also collected, including smartphone ownership, internet access, frequency of internet use, preferred language for obtaining health information, frequency of social-media use and previous experience with AI-based tools**.** Information regarding access to health facilities was also collected to understand whether physical access to healthcare may influence digital health-information-seeking practices. To protect participants' privacy and confidentiality, the **study** did not collect unnecessary personally identifying information, such as names, contact numbers, Aadhaar numbers or exact residential addresses.

21. PILOT STUDY

Before the main survey, conduct a pilot study with approximately:

30–50 respondents

The pilot should evaluate:

- clarity;
- cultural appropriateness;
- translation;
- response time;
- reliability;
- ambiguous questions;
- respondents' understanding of "AI-based health information";
- comprehension of misinformation-related questions.

Items with poor reliability or weak conceptual clarity should be revised before the main survey.

22. RELIABILITY AND VALIDITY

8.6 Measurement Model Assessment

The measurement model was assessed to establish the reliability, convergent validity and discriminant validity of the constructs used in the study. Since the proposed research employed latent constructs such as Digital Health Literacy, Trust in Digital Health Information, Perceived Information Credibility, Health Misinformation Exposure, AI-Based Health Information Use, Trust in AI, Perceived Usefulness, Health Information-Seeking Behaviour and Verification Behaviour, measurement-model assessment was necessary before testing the structural relationships.

8.6.1 Factor Loadings

The standardised outer loading of each questionnaire item on its respective latent construct was examined. An outer loading of **0.70 or above** was considered desirable, indicating that the item adequately represented the underlying construct.

The loading can be expressed mathematically as:

$$\lambda_i = Cov(X_i, \eta) / \sqrt{[Var(X_i) \times Var(\eta)]}$$

where λ_i represents the factor loading of item *i*, X_i represents the observed item and η represents the latent construct.

Items with loadings between **0.40 and 0.70** were not automatically deleted. Their removal was considered only when deletion improved reliability or validity and was theoretically justified. Items with very low loadings were considered for removal.

8.6.2 Internal Consistency Reliability

Internal consistency was examined using **Cronbach's Alpha (α), Composite Reliability (CR) and rho_A**.

Cronbach's Alpha

Cronbach's Alpha was calculated to determine the internal consistency of the items measuring each construct:

$$\alpha = [k/(k-1)]\ [1 - (\Sigma\sigma_i^2 / \sigma_t^2)]$$

where:

- **k** = number of measurement items;
- σ_i^2 = variance of each individual item; and
- σ_t^2 = variance of the total scale.

A value of **0.70 or above** was considered acceptable for established constructs.

Composite Reliability

Composite Reliability was calculated as:

$$CR = (\Sigma\lambda_i)^2 / [(\Sigma\lambda_i)^2 + \Sigma(1-\lambda_i^2)]$$

where λ_i represents the standardised outer loading of each indicator.

A **CR value ≥ 0.70** indicated satisfactory internal consistency reliability.

rho_A

rho_A was also examined as an additional reliability estimate in the PLS-SEM analysis. Values of approximately **0.70 or higher** were considered indicative of acceptable reliability.

8.6.3 Convergent Validity

Convergent validity was assessed using the **Average Variance Extracted (AVE)**. AVE measures the proportion of variance in the observed indicators that is explained by their underlying latent construct.

The formula was:

$$AVE = \Sigma\lambda_i^2 / [\Sigma\lambda_i^2 + \Sigma(1-\lambda_i^2)]$$

An **AVE value ≥ 0.50** was considered satisfactory, indicating that the construct explained at least 50% of the variance in its indicators.

8.6.4 Discriminant Validity

Discriminant validity was examined to determine whether the constructs represented conceptually distinct phenomena. Two complementary approaches were used: the **Fornell–Larcker criterion** and the **Heterotrait–Monotrait (HTMT) ratio**.

Under the Fornell–Larcker criterion, the square root of the AVE of each construct should be greater than its correlations with other constructs. Thus:

$$\sqrt{AVE_i} > r_{ij}$$

where r_{ij} represents the correlation between constructs *i* and *j*.

Natural Resources for
Human Health
eISSN: 2583-1194

The HTMT ratio was additionally examined because it provides a more sensitive assessment of discriminant validity. HTMT values below **0.85** were considered strong evidence of discriminant validity, while values below **0.90** were considered acceptable in less conservative applications.

8.6.5 Decision Criteria

The measurement model was considered satisfactory when the following conditions were broadly achieved:

Assessment	**Statistical Criterion**
Outer/Factor Loading	Preferably ≥ 0.70
Cronbach's Alpha	≥ 0.70
Composite Reliability	≥ 0.70
rho_A	≥ 0.70
AVE	≥ 0.50
Fornell–Larcker	√AVE greater than inter-construct correlations
HTMT	Preferably < 0.85; acceptable < 0.90

The measurement model was evaluated before testing the structural model and hypotheses (H1–H10)**.** Items that did not meet the recommended thresholds were reviewed in conjunction with their theoretical relevance, content validity and effect on the overall reliability and validity of the respective construct. This approach ensured that statistical criteria were not applied mechanically and that the final measurement model remained theoretically meaningful.

23. DATA ANALYSIS

The collected data were analysed using **IBM SPSS Statistics** and **SmartPLS/AMOS**. SPSS was used for data screening, descriptive statistics, reliability analysis and preliminary assessment, while SmartPLS/AMOS was used to examine the measurement and structural models. The analysis was conducted in sequential stages to empirically test the proposed hypotheses.

Stage 1: Data Screening and Descriptive Statistics

Initially, the dataset was examined for **missing values, outliers, inconsistent responses and data-entry errors**. Descriptive statistics were subsequently calculated to describe respondents' demographic characteristics and digital health-information practices.

For categorical variables, **frequency and percentage** were calculated:

Percentage (P) = (f / N) × 100

where **f** represents the observed frequency and **N** represents the total number of valid respondents.

For continuous or Likert-scale variables, the **mean (M)** and **standard deviation (SD)** were calculated.

The mean was calculated as:

M = ΣX / N

where **X** represents an individual observation and **N** represents the number of observations.

The standard deviation was calculated to determine the dispersion of responses around the mean.

Stage 2: Reliability Analysis

The internal consistency of the measurement scales was assessed using **Cronbach's Alpha (α)** and **Composite Reliability (CR)**.

Cronbach's Alpha was calculated as:

$$\alpha = [k/(k-1)]\ [1 - (\Sigma\sigma^2_i / \sigma^2_t)]$$

where **k** is the number of items, σ^2_i is the variance of each item and σ^2_t is the variance of the total scale.

A value of $\alpha \geq 0.70$ was considered acceptable for established constructs.

Composite Reliability was calculated as:

$$CR = (\Sigma\lambda_i)^2 / [(\Sigma\lambda_i)^2 + \Sigma(1-\lambda_i^2)]$$

where λ_i represents the standardised factor loading of each indicator. A CR value of approximately **0.70 or above** was considered satisfactory.

Stage 3: Measurement Model

The measurement model was evaluated using **Confirmatory Factor Analysis (CFA)** and, where PLS-SEM was used, the corresponding PLS measurement-model assessment.

The analysis examined:

1. **Indicator reliability**
2. **Internal consistency reliability**
3. **Convergent validity**
4. **Discriminant validity**

3.1 Indicator Reliability

Standardised factor loadings were examined for each questionnaire item. Loadings of approximately **0.70 or higher** were considered desirable. Items with lower loadings were not removed automatically; their theoretical relevance and contribution to construct validity were considered before any deletion.

3.2 Convergent Validity

Convergent validity was assessed using **Average Variance Extracted (AVE)**:

$$AVE = \Sigma\lambda_i^2 / n$$

where λ_i represents the standardised factor loading and **n** represents the number of indicators.

An **AVE ≥ 0.50** indicates that a construct explains at least 50% of the variance of its indicators.

3.3 Discriminant Validity

Discriminant validity was assessed using the **Fornell–Larcker criterion** and **Heterotrait–Monotrait ratio (HTMT)**.

Under the Fornell–Larcker criterion:

√AVE of each construct > its correlations with other constructs.

The HTMT ratio was additionally examined, with values generally below **0.85–0.90** considered supportive of discriminant validity, depending on the conceptual similarity of the constructs.

Stage 4: Structural Model

After establishing satisfactory reliability and validity, the structural model was estimated to test the hypothesised relationships.

The general structural equation can be represented as:

$$\eta = \beta\eta + \Gamma\xi + \zeta$$

where:

- η = endogenous latent variable;

- ξ = exogenous latent variable;
- β = relationship among endogenous variables;
- Γ = effect of exogenous variables on endogenous variables; and
- ζ = residual/error term.

For the present study, the principal equation for health-information-seeking behaviour can be expressed as:

$$\mathbf{HISB} = \beta_1\mathbf{DHL} + \beta_2\mathbf{TDH} + \beta_3\mathbf{PIC} + \beta_4\mathbf{AHI} + \varepsilon$$

where:

- **HISB** = Health Information-Seeking Behaviour;
- **DHL** = Digital Health Literacy;
- **TDH** = Trust in Digital Health Information;
- **PIC** = Perceived Information Credibility;
- **AHI** = AI-Based Health-Information Use; and
- ε = unexplained error.

A separate equation can be specified for verification behaviour:

$$\mathbf{IVB} = \beta_5\mathbf{DHL} + \beta_6\mathbf{HME} + \varepsilon$$

where:

- **IVB** = Information Verification Behaviour;
- **DHL** = Digital Health Literacy; and
- **HME** = Health Misinformation Exposure.

For AI-based information use:

$$\mathbf{AIU} = \beta_7\mathbf{TAI} + \beta_8\mathbf{PUA} + \varepsilon$$

where:

- **AIU** = AI-Based Health-Information Use;
- **TAI** = Trust in AI-Based Health Information; and
- **PUA** = Perceived Usefulness of AI-Based Health Information.

4.1 Path Coefficients

The hypothesised relationships were assessed using **standardised path coefficients (β)**.

A positive β indicates a positive relationship, whereas a negative β indicates an inverse relationship.

For example:

DHL → HISB: $\beta = 0.35$

would indicate that higher digital health literacy is associated with higher health-information-seeking behaviour, subject to statistical significance.

4.2 Significance Testing

The statistical significance of the hypothesised relationships was assessed using **bootstrapping**.

The t-statistic can be expressed as:

$$t = \beta / SE(\beta)$$

where β represents the estimated path coefficient and **SE(β)** represents its standard error.

A relationship was considered statistically significant when:

$$p < 0.05$$

The decision rule was:

$p < 0.05$ → Hypothesis supported

$p \geq 0.05$ → Hypothesis not supported

For greater robustness, 95% bootstrap confidence intervals were also reported. A mediation or path effect was considered statistically significant when its bootstrap confidence interval did not include zero.

Stage 5: Coefficient of Determination (R^2)

The explanatory power of the structural model was evaluated using **R^2**.

The coefficient of determination is expressed as:

$$R^2 = 1 - (SSres / SStotal)$$

where:

- **SSres** = residual sum of squares; and
- **SStotal** = total sum of squares.

R^2 indicates the proportion of variance in an endogenous construct explained by its predictors.

For example, an:

$$R^2 = 0.45$$

would indicate that the predictors collectively explain **45% of the variance** in the dependent construct.

Stage 6: Effect Size (f^2)

The individual contribution of each predictor was examined using **effect size (f^2)**:

$$f^2 = (R^2included - R^2excluded) / (1 - R^2included)$$

This determines how strongly an individual predictor contributes to the explanatory power of the endogenous variable.

The commonly used interpretation is approximately:

- **0.02 = small effect**
- **0.15 = medium effect**
- **0.35 = large effect**

These values should be interpreted in conjunction with the substantive context of the study.

Stage 7: Predictive Relevance (Q^2)

The predictive relevance of the model was assessed using **Stone–Geisser's Q^2**, particularly through a blindfolding procedure where appropriate for the selected PLS-SEM model.

The general expression is:

$$Q^2 = 1 - (SSE / SSO)$$

where:

- **SSE** = sum of squared prediction errors; and
- **SSO** = sum of squared observations.

A **$Q^2 > 0$** indicates predictive relevance for the endogenous construct.

Where the selected software/method does not support traditional blindfolding for the specific model, an appropriate out-of-sample predictive assessment should be reported instead.

Stage 8: Mediation Analysis

Hypothesis **H9** proposes that information-verification behaviour mediates the relationship between digital health literacy and responsible health-information use.

The proposed mediation model is:

Digital Health Literacy → Verification Behaviour → Responsible Health-Information Use

The **indirect effect** is calculated as:

Indirect Effect = a × b

where:

- **a** = effect of Digital Health Literacy on Verification Behaviour;
- **b** = effect of Verification Behaviour on Responsible Health-Information Use.

The total effect is:

Total Effect = Direct Effect + Indirect Effect

Bootstrapping with, preferably, **5,000 resamples** will be used to estimate the indirect effect and its 95% confidence interval.

If the confidence interval **does not include zero**, the indirect effect will be considered statistically significant.

Stage 9: Moderation Analysis

Hypothesis **H10** proposes that digital health literacy moderates the relationship between trust in digital health information and health-information-seeking behaviour.

The moderation equation is:

$$\mathbf{HISB = \beta_0 + \beta_1 TDH + \beta_2 DHL + \beta_3 (TDH \times DHL) + \varepsilon}$$

where:

- **TDH** = Trust in Digital Health Information;
- **DHL** = Digital Health Literacy;
- **TDH × DHL** = interaction term; and
- β_3 = moderation effect.

The hypothesis will be supported if the **interaction coefficient (β_3)** is statistically significant.

The interaction should subsequently be examined using a **simple-slope analysis** to determine whether the relationship between trust and information seeking differs at low, medium and high levels of digital health literacy.

Stage 10: Group Comparison

Where the final sample size permits, subgroup differences may be examined according to:

- gender;
- educational qualification;
- level of digital access;
- frequency of internet use; and
- AI-user versus non-AI-user status.

For latent-variable comparisons using SEM, measurement invariance should first be established. In PLS-SEM, the **MICOM procedure** may be used before conducting multi-group analysis. In covariance-based SEM, appropriate measurement-invariance testing should be conducted before comparing structural relationships.

For example:

AI Users vs. Non-AI Users

can be compared to determine whether the relationships between trust, perceived usefulness and information-seeking differ between the two groups.

25. EMPIRICAL RESULTS: ILLUSTRATIVE STATISTICAL MODEL

25.1 Respondent Profile

For illustration, a simulated dataset of 450 young Santal adults aged 18–30 years was considered. The sample comprised 252 males (56.0%) and 198 females (44.0%). The mean age of respondents was **24.1 years (SD = 3.42)**. Regarding education, 28.4% had completed higher secondary education, 46.7% had undergraduate qualifications, and 24.9% had postgraduate or equivalent qualifications.

25.2 Digital Access

In the illustrative dataset, 414 respondents (92.0%) reported smartphone ownership, while 36 (8.0%) did not own a smartphone but reported access to a shared digital device. **389** respondents (86.4%) reported daily internet use, whereas 61 (13.6%) accessed the internet less frequently.

The preferred language for obtaining health information was Bengali for **247 respondents (54.9%)**, Santali for **126 (28.0%)**, and English for **77 (17.1%)**. Social-media use was reported by **397 respondents (88.2%)**.

25.3 Health-Information Sources

Health-information source	Frequency (n)	Percentage (%)
Search engines	331	73.6
Social media	309	68.7
YouTube	287	63.8
Government websites	143	31.8
Health professionals	278	61.8
Family/friends	246	54.7
AI chatbots	156	34.7

Note: Multiple responses were permitted; therefore, percentages do not total 100%.

In this illustrative dataset, search engines represented the most frequently reported digital source (73.6%), followed by social media (68.7%) and YouTube (63.8%). AI chatbots were reported by 34.7% of respondents.

eISSN: 2583-1194

25.4 AI-Based Health-Information Use

Among the 450 respondents, 156 (34.7%) reported having used an AI-based system for health-related information, while 294 (65.3%) reported no such use.

Among AI users:

Purpose of AI use	n	%
Understanding symptoms/general health issues	119	76.3
Understanding medical terminology	101	64.7
General preventive-health information	94	60.3
Preparing questions for a health professional	72	46.2
Treatment-related information	51	32.7

Among AI users, 71 respondents (45.5%) reported checking AI-generated information against another source, whereas 85 (54.5%) did not consistently undertake independent verification.

25.5 Measurement Model

The constructs were assessed using Likert-scale measurement items. Reliability and convergent validity were evaluated through Cronbach's alpha (α), Composite Reliability (CR), and Average Variance Extracted (AVE).

The mathematical expression for Cronbach's alpha is:

$$\alpha = [k/(k-1)]\ [1 - (\Sigma\sigma^2_i/\sigma^2_t)]$$

where k represents the number of items, σ^2_i represents individual-item variance and σ^2_t represents total-score variance.

Composite reliability may be expressed as:

$$CR = (\Sigma\lambda_i)^2 / [(\Sigma\lambda_i)^2 + \Sigma(1-\lambda_i^2)]$$

where λ_i represents the standardised factor loading.

AVE is calculated as:

$$AVE = \Sigma\lambda_i^2 / n$$

where n represents the number of indicators.

Illustrative Reliability and Validity Results:

Construct	Cronbach's α	CR	AVE
Digital Health Literacy (DHL)	0.893	0.921	0.701
Trust in Digital Health (TDH)	0.888	0.918	0.692
Perceived Information Credibility (PIC)	0.865	0.909	0.714
Health Misinformation Exposure (HME)	0.881	0.912	0.675
AI Health Information Use (AHI)	0.892	0.919	0.694
Trust in AI Health Information (TAI)	0.873	0.914	0.727
Perceived Usefulness of AI (PUA)	0.857	0.903	0.700
Health Information-Seeking Behaviour (HISB)	0.902	0.927	0.718
Information Verification Behaviour (IVB)	0.841	0.895	0.681

Natural Resources for
Human Health
eISSN: 2583-1194

All illustrative Cronbach's alpha and CR values exceed the commonly applied **0.70 threshold**, while all AVE values exceed **0.50**, indicating satisfactory internal consistency and convergent validity.

25.6 Structural Model and Hypothesis Testing

The structural model can be estimated using **PLS-SEM**. For each hypothesised path, the standardised path coefficient (β), bootstrapped *t*-statistic and *p*-value should be reported.

The basic structural equation for health-information seeking can be represented as:

$$\mathbf{HISB} = \beta_0 + \beta_1\mathbf{DHL} + \beta_2\mathbf{TDH} + \beta_3\mathbf{PIC} + \beta_4\mathbf{AHI} + \varepsilon$$

For verification behaviour:

$$\mathbf{IVB} = \beta_0 + \beta_1\mathbf{DHL} + \beta_2\mathbf{HME} + \varepsilon$$

For AI information use:

$$\mathbf{AHI} = \beta_0 + \beta_1\mathbf{TAI} + \beta_2\mathbf{PUA} + \varepsilon$$

Illustrative Hypothesis-Test Results

Hypothesis	Relationship	β	t	p	Decision
H1	DHL → HISB	0.279	6.952	<0.001	Supported
H2	TDH → HISB	0.394	9.125	<0.001	Supported
H3	PIC → HISB	0.173	4.377	<0.001	Supported
H4	DHL → IVB	0.400	9.928	<0.001	Supported
H5	HME → IVB	−0.247	−6.086	<0.001	Supported
H6	TAI → AHI	0.341	8.070	<0.001	Supported
H7	PUA → AHI	0.279	6.430	<0.001	Supported
H8	AHI → HISB	0.202	5.445	<0.001	Supported
H9	DHL → IVB → Responsible Use	0.244*	—	<0.001*	Supported*
H10	DHL × TDH → HISB	0.002	0.043	0.966	Not Supported

*Illustrative indirect effect based on simulated mediation analysis; the final paper should report the bootstrapped indirect effect, standard error, confidence interval and VAF, rather than only β.

25.7 Interpretation of the Illustrative Structural Model

The simulated results indicate that trust in digital health-information sources ($\beta = 0.394$, $p < 0.001$) would be the strongest direct predictor of health-information-seeking behaviour among the principal predictors. Digital health literacy also demonstrated a significant positive association with information seeking ($\beta = 0.279$, $p < 0.001$).

Perceived information credibility had a positive and statistically significant relationship with health-information seeking ($\beta = 0.173$, $p < 0.001$). Digital health literacy was strongly associated with verification behaviour ($\beta = 0.400$, $p < 0.001$), suggesting that individuals with greater digital health literacy may be more likely to evaluate and verify health information.

Health misinformation exposure demonstrated a negative relationship with verification behaviour ($\beta = -0.247$, $p < 0.001$). This simulated result would be consistent with the proposed argument that repeated exposure to misleading information may undermine effective verification practices.

Within the AI component, trust in AI-based health information ($\beta = 0.341$, $p < 0.001$) and perceived usefulness of AI ($\beta = 0.279$, $p < 0.001$) were positively associated with AI-based health-information use. AI use subsequently showed a significant positive relationship with overall digital health-information seeking ($\beta = 0.202$, $p < 0.001$).

The moderation effect in H10 was not statistically significant in this illustrative model ($\beta = 0.002$, $p = 0.966$). Therefore, digital health literacy did not significantly moderate the relationship between trust and health-information seeking in the simulated dataset.

25.8 Model Explanatory Power

The simulated regression/structural model produced an illustrative:

- **$R^2 = 0.438$ for HISB**
- **$R^2 = 0.274$ for IVB**
- **$R^2 = 0.250$ for AHI**

Thus, approximately 43.8% of the variance in health-information-seeking behaviour, 27.4% of the variance in verification behaviour and 25.0% of the variance in AI-based health-information use was explained by the respective predictors in the simulated model. The final manuscript should additionally report f^2, Q^2, SRMR, confidence intervals and predictive assessment, depending on the SEM software and estimation procedure used.

26. DISCUSSION

The discussion should focus on empirical relationships rather than merely reporting statistics. Digital health-information **access** should examine whether young Santal adults rely mainly on search engines, social media, health professionals, health websites, or community networks. **Trust** should compare confidence in doctors, government sources, hospitals, family, social media, and AI. Misinformation should assess whether exposure encourages verification or produces uncertainty and uncritical acceptance. **AI** should examine its emerging role as a health-information source and whether users verify AI-generated information. Finally, digital health literacy should test whether higher literacy reduces misinformation vulnerability and increases verification behaviour.

27. POLICY AND PRACTICAL IMPLICATIONS

The findings could have implications for:

27.1 Digital health literacy

Develop community-based digital health-literacy programmes focusing on:

i. Source evaluation;

ii. Misinformation detection;

iii. Evidence verification;

iv. Responsible AI use.

27.2 Community health workers

ASHAs and other frontline health workers could be trained to address common online health misinformation encountered within communities.

27.3 Local-language information

Reliable health-information materials should be made available in languages accessible to Santal communities, including appropriate use of Santali and Ol Chiki where relevant.

27.4 AI health information

AI-based systems should be designed to:

i. Clearly disclose uncertainty;

ii. Provide reliable sources;

iii. Discourage self-diagnosis;

iv. Encourage professional consultation;
v. Support multilingual access.

27.5 Community information systems

Community-level digital information services could connect:

Young adults → reliable information → health workers → health institutions

rather than leaving users dependent on unverified social-media information.

28. Ethical Considerations

Because the research involves human participants and health-related information, ethical safeguards are essential.

The study should obtain:

1. Institutional Ethics Committee approval before data collection;
2. informed consent;
3. voluntary participation;
4. confidentiality;
5. anonymity/pseudonymisation;
6. the right to withdraw;
7. secure data storage.

The questionnaire should not collect unnecessary medical diagnoses or sensitive health histories. Participants should also be informed that the study is not providing medical advice.

29. LIMITATIONS

The study may have several limitations.

First, the cross-sectional design will limit causal inference.

Second, self-reported information-seeking behaviour may be affected by recall and social-desirability bias.

Third, the study may not represent all Santal communities across West Bengal.

Fourth, digital access may vary substantially between villages.

Fifth, AI technologies change rapidly, meaning patterns of AI-based health-information use may change over time.

Finally, the study measures respondents' perceptions of misinformation and credibility rather than independently verifying every health claim encountered by participants.

30. CONCLUSION

Digital transformation is changing how young adults access and evaluate health information. For communities experiencing differences in access to health services, education and digital resources, these changes create both opportunities and risks. The proposed study focuses on young adults in the Santal community of West Bengal and examines the intersection of health information-seeking behaviour, trust, misinformation, digital health literacy and AI-based health information.

The empirical model proposed in this study moves beyond measuring simple internet use by examining the mechanisms through which individuals find, trust, evaluate, verify and potentially act upon health information. The study may therefore contribute to the emerging literature on digital health information behaviour while providing evidence for culturally responsive community information services.

REFERENCES:

[1] Basnyat, I., Nekmat, E., Jiang, S., & Lin, J. (2018). Applying the modified Comprehensive Model of Information Seeking to online health information seeking in the context of India. Journal of Health Communication, 23(6), 563–572.

[2] Srivastava, P., & Srivastava, M. (2026). Patterns and predictors of online health information seeking and credibility assessment among youths in India. Discover Public Health, 23, 534.

[3] Taraphdar, P., Vasudeva, A., & Sheikh, N. A. (2022). Understanding health care seeking behavior in a tribal setting in West Bengal. Journal of Family Medicine and Primary Care, 11(4), 1443–1449.

[4] Widmer, D., Ouvrard, P., Vaucher, C., et al. (2020). Thinking rural health in Santal community in West Bengal: An interprofessional bottom-up approach to rural health. Health Systems & Reform, 6(1).

[5] Office of the Registrar General & Census Commissioner, India. (2011). Population Census 2011: Scheduled Tribe population tables, West Bengal.

[6] National Health Mission, Ministry of Health & Family Welfare, Government of India. (2026). Report of the Expert Committee on Tribal Health.

[7] Online Health Information–Seeking in the Era of Large Language Models: Cross-Sectional Web-Based Survey Study. (2025). Journal of Medical Internet Research, 27, e68560.

[8] Exploring Online Health Information–Seeking Behavior Among Young Adults: Scoping Review. (2025). Journal of Medical Internet Research, 27, e70379.