

Digital Health Literacy and Ai-Generated Health Information: Its Influence on Health Decision-Making Among Adolescents

Sakshi Kumar¹, Pooja Saxena², Anurag Tewari³, Vallam Reddy Bhargavi Reddy⁴, Syed Tanzeel Rabani⁵, Arpita Singh⁶, Desh Ratan⁷

¹Assistant Professor & HOD, Department of MCA, Noida Institute of Engineering and Technology, Greater Noida. Email: sachdevasakshi8@gmail.com

²Assistant Professor, Department of Computer Science, GNIOT Institute of Professional Studies, Greater Noida.

saxena06pooja@gmail.com

³Associate Professor, Department of School of Computing and AI, Lovely Professional University, Phagwara, Punjab. Email: tiwarianurag23dec@gmail.com

⁴Assistant Professor, Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Andhra Pradesh, 522302, India. Email: vallamreddybhargavireddy@gmail.com

⁵Assistant Professor, Department of CSE APEX, CGC University, Mohali. Email: syed.rabani91@gmail.com

⁶Assistant professor, Department of CSE (AI&ML), Rajkumar Goel Institute of Technology, 5 KM Stone, Delhi-Meerut Road, Ghaziabad, Uttar Pradesh. Email: singharpita2009@gmail.com

⁷Assistant Professor, Department of Management, Mangalmay Institute of Management and Technology Greater Noida. Email: desh.ratan09@gmail.com

ABSTRACT

Generative artificial intelligence (AI) is reshaping the landscape of health information access and adoption for adolescents. Even fluent systems can be used to explain something tricky, but a fluent response can hide unsupported claims, lack context, or may give false confidence. This narrative review aims to explore the potential impact of digital health literacy on adolescents' evaluations of AI-generated health information and their decision-making processes. This draws on literacy instruments, reviews of adolescent online information appraisal and trust, surveys of chatbot use, response-safety evaluations, observational studies, participatory research, and international advice. The literature shows that there is some degree of imbalance in the way people are digitally health literate, but helpful and understandable answers are not necessarily right or safe. Observational evidence shows a correlation between health knowledge acquisition via AI and subsequently physical activity, but fails to confirm causation or directly assess quality of decision. The review indicates that the exposure to comprehension, credibility, independent verification, calibrated trust and action is influenced by the quality of the system, emotional condition, social support, privacy, and access to health care. The existing data does not specify that digital health literacy has a causal protective effect against adolescents who get harmful AI-informed guidance. Responses should demonstrate competency in applying appraisal skills, explain and secure AI systems and provide support that is easily accessible. However, it remains important to measure verification and decision quality prospectively and to separate the literacy intervention effects from improvements in AI design in future studies.

Keywords: Digital Health Literacy, Generative Artificial Intelligence, Adolescents, Health Information, Decision-Making, Trust Calibration, Misinformation.

1. INTRODUCTION

1.1 Background of the Study

As adolescents navigate digital environments where search engines, social media, video platforms, and conversational systems blend, they are more likely to come across health claims. A young person can view a nutrition claim on a brief video, engage a chatbot to break it down for them, and then share it back with their peers. The source, tone, and authority of information may shift with each transition. Generative artificial intelligence (AI) thus adds another means of engaging with health information: the way explanations are made personal, uncertainty conveyed and a claim made relevant to a young person's decision.

Previous reviews summarise the varied online health information needs of adolescents, and how they access and evaluate health information when seeking it online (Freeman et al., 2018; Park and Kwon, 2018). Studies related to trust suggest that young people can feel untrustful about information on the Internet even as they keep consuming it, and that their assessment of the trustworthiness of the information can influence their behavior (Freeman et al., 2020). This existing chat is beefed up with asking follow up questions in real time and adjusting the language to make it more relevant. These features can assist adolescents with understanding jargon or with planning for a clinical consultation. But this is only useful if the answer is correct, complete, relevant to their particular query, and includes suitable restrictions.

1.2 Problem Statement

One of the challenges is that technology use, confidence, satisfaction, perceived helpfulness, and informed decision-making are mutually related but separate outcomes. An adolescent can know the wrong answer, feel good about it, and put off getting more information. On the other hand, uncertainty can raise an initial feeling of worry and bring about a better decision. Evaluations based simply on levels of confidence or satisfaction may thus misrepresent benefit or harm.

Digital health literacy offers a valuable lens to explore the challenge but many tools are outdated in relation to the advent of generative AI. Conversational systems can generate plausible explanations and references without assuring reliable, context appropriate information. It is unclear whether traditional literacy skills are sufficient to support decision making in these environments and whether literacy-related behaviours of appreciation/confirmation are indicative of the quality of the adolescent's subsequent choice.

1.3 Objectives of the Study

The goals of this review are fourfold:

1. to explore the skills needed to assess AI-generated health information,
2. to summarise evidence on adolescent digital health literacy, online trust, and chatbot engagement and response safety,
3. to outline a conceptual model relating adolescent appraisal of health information created by ChatGPT to verification, calibration of trust, and health-related action, and
4. to identify educational, clinical, design, and research priorities.

The biggest contribution does not lie in the perceived assistance being equal to the decision quality demonstrated, since verification and seeking help is a visible behaviour that cannot be measured directly.

1.4 Research Questions

The analysis focuses on three questions:

1. How do established digital health literacy concepts apply to AI-generated health information?
2. How could literacy feed into credible judgments, verification, trust and following through by the adolescent?
3. How much evidence and safeguards are required to support assertions that health decision-making has been improved?

Adolescents are defined as individuals between the ages of 10 to 19 (World Health Organization [WHO] 2020). Where appropriate, studies of younger children or young adults are noted, but not necessarily assumed to be equivalent to the adolescent population.

2. REVIEW OF LITERATURE

2.1 Digital Health Literacy and Adolescent Competence

The eHealth Literacy Scale (eHEALS) was created by Norman and Skinner (2006) to measure perceived skills in locating, evaluating, and using electronic health information. Practical value in this is that it can highlight any confidence or perceived difficulty; however, a self-report score is not indicative of whether a person will identify an unsupported claim when it actually happens during an information encounter. The distinction is important when conversational fluency can produce intelligently sounding output that does not enhance output correctness.

The Digital Health Literacy Instrument (DHLI) measures operational skills, navigation skills, searching skills, appraisal of information reliability, relevance, content generation, and privacy (van der Vaart & Drossaert, 2017). These domains offer a helpful starting point for looking at AI-mediated health information. That is not without adaptation: researchers need to verify that adolescents are able to find missing background, verify the evidence they cite, identify uncertainty, understand privacy implications and know when a chatbot is not a suitable source. Without validation, there is a risk of assuming that an established instrument measures AI-specific competence.

2.2 Online Health Information, Trust, and Adolescent Contexts

Freeman et al. (2018) conducted a systematic review of the way that adolescents aged 13-18 search for and appraise health information online. Park and Kwon (2018) conducted a review of health-related Internet use among children and adolescents, including evidence that was somewhat older than the typical age of adolescents. Overall, these reviews address various information needs and limitations in appraisal, as well as gaps in the underlying research. When it comes to media platforms, knowing a little does not mean you can make good judgements and use it safely.

In particular, Freeman et al. (2020) examined trust in the appraisal of online health information among adolescents. They synthesized 22 studies that did focus on young people (13–18), and they found patterns of continued use of online information despite distrust, reliance on heuristics in assessing trustworthiness, differences in trust between websites and social media, and connections between trust and subsequent response. The present review is relevant, since it links up trust judgements with action. But it's pretty much just about traditional online information instead of current gen AI. Translating these findings into using chatbots with teens is then a logical follow-on, but not direct proof that they may do so in the same manner.

In 2022, Stauch et al. (2025) conducted a representative study in Germany with 1,448 participants aged 9–18 years. 30.76% (1.362) of participants were classified as having problematic digital health literacy, while 4.63% (33) were

classified as having inadequate digital health literacy. Some of the challenges were related to privacy and navigation, as well as literacy, which varied by sociodemographic. These findings reflect variation and potential inequity; however, they assess overall digital health literacy instead of performance in evaluating current AI-based health advice. For all adolescents and systems, direct interpretation is limited because of the inclusion of children under age 10 and the collection period.

Information from India offers a complementary yet indirect perspective on the verification behaviour. Srivastava & Srivastava (2026) carried out a hospital-based cross-sectional study involving 196 young adults aged between 16 to 24 years in Varanasi. 87.8% said they looked for health information online, but only 5.1% actively checked for credibility and 68.9% did not check for sources. Youth (school and dropout participants), younger participants, and participants who owned their device were also less likely to check the information. While this study is relevant to the local context, it did not focus on AI-generated information; it was a single-hospital, mixed-age sample, so generalisation to the rest of the sample of adolescents in India is limited.

2.3 AI-Generated Information, Credibility, and Response Safety

Ayers et al (2023) asked physicians and a chatbot to respond to 195 patient questions in a public online forum. There were 585 evaluator ratings, with evaluators preferring the chatbot 78.6% of the time and positively rating chatbot answers for quality and empathy. This finding illustrates the communicative potential of generated responses for a given evaluation design. It does not say that these responses were more clinically accurate, that adolescents would experience them — or value them — in the same way, or that their use would enhance health. The chatbot answers were also considerably longer than those of physicians, which makes simple comparisons more difficult.

Brewster et al. (2025) assessed 75 chatbot answers from 25 consumer chatbots on simulated emergent adolescent queries about health. 81.3% of all responses were rated as highly understandable, and 46.7% were rated as highly appropriate. This comparison reveals that understandable language and appropriate supports are not the same thing. Performance also varied according to the type of chatbot. While these scenarios are simulated and do not cover every product, version, or actual conversation, they demonstrate that readability or perceived empathy can't be the only safety measure.

Chen et al. (2026) conducted a study to assess the perceived trustworthiness of AI voices in health-learning modules among 104 adolescents (11–17 years old) and their parents/guardians. For all the modules assessed, social impression and sound appeal accounted for 71% to 75% of the variance in the ratings of trustworthiness. This finding indicates that interface presentation can impact perceived trust, and the end result was a rating of trust rather than accurate information, safe reliance, or future health choices. It also involved AI voices in learning modules and should not be interpreted as a direct test of text-based generative health advice.

The WHO advice on ethics and governance of large multi-modal models defines risks which are relevant to health settings, including unreliable outputs, bias, automation bias, privacy and cybersecurity (WHO, 2024). The guidance was not specific to adolescents and reaffirms that system design and governance should be considered along with user knowledge. Any strategy that focuses exclusively on literacy risks overloading youth with the responsibility to self-identify failures that may not be within their scope or means to detect.

2.4 Health Decisions and Emerging Behavioural Evidence

McBain et al. (2026) presented findings on a nationally representative survey of 1,009 people in the United States between the ages of 12 and 21. Based on weighted estimates, 19.2% had consulted an AI chatbot for mental health information. Of users, 91.7% rated the advice as somewhat to very helpful, and 63.3% had said they had revealed this use to no one. The results are based on reported participation and perceived usefulness, rather than the correctness of the information or therapeutic efficacy. It includes young adults, and the self-reported measures do not indicate how information was verified or if the advice influenced follow-up choices.

Feng et al. (2025) conducted a meta-analysis of randomised controlled trials of AI chatbots to reduce mental distress and for health behaviour change, in individuals aged 15–39 years. They found 31 trials, including 29,637 participants, of which 26 studies were included in meta-analyses. Pooled effects indicated a small-to-moderate effect on mental distress and a small effect on health behaviours, and the review found that evidence on the effectiveness of generative systems was inconclusive. The age range included was large and adult populations are eligible as well, and the types of interventions were not well controlled for in this study, so these findings do not specifically reflect the effectiveness of the interventions for adolescents utilizing AI for individual health choices.

Sun et al., (2026) analyzed 2,306 questionnaires distributed to secondary school students among 10 schools in five cities of southern China. The direct path from AI-related health knowledge acquisition to physical activity was significant and positive, while exercise self-efficacy served as an intervening variable that mediated some of the relationship, and health information literacy appeared to act as a moderator for one path in the model. This is a straightforward link between

knowledge/health literacy and an AI related health behaviour but not a direct test of decision quality or clinical reliability. No causal claims, including temporal causality, ruling out reverse causality, or effects caused by the AI information, can be made from the cross-sectional design. Data should be interpreted as relationships that need to be confirmed in future studies.

There is a participatory example by Aloï et al. (2025) who co-designed an online educational tool to enhance digital health literacy among adolescents. A total of 44 young people participated in initial workshops and 8 in beta testing. This work validates the importance of engaging adolescents in resource development and contributes to the issues of concern to adolescents: influencers, family misinformation, and AI tools. Positive feedback and acceptability are not synonymous with proven long-term improvements in appraisal competence or the quality of health decisions.

2.5 Research Gap and Analytical Propositions

The primary gap is between the appraising and the doing of consequential choices based on that. There are parallel sections of the pathway in the current evidence: measurement corresponds to literacy; reviews correspond to trust and information-seeking; surveys correspond to chatbot engagement; evaluation corresponds to properties of chatbot responses. Fewer studies have seen all of these variables – demonstrated literacy, verification, calibrated trust, adolescent decisions – together. The present review, therefore, constructs three propositions for further testing, but does not presume effect.

P1: Self-reported confidence is not as effective as demonstrated appraisal competence in predicting the behaviour of verification. P2: Associations of digital health literacy and decision quality may be mediated by verification and calibrated trust. P3: Emotional vulnerability and barriers to professional support can also diminish the protective role of literacy. These are ideas and not tested. Their worth lies in being converted to measurable constructs and the testing of alternative explanations such as previous health knowledge, socioeconomic factors, and access to care.

3. RESEARCH METHODOLOGY

3.1 Research Design and Selection of Studies

This study uses a critical narrative review with conceptual synthesis. The design reflects an emerging topic with evidence ranging from instrument-development studies, systematic reviews, surveys, observational research, simulated chatbot evaluation, participatory design, and governance guidance. The aim is to compare the things that these pieces of evidence can and cannot prove, and create an explanatory model. No participants were recruited, no original clinical data set analysed, no pooled effect size calculated.

Original manuscript included purposive selection of traceable publications that explicitly discussed adolescent digital health literacy or contact with health-related AI. When it was helpful in defining a construct or a possible mechanism, foundational measurement studies, conventional online-information reviews, a response comparison using adult-context questions, and international guidance were retained. A directly relevant systematic review of adolescents' trust in online health information was added for the revised version, and the original references were cross-referenced with the websites of the publishers, PubMed, and official WHO information. This is not an exhaustive or systematic search; sources were prioritised for direct relevance, traceability and clarity of study design.

3.2 Data Sources and Collection Procedures

Targeted web searches were conducted on 15 October 2026 with combinations of words connected to adolescents, digital health literacy, generative AI, chatbots, health information, and health decision making. Additional target web searches and source checks were conducted during the revision process to verify important publication information and conclusions and bring in terms such as trust, credibility, verification, adolescent AI-mediated health communication, and the effect of chatbot interventions. Where possible, publisher pages, PubMed records, accessible article texts, and official WHO material were consulted.

There are 16 sources in the final purposive collection: empirical studies, systematic reviews, foundational measurement papers, a meta-analysis, and policy guidance. The corpus is designed to help to develop a critical conceptual synthesis; it is not designed to estimate prevalence or pooled effect of a defined exposure. The provided manuscript made no attempt to rebuild a complete query log per database or record level exclusion log. The review is in English only, which may introduce a selection bias. No assertions are made regarding exhaustive retrieval, duplicate independent screening, preregistration, or formal risk-of-bias scoring.

3.3 Analytical Categories and Interpretive Procedure

Sources were compared on 6 analytical categories:

1. access,
2. demonstrated or perceived appraisal competence,

3. credibility, trust,
4. verification,
5. health-related decision, behaviour and
6. contextual issues such as privacy, equity, access to professional support.

For each source, the synthesis identified whether evidence pertained to self-reported perceptions, demonstrated skills, reported behaviour, observed response properties, or normative guidance. These outcome types were not viewed as interchangeable.

Conceptual propositions were created by identifying potential relationships between separately measured constructs from the literature, as well as considering alternative explanations. The synthesis represents the difference between the reported result of a source, the indirect interpretation of multiple sources, and a proposition that awaits testing. Distinctions can be made easier by using the table of source characteristics.

3.4 Reliability, Scope, and Ethical Considerations

Traceable reference to texts, clear description of the phenomena studied and the nature of the study, preservation of the sample numbers, and the distinction between findings and concepts deepen the interpretive reliability. The review did not involve systematic screening or quality-scoring of papers; as a result, it should be understood as a critical synthesis, rather than a comprehensive mapping of all of the available evidence. AI system modifications also restrict the transferability of individual chatbot evaluations.

The review relied on publicly available literature and no literature specific to adolescents was identified. Nevertheless, ethical interpretation must focus on issues of privacy, autonomy, and differential access to resources. Special attention is paid to the fact that adolescents are not solely responsible for failures in the AI. It also acknowledges that research involving minors would require age-appropriate consent or assent, privacy safeguards, ethics review, and protocols for responding to reports of distress or emergencies involving a minor's health.

4. SYNTHESIS AND DISCUSSION

4.1 Profile of the Comparative Evidence

Different types of evidence are included in the literature. Literacy instruments and population surveys report capacities and gaps; systematic reviews synthesize literature on online information-seeking and trust; Chatbot evaluations test the properties of these responses; and an observational study examines an association with physical activity; participatory work can contribute to the design of educational initiatives; international guidance focuses on system level governance. Together they are an explanation of potential influence, not evidence for or against the overall effect of AI on adolescent health.

It also shows why seemingly contradictory results should not be combined into a single result. With unsuitable outputs, high satisfaction is compatible because users' perceptions and responses to appropriateness are distinct constructs. Likewise, a correlation with literacy and behaviour may reflect background knowledge, education experiences, or access to supports. Summary information on the sources used in a review, including contributions, populations, designs, and boundaries for interpretation.

Table 1. Characteristics and Principal Contributions of the Included Literature

Source and evidence type	Population or scope	Contribution to this review	Key interpretive boundary
Aloi et al. (2025) Participatory design study	Young people involved in resource co-design; 44 in workshops and 8 in beta testing	Shows the value of adolescent participation in designing relevant digital health literacy resources.	Acceptability and co-design do not establish sustained skill gains or improved decisions.
Ayers et al. (2023) Comparative response evaluation	195 public-forum patient questions; 585 evaluations	Chatbot responses were preferred in 78.6% of evaluations and rated favourably for quality and empathy.	Preference ratings are not clinical outcomes; respondents were not a defined adolescent sample.

Source and evidence type	Population or scope	Contribution to this review	Key interpretive boundary
Brewster et al. (2025) Simulated chatbot evaluation	75 responses from 25 consumer chatbots to emergent adolescent health scenarios	Highest understandability ratings were more frequent than highest appropriateness ratings (81.3% vs. 46.7%).	Simulated prompts and tested products do not represent every model or real-world encounter.
Chen et al. (2026) Cross-sectional AI-voice study	104 participants, including adolescents aged 11–17 and parents or guardians	Social impression and sound appeal explained 71%–75% of variance in perceived voice trustworthiness.	Perceived trust is not evidence of accuracy or safe decisions; study used voice-based learning modules.
Feng et al. (2025) Systematic review and meta-analysis	31 RCTs; 29,637 participants aged 15–39; 26 studies meta-analysed	Reports overall effects of chatbot interventions on distress and health behaviours; generative-system effectiveness remained inconclusive.	Broad age range and mixed interventions limit adolescent-specific and generative-AI conclusions.
Freeman et al. (2018) Systematic review	Adolescents aged 13–18 in included studies	Reviews adolescent online health-information searching and appraisal.	Published before contemporary generative AI; evidence and methods varied across studies.
Freeman et al. (2020) Systematic review	22 studies concerning adolescents aged 13–18	Synthesises trust heuristics and links between trust in online information and adolescents' responses.	Focuses on conventional online information, not current generative AI chatbots.
McBain et al. (2026) National survey	1,009 US respondents aged 12–21	Estimates chatbot use for mental health advice and users' perceived helpfulness and disclosure.	Includes young adults; self-report does not establish advice accuracy or decision quality.
Norman and Skinner (2006) Instrument development	Foundational eHealth literacy measurement	Provides a self-report measure of perceived ability to use electronic health information.	Perceived ability is not the same as demonstrated performance; predates generative AI.
Park and Kwon (2018) Systematic review	Children and adolescents; some included evidence extends beyond age 19	Maps health-related internet use and identifies diverse needs and research limitations.	Broad age scope and pre-generative-AI publication date constrain direct application.
Srivastava and Srivastava (2026) Hospital-based cross-sectional study	196 participants aged 16–24 at one hospital in Varanasi, India	Reports widespread online health-information seeking but low self-reported credibility verification.	Mixed-age, single-hospital sample; information seeking was not specific to generative AI.
Stauch et al. (2025) Representative cross-sectional study	1,448 German participants aged 9–18; literacy classified for 1,362	Documents variation in digital health literacy and sociodemographic associations.	Data were collected in 2022 and measured general literacy rather than AI appraisal.

Source and evidence type	Population or scope	Contribution to this review	Key interpretive boundary
Sun et al. (2026) Cross-sectional questionnaire study	2,306 students from 10 secondary schools in five cities in southern China	Finds an association between AI-related health knowledge acquisition and physical activity, with modelled mediation and moderation.	Cross-sectional design cannot establish causation; physical activity is not a direct decision-quality measure.
van der Vaart and Drossaert (2017) Instrument development	Development and evaluation of the Digital Health Literacy Instrument	Broadens assessment to navigation, searching, reliability, relevance, content generation, and privacy.	AI-specific validity must be established rather than assumed.
WHO (2020) Public health definition	International definition of adolescence	Supports the age range of 10–19 used as the review’s main population definition.	Age definitions do not erase meaningful developmental and contextual differences.
WHO (2024) Governance guidance	Large multi-modal models in health and related settings	Identifies system-level concerns such as inaccurate outputs, bias, automation bias, privacy, and cybersecurity.	Guidance is normative and not an adolescent-specific outcome evaluation.

Combined, the studies argue the need to separate the constructs under comparison carefully. Evaluators might like a chatbot response, a young person might say that advice helped and a response can be understandable, neither of which can be used alone to determine that a decision was safe or well informed. The stakes of the decision also influence comparability. There is a difference between evaluating a general wellness claim and determining whether symptoms must be assessed urgently. Assessments should therefore be structured to classify the decision stakes in advance and to provide a separate performance report on safety-critical decisions, rather than average scores, which may hide significant failures.

4.2 Literacy, Trust, and Verification

Digital health literacy can influence decision-making by allowing young people to deal with uncertainty. A competent user should be able to recognise the difference between general education and personalised clinical evaluation, where there is insufficient support for an answer and where it needs to be sought elsewhere from a professional. The right goal is appropriate trust: the dependence on the quality of evidence, the short-comings of the system, and the cost of error. Blanket assurance and blanket dismissal, both are detrimental to productive interaction.

Verification is a potential bridge between comprehension and action, proposed in this review. Checking against some other source, preferably independent and appropriate, does not simply mean asking the same system to repeat or reinforce its original response. There needs to be a created citation and a supporting claim; the name of a recognised organisation does not mean that the linked content is applicable to the situation of the adolescent. Instead of just asking participants to repeat general warnings about misinformation, evaluations must test source checking, claim-to-evidence matching, context, and applicability.

An educational evaluation would have two equally readable explanations of a fictional health claim, one with the proper evidence and the other one partial or lacking evidence. Students would select the explanation, review the sources, state what they still didn't understand, and identify a reasonable next step. The task is suggested as a measurement strategy, rather than as something that has already demonstrated effectiveness in enhancing health outcomes. It should include unfamiliar examples to find out whether skills transfer outside what is instructed.

Competence is also about knowing when not to make a confident judgment. Good decision-making by an adolescent does not require guessing the diagnosis based on lack of information. It may be more relevant to recognize that a chatbot is not able to analyze the entire situation, plan the questions to be asked to a clinician, or find an immediate support than to provide a confident answer. Decision measures should thus reward legitimate uncertainty and relevant escalation of the situation where the evidence is not able to justify a definite choice.

Confidence and correctness should be measured independently by researchers at the individual level. Confidence ratings on each appraisal task will be able to recognize confidently incorrect answers and cautious correct answers. It is possible that a programme that raises average knowledge and confidence in remaining errors will provide less

protection than the overall score indicates. Likewise, merely counting the number of times a source was cited is not enough: the research should be done to see if the source was relevant, understood and reflected in the final argument.

4.3 Implications for Health Decision-Making

4.3.1 Educational Benefits and Risks of Overreliance

Conversational tools can assist with health education by providing explanations of terminology, restructuring complex information, translating material, or assisting adolescents in expressing their concerns. Such applications can be justified most when AI complements trustworthy data and maintains important qualifications. Simplification and translation are not supposed to be neutral: a clearer sentence can leave out uncertainty or delete a vital exception, or can make general information seem to apply to each case separately.

One of the risks is premature closure: accepting a seemingly comprehensive explanation without looking at missing information, alternatives, or seeking professional counsel. The repetition of the dialogue can strengthen the wrong assumption when the system reflects the words of the user or does not question the early premise. Evidence may be requested, but even an appearance of a valid-looking reference is not evidence that the referenced source exists or bears out the claim it generates.

Appraisal can also be affected by emotional urgency. Even a knowledgeable adolescent may find it hard to verify because of distress, embarrassment, fear of judgement or a need to be reassured. This is still just a hypothesis that needs to be researched. It favours the application of imaginary, age-related pedagogical situations which practice uncertainty without submitting learners to specific harmful recommendations or distressing a personal failure.

A learning sequence in a school might progress through identifying the source of a claim, verifying the supporting evidence of the claim, recording doubt, safeguarding personal information, and choosing a reasonable follow-up. Educators were able to contrast answers to the same questions with different wording and demonstrate how omission of context alters meaning. It must be assessed with unfamiliar examples and with follow-up to ensure that success is not based on short-term memory alone. This teaching will need to be integrated with health education rather than being a technical checklist alone.

4.3.2 Privacy, Equity, and Supported Autonomy

AI interactions produce disclosure decisions and health decisions. Teens can disclose symptoms, relationships, or other sensitive events since a discussion feels confidential. Privacy competence involves an awareness that information can be saved or processed, checking privacy settings where possible, and not using unnecessary identifying information. WHO recommendations on health-related large multi-model models also focus on the role of governance, risk management, and protection against harms that emerge due to model behaviour and deployment (WHO, 2024).

The feasibility of verification is also unequal. Access to reliable evidence may require language skills, available formats, good connection, time, and knowledge about local services. If the teen is able to recognize the ambiguity, that may still not mean that they have a confidential and affordable way to reach out to professional advice. Individual differences in decision quality should not be automatically explained by individual literacy, however; the research needs to quantify what support and alternatives were actually available.

Universal surveillance is not the right solution compared to a supported autonomy. Verification can be taught in schools without any personal disclosure; clinicians can ask adolescents to talk about the outputs of chatbots without mocking them; and families can assist in finding reliable sources without violating age-related privacy. Automatic parental monitoring is not a one-fit-all solution, especially when disclosure can put a youngster into danger. Protections must maintain access to confidential professional support where needed.

Institutional strategies should also look at the load they put on users. The instructions that need a long search can be inconvenient in shared devices or when the connection is weak. Available options include short snippets of evidence, easy-to-understand backlinks, clear explanations about uncertainty, as well as ways to find competent help. These suggestions need usability and outcome testing with the youths who are supposed to use them. Meaningful participation, completion, and decision outcomes should be used to evaluate equity rather than access statistics.

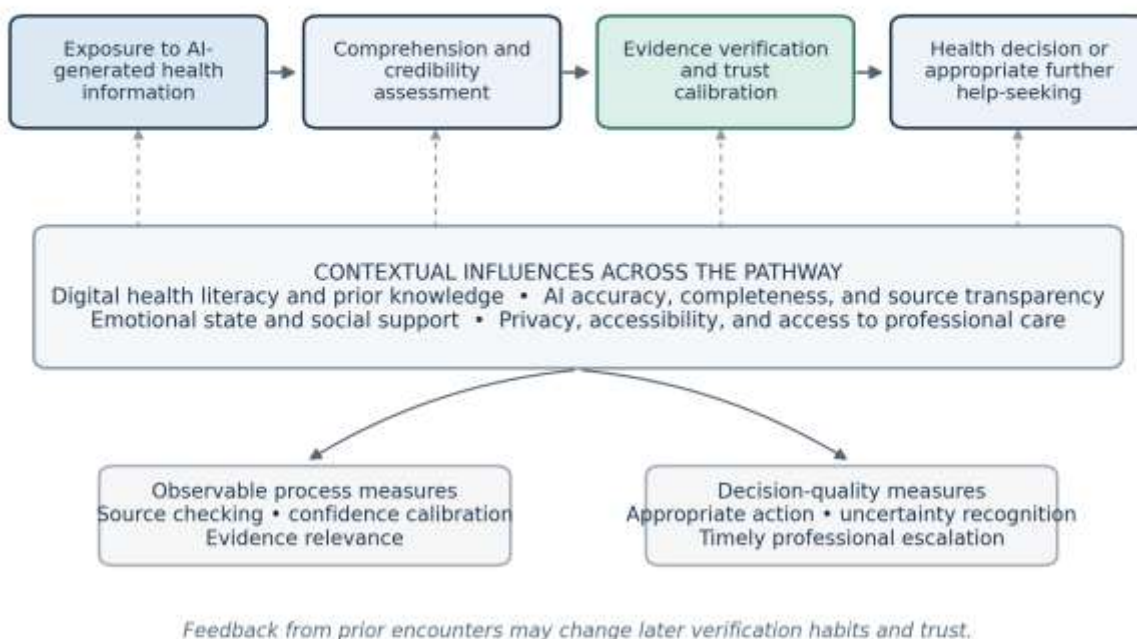
4.4 Proposed Relationship Between Literacy, AI Information, and Decisions

The presented model connects exposure to AI-generated health information to comprehension, assessment of credibility, independent verification, trust calibration, and action. Digital health literacy can affect various transitions, and the quality and completeness of information and the interface transparency can determine the requirements of the user. The pathway may also depend on emotional state, social support, privacy limitations, and availability of services. Past experience can either reinforce checking behavior or help to maintain overconfidence, so the process can be iterative instead of linear.

Figure 1. Proposed Conceptual Framework Linking Digital Health Literacy to Adolescent Health Decision-Making in AI-Mediated Environments

PROPOSED DECISION PATHWAY

A conceptual model for empirical testing, not an established causal sequence



Note. The figure presents a conceptual pathway for empirical testing. Arrows indicate proposed relationships rather than established causal effects. Contextual factors may influence more than one stage.

The framework identifies two potential roles of literacy. Moderation Literacy can modify the strength of the association between exposure to AI-generated information and acceptance of the information. Literacy can also act as an antecedent and prompt verification which can be linked to action. These are not identical statistical questions and cannot be lumped into a single generic literacy effect. Moreover, verification can be determined in part by the availability of reliable evidence. Without the appropriate information provided, a competent adolescent cannot simply analyze a source that is not given or determine an omitted qualification on their own. Transparency of the system and competence of the user should then be looked at together.

It is important to assess the quality of decisions in relation to defensible standards defined before data were collected. In educational contexts, a multidisciplinary panel might establish acceptable options, qualification requirements, or situations where additional information or clinical evaluation is needed. When it is justified by evidence or personal inclination, more than one plausible response must be permitted. Realistic scenarios and outcome measures should be developed with the contribution of adolescents. Researchers are advised to disclose scoring inconsistencies and test how inferences are altered under reasonable alternative conditions.

The framework identifies roles within multiple tiers. Educators can measure demonstrated checking, not confidence alone, clinical services can offer verification and follow-up options, and developers can offer a dependable path to human assistance, meaningful uncertainty, and privacy safeguards—all of which are clearly traceable sources. A warning or citation alone is not enough if the adolescent does not comprehend it or know how to apply it. Success must be judged by safer and better-supported decisions, not just longer engagement or more trust alone.

5. CONCLUSIONS, LIMITATIONS, AND FUTURE RESEARCH

5.1 Conclusions

The health information produced by AI can affect adolescents by affecting their understanding, credibility, assurance, and belief that they should seek additional assistance. Digital health literacy is a reasonable factor in this direction,

particularly by means of verification and calibrated trust. Nonetheless, the evidence available does not prove a general causal protective effect. Access, confidence, satisfaction, and perceived helpfulness must be independent of demonstrated appraisal competence and safe action.

Responsible response includes practical literacy instruction as well as reliable and self-assessed systems and human supports. We must acknowledge that there are some misleading responses, confusing interfaces, privacy concerns, and accessibility obstacles that simply cannot be solved with literacy alone. What this review offers, as the result, is a testable decision pathway and collective responsibility sharing, not a belief that allocating health decisions to technological familiarities will lead to informed health choices.

5.2 Limitations of the Study

Instead of systematic screening, the review relies on a purposive and English-language-based corpus. A complete database-by-database search log or record-level exclusion log was not included in the provided manuscript, nor did it have a pooled analysis or systematic risk of bias assessment. It is important to recognize that the studies sampled are of different ages, in different regions, with differing designs, and testing different outcomes so direct comparisons are limited. Other sources are self-reported perception, other sources are simulated answers, overall digital literacy, or health behaviour and not decision quality. Results of adult-context testing and mixed-age groups should not be automatically applied to adolescents. Even AI models and products evolve, making certain system-specific safety discoveries less durable. Lastly, the proposed framework has not been empirically tested and the targeted search could have overlooked relevant studies.

5.3 Future Research

Future researchers would need to measure literacy prior to or at AI exposure, report information encountered, and track follow-up decisions. Ethically designed experiments would be able to compare matched explanations which differ in the quality of evidence, uncertainty communication, source transparency, or support of verification. Detection of unsupported claims, quality of source inspection, calibration of confidence, proper referral decisions, and retention delays are all outcomes. Clinical outcomes must be assessed independently and followed up accordingly; they cannot be determined based on immediate satisfaction or desired behaviour.

Researchers are encouraged to differentiate between the impact of user-literacy interventions and safer interfaces and to compare both individually and jointly. Potential confounders or effect modifiers should be measured by prior health knowledge, experience with AI, social support, socioeconomic resources, and available care. Preregistered outcome definitions and clear scoring rules would minimize the probability of emphasizing better confidence and ignoring unchanged or worse decisions.

Future efforts should differentiate between younger and older adolescents and consider underserved linguistic, rural, and disability communities. Self-report scales should be supplemented with validated performance tasks and adolescents should assist in defining realistic scenarios and meaningful decision outcomes. Studies involving minors must have proper ethical review, consent and assent processes, privacy measures, and clear protocols to deal with distress or emergent disclosures. To ensure that future findings can be interpreted as systems change, authors should document the AI model, version, prompts, and evaluation conditions.

References

1. Aloï, M. A., Caldwell, P. H. Y., Taba, M., Kang, M., Skinner, S. R., Bray, L., & Scott, K. M. (2025). Co-designing an online educational resource to help adolescents improve their digital health literacy. *BMC Public Health*, 25, Article 1870. <https://doi.org/10.1186/s12889-025-22949-0>
2. Ayers, J. W., Poliak, A., Dredze, M., Leas, E. C., Zhu, Z., Kelley, J. B., Faix, D. J., Goodman, A. M., Longhurst, C. A., Hogarth, M., & Smith, D. M. (2023). Comparing physician and artificial intelligence chatbot responses to patient questions posted to a public social media forum. *JAMA Internal Medicine*, 183(6), 589–596. <https://doi.org/10.1001/jamainternmed.2023.1838>
3. Brewster, R. C. L., Zahedivash, A., Tse, G., Bourgeois, F., & Hadland, S. E. (2025). Characteristics and safety of consumer chatbots for emergent adolescent health concerns. *JAMA Network Open*, 8(10), Article e2539022. <https://doi.org/10.1001/jamanetworkopen.2025.39022>
4. Chen, J.-L., Lin, C.-X., Koester, K. A., Raymond-Flesch, M., Xu, L., Thapar, A. C., & Thapar, J. C. (2026). Optimizing AI voice for adolescent health: Preferences and trustworthiness across teens and parent. *Journal of Pediatric Health Care*, 40(4), 504–512. <https://doi.org/10.1016/j.pedhc.2026.03.003>
5. Feng, X., Tian, L., Ho, G. W. K., Yorke, J., & Hui, V. (2025). The effectiveness of AI chatbots in alleviating mental distress and promoting health behaviors among adolescents and young adults: Systematic review and meta-analysis. *Journal of Medical Internet Research*, 27, Article e79850. <https://doi.org/10.2196/79850>

6. Freeman, J. L., Caldwell, P. H. Y., Bennett, P. A., & Scott, K. M. (2018). How adolescents search for and appraise online health information: A systematic review. *The Journal of Pediatrics*, 195, 244–255.e1. <https://doi.org/10.1016/j.jpeds.2017.11.031>
7. Freeman, J. L., Caldwell, P. H. Y., & Scott, K. M. (2020). The role of trust when adolescents search for and appraise online health information. *The Journal of Pediatrics*, 221, 215–223.e5. <https://doi.org/10.1016/j.jpeds.2020.02.074>
8. McBain, R. K., Cantor, J. H., Breslau, J., Diliberti, M., Zhang, L. A., Zhang, F., Burnett, A., Kofner, A., Rader, B., Pataranutaporn, P., Stein, B. D., Mehrotra, A., & Yu, H. (2026). AI chatbot use and disclosure for mental health among US adolescents and young adults. *JAMA Pediatrics*, 180(8), 884–890. <https://doi.org/10.1001/jamapediatrics.2026.2015>
9. Norman, C. D., & Skinner, H. A. (2006). eHEALS: The eHealth Literacy Scale. *Journal of Medical Internet Research*, 8(4), Article e27. <https://doi.org/10.2196/jmir.8.4.e27>
10. Park, E., & Kwon, M. (2018). Health-related internet use by children and adolescents: Systematic review. *Journal of Medical Internet Research*, 20(4), Article e120. <https://doi.org/10.2196/jmir.7731>
11. Srivastava, P., & Srivastava, M. (2026). Patterns and predictors of online health information seeking and credibility assessment among youths in India. *Discover Public Health*, 23, Article 534. <https://doi.org/10.1186/s12982-026-01749-w>
12. Stauch, L., Renninger, D., Rangnow, P., Hartmann, A., Fischer, L., Dadaczynski, K., & Okan, O. (2025). Digital health literacy of children and adolescents and its association with sociodemographic factors: Representative study findings from Germany. *Journal of Medical Internet Research*, 27, Article e69170. <https://doi.org/10.2196/69170>
13. Sun, Q., Gou, H., Xiang, L., Hu, C., & Fang, Y. (2026). The impact of generative AI health knowledge acquisition on adolescents' physical activity: The mediating role of exercise self-efficacy and the moderating effect of health information literacy. *Frontiers in Public Health*, 13, Article 1705265. <https://doi.org/10.3389/fpubh.2025.1705265>
14. van der Vaart, R., & Drossaert, C. (2017). Development of the Digital Health Literacy Instrument: Measuring a broad spectrum of Health 1.0 and Health 2.0 skills. *Journal of Medical Internet Research*, 19(1), Article e27. <https://doi.org/10.2196/jmir.6709>
15. World Health Organization. (2020, October 19). Adolescent health and development. <https://www.who.int/news-room/questions-and-answers/item/adolescent-health-and-development>
16. World Health Organization. (2024). Ethics and governance of artificial intelligence for health: Guidance on large multi-modal models. <https://www.who.int/publications/i/item/9789240084759>