



## OPEN ACCESS

### EDITED BY

Marvin Kopka,  
Technical University of Berlin, Germany

### REVIEWED BY

Tze Wei Liew,  
Multimedia University, Malaysia  
Arman Miri,  
Hamadan University of Medical  
Sciences, Iran

### \*CORRESPONDENCE

Fuzhong Wu  
✉ 674656753@qq.com

RECEIVED 15 December 2025

REVISED 24 February 2026

ACCEPTED 25 February 2026

PUBLISHED 09 March 2026

### CITATION

Wang Y, Wu F and Ni B (2026) The detachment of responsibility: how and when AI-powered digital avatars undermine behavioral intentions in cancer screening persuasion. *Front. Commun.* 11:1767975. doi: 10.3389/fcomm.2026.1767975

### COPYRIGHT

© 2026 Wang, Wu and Ni. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

# The detachment of responsibility: how and when AI-powered digital avatars undermine behavioral intentions in cancer screening persuasion

Yixi Wang<sup>1</sup>, Fuzhong Wu<sup>2\*</sup> and Binbin Ni<sup>3</sup>

<sup>1</sup>School of Broadcasting and Anchoring, Communication University of China, Beijing, China, <sup>2</sup>School of Communication, Soochow University, Suzhou, China, <sup>3</sup>School of Journalism and Communication, Tsinghua University, Beijing, China

As artificial intelligence technologies advance, humanlike digital avatars are increasingly used in health communication. Yet, whether—and how—avatars embodying real humans influence the persuasive impact of health messages remains largely unexplored. Drawing on communicative responsibility theory, this study proposes that, compared with real human communicators, audiences perceive digital avatars as less capable of assuming communicative responsibility, thereby weakening their persuasive effectiveness. Furthermore, this negative effect is expected to be stronger when the avatar represents a high-expertise identity (i.e., a doctor) rather than a low-expertise identity (i.e., a blogger). Using pancreatic cancer screening as the persuasion context, we conducted a 2 (communicator type: digital avatar vs. human) × 2 (source identity: high-expertise vs. low-expertise) between-subjects experiment with 211 Chinese participants. The findings revealed that (1) messages delivered by the digital avatar led to lower behavioral intentions than those delivered by its human originator; (2) perceived communicative responsibility negatively mediated this relationship; and (3) the negative effect of a digital avatar on perceived communicative responsibility was stronger when it represented a high-expertise identity rather than a low-expertise identity, which in turn led to lower behavioral intentions. These results advance theoretical understanding of avatar-mediated persuasion and offer practical guidance for the design of AI-based health communication.

### KEYWORDS

behavioral intention, communicative responsibility theory, digital avatar, health communication, persuasive effects

## Introduction

The growing need to educate the public about potential health risks and to encourage preventive or corrective health behaviors has led both medical experts (e.g., doctors and scientists) and non-experts (e.g., media influencers and broad science educators) to increasingly engage in health communication (Kaňková et al., 2024). Among the various channels available, audio-visual media plays a particularly important role, as it allows for parasocial interaction and vivid, engaging presentation of health information, enhancing attention, comprehension, and persuasion. However, not all potential health communicators—especially medical experts—are willing or able to participate in such media. Creating high-quality audio-visual content requires substantial time and effort, familiarity with camera presence, and a willingness

to undergo portrait exposure, all of which may pose significant barriers.

These individual-level barriers are compounded by structural pressures. As digital health communication becomes increasingly platformized across social media, mobile applications, and telehealth systems, the demand for infrastructures capable of sustaining user engagement and responsiveness at scale continues to intensify (Aljaroodi et al., 2019). Yet the human resources available for producing and disseminating health content are limited, particularly given the breadth of topics and the scale of target audiences. This tension between escalating communicative demand and finite human supply creates a gap that conventional digital tools alone are insufficient to close (Ahmed et al., 2025).

Digital avatars have emerged as a promising response to both sets of challenges. A digital avatar is a visually embodied agent capable of delivering information in an engaging and accessible manner (Aljaroodi et al., 2019). Unlike traditional human-centered approaches, digital avatars enable around-the-clock delivery of health guidance and extend communicative reach at substantially lower marginal costs. Among the various types of digital avatars, humanlike avatars are particularly noteworthy because they maintain a strong operational link to their human originators, distinguishing them from fully automated virtual agents. This connection operates on two levels. First, avatars are often designed to visually resemble the professionals or influencers they represent, supporting the creation of a recognizable and trustworthy personal brand. Second, and more critically, the content delivered by these avatars is primarily authored or rigorously vetted by humans, ensuring the accuracy, credibility, and reliability of the information. In China, platforms such as Douyin (TikTok), Kuaishou, and WeChat Video Channels increasingly feature “virtual doctors” and “virtual nurses” resembling real people to disseminate reliable health information.

Despite this growing adoption, the communicative effectiveness of humanlike avatars relative to human communicators remains an open and contested empirical question. Recently, scholarly interest in AI- and avatar-mediated persuasion has grown steadily, with prior studies examining comparative effectiveness across domains including marketing (Nalivaiké and Miliukaitė, 2024), environmental communication (Ni et al., 2023), education (Hajeer et al., 2024), as well as health communication (Xia et al., 2025). These investigations have produced mixed results, revealing both cases where AI avatars rival or even surpass human communicators and cases where human sources remain more persuasive. These inconsistent results suggest that the relative effectiveness of AI versus human communicators is not fixed and may be contingent on contextual and psychological factors. Researchers have also begun to probe the mechanisms behind these discrepancies, examining factors such as cognitive and emotional trust (Nalivaiké and Miliukaitė, 2024), risk and identity oneness perception (Ni et al., 2023), and benefit perception (Hajeer et al., 2024).

Yet, three interrelated gaps limit the applicability of existing work to the present context. The first gap concerns the distinctive nature of health-related persuasion as a domain. Unlike the consumer, environmental, and educational contexts examined in prior work, health decisions frequently carry serious personal consequences, require sustained attention, and involve intricate cognitive processing (Masic, 2022). This raises the possibility that the mechanisms governing avatar effectiveness in other domains may not transfer straightforwardly to health communication. The second concerns the type of avatar examined. Unlike generic AI agents that autonomously generate content,

humanlike avatars in current practice function as the embodied spokespersons of identifiable human originators, delivering content authored or rigorously vetted by those originators. Yet, direct comparisons between such avatars and their specific human originators remain scarce in health persuasion contexts. The third gap concerns the mechanisms driving avatar-mediated persuasion in health contexts specifically. Prior research comparing AI and human doctors has focused primarily on perceived warmth, empathy, relational depth, accuracy, and patients’ desire for uniqueness (Longoni et al., 2019; Seitz, 2024). While informative, these factors overlook the extent to which an avatar can embody the communicative responsibility of its human originator—that is, the degree to which it is perceived as accountable for delivering attentive, accurate, and context-sensitive guidance, as well as for the potential consequences of the messages it conveys. We argue that this construct is a critical factor in understanding and enhancing the effectiveness of avatar-mediated health communication.

Therefore, this study seeks to address this gap by investigating how humanlike digital avatars, when used as intermediaries in health communication, influence persuasive outcomes. Specifically, we aim to understand how the digital translation of human communicators through avatars—whether it diminishes, enhances, or maintains the effectiveness of health communication—affects persuasion, particularly in high-stakes contexts such as cancer screening. We propose perceived communicative responsibility as a key mechanism explaining the persuasive effect of humanlike avatars relative to their human originators. We hypothesize that, even when a digital avatar embodies the appearance and identity of a human spokesperson, it may fail to establish a clear responsibility link between the communicator and the accountability for the message. This, in turn, may reduce perceived communicative responsibility and subsequently lower behavioral compliance. Furthermore, we predict that this effect will be more pronounced when the avatar represents a high-expertise identity, such as a medical professional, compared to a lower-expertise source, such as a blogger.

By examining these dynamics, this study aims to illuminate how digital intermediaries, such as avatars, shape audience perceptions and behavior, while advancing understanding of potential underlying mechanisms. Practically, the study seeks to provide empirical evidence on whether digital avatars, despite their growing adoption and advantages in health communication, can effectively replace human communicators, and how any potential negative effects of avatar-mediated communication can be mitigated.

## Literature review

### The persuasive effect of digital avatar relative to human

The growing integration of artificial intelligence (AI) in communication has prompted significant scholarly interest in understanding the persuasive effects of AI-mediated agents. A central question in this field is whether AI communicators can achieve persuasive outcomes comparable to, or even surpass, human communicators under certain conditions (Huang and Wang, 2023). This question has been examined across various AI-driven communicative entities, including humanoid robots, voice assistants, and digital avatars. This study specifically focuses on humanlike digital avatars, defined as embodied AI

agents designed to visually replicate real humans using advanced technologies such as facial modeling, motion capture, and deepfake techniques (Biocca, 1997; Kim et al., 2023).

To explain the persuasive effect of AI communicator relative to human communicator, this study draws on MAIN (Modality—Agency—Interactivity—Navigability) model (Sundar, 2008). Originally developed to explain how technological affordances shape source credibility judgments, the model was later extended to broader persuasive outcomes, including attitudes and behavioral intentions (Sundar et al., 2019). The MAIN model proposes that technological features of an interface function as heuristic cues that influence users' cognitive processing, often independent of message content. In AI-mediated communication, even when identical messages are presented, cues signaling machine agency versus human agency may activate distinct heuristics, thereby shaping credibility assessments, affective responses, and ultimately persuasive outcomes (Sundar et al., 2015).

Consistent with the MAIN framework, prior research indicates that, under certain conditions, AI communicators can produce favorable persuasive outcomes compared to human communicators. For example, AI journalists have been found to mitigate audience bias toward distrusted media outlets (Cloudy et al., 2022). AI-generated voices, compared with human voices, can increase climate change risk perception by eliciting auditory fear, thereby promoting pro-environmental behavioral intentions (Ni et al., 2023). Similarly, digital avatars, such as virtual influencers, have been shown to outperform human influencers in promoting product purchases (Sakuma et al., 2023). However, evidence from other contexts points in the opposite direction. For instance, audiobooks narrated by human voices are perceived as more credible and persuasive than those narrated by AI (Rodero and Lucas, 2023), and news delivered by human presenters is often favored over AI-generated anchors (Kim et al., 2025). In addition, digital avatars such as synthetic spokespersons perform comparably to humans in training and knowledge-transfer settings (Lind, 2025). These inconsistent findings suggest that the relative persuasive effectiveness of AI communicators, including digital avatars, may vary depending on the context.

In the health decision-making context, evidence suggests persistent user reactance toward AI-provided medical services. Consumers generally prefer human providers (Longoni et al., 2019) and trust recommendations from human physicians more than those generated by computer algorithms (Li S. et al., 2025; Riedl et al., 2024). Resistance in medical decision-making contexts partly stems from the belief that machines lack the warmth, empathy, and relational depth required in patient-provider communication (Seitz, 2024). Longoni et al. (2019) further notes that individuals often assume AI diagnoses fail to account for personal uniqueness, reinforcing reluctance to accept AI-based healthcare advice. Taken together, these findings suggest that in health decisions—particularly those involving risk, safety, or potentially life-threatening consequences—individuals tend to be more receptive to persuasion and guidance from human providers than from AI systems. Consequently, when encountering AI-powered humanlike avatars, people may transfer their preexisting impressions of AI versus human medical communicators and exhibit lower behavioral compliance with avatar-delivered recommendations. Accordingly, this study proposes the following hypothesis:

*H1:* Compared with messages delivered by real humans, health persuasion messages based on digital avatars will elicit lower behavioral intention among audiences.

## Unpacking the persuasive gap between digital avatars and humans from a communicative responsibility perspective

Building on the preceding discussion, this study further explores the underlying mechanism behind the potential persuasive gap between digital avatars and real human communicators. Unlike entertainment or consumption-oriented contexts, health communication is a trust- and authority-dependent domain characterized by its seriousness. In such settings, audiences not only care about whether the communicator possesses relevant professional qualifications but also expect the communicator to assume responsibility for the accuracy and potential consequences of the information conveyed. Accordingly, this study introduces the concept of perceived communicative responsibility as a mediating mechanism to explain how different types of communicators may influence behavioral compliance with health recommendations.

Communicative responsibility refers to the obligation to ensure mutual understanding between interlocutors and to uphold relevant social expectations when participating in communication (Li et al., 2023; Youk and Park, 2023). Rooted in speech act theory and moral philosophy of language, communicative responsibility is premised on the idea that communication is not merely an exchange of information but a socially embedded practice in which participants are bound by normative expectations of sincerity, accuracy, and accountability (Habermas, 1985; Searle, 1969). These expectations become especially salient in institutional and mediated communication contexts—such as health, legal, or organizational settings—where asymmetries in expertise and power render audiences particularly dependent on communicators to act in good faith and to stand behind the content they convey (e.g., Grieshofer, 2022; Kuutila et al., 2024; Stivers and Timmermans, 2020). In AI-driven or algorithmically mediated environments, these expectations persist: users continue to expect AI advisors to uphold communicative responsibility of delivering accurate, reliable, and contextually appropriate information (Li et al., 2023; Sundar, 2020).

According to communicative responsibility theory, individuals in communicative contexts make systematic judgments regarding the degree to which each participant is accountable for contributing to the creation of shared understanding in an interaction (Aune et al., 2005). In this framework, audiences are likely to attribute communicative responsibility to a communicator only when they perceive that the communicator possesses both the normative awareness—an understanding of the social and moral expectations associated with the communicative role—and the behavioral autonomy—the capacity to intentionally act in a manner that upholds these responsibilities (Ayad and Plaks, 2025; Malle et al., 2014; Shoemaker, 2011). Together, these two dimensions constitute the foundational basis upon which audiences assess whether a communicator can legitimately be held responsible for the content and consequences of a communicative act.

Existing studies imply that digital avatars, which are representations of human originators, may be perceived as less capable of fulfilling communicative responsibilities along both dimensions. With respect to normative awareness, communicative responsibility requires that a communicator acts with genuine understanding and commitment to social and moral norms (Aune et al., 2005). Although avatars can simulate humanlike expressions and verbal behavior through anthropomorphic designs, their outputs are generated algorithmically or guided by human operators rather than through deliberative moral

reasoning (Misselhorn, 2020; Zhang et al., 2023). As a result, audiences attribute lower intentionality and moral comprehension to avatars, perceiving them as tools rather than normatively aware agents. Regarding behavioral autonomy, responsibility attribution depends on the perception that an entity can initiate actions independently. Avatars are typically constrained by algorithms or operator control, limiting perceived agency (Sundar and Kim, 2019), and their non-sentient nature prevents them from experiencing consequences or learning from outcomes, which further reduces the perception of autonomous action (Coeckelbergh, 2020). Moreover, accountability for any errors is diffused across underlying systems or controlling organizations rather than residing in the avatar itself (Gunkel, 2020), heightening users' concerns that the avatar cannot be held directly responsible.

In the context of health communication, such differences in perceived communicative responsibility become even more salient. In traditional healthcare systems, the chain of communicative responsibility is clear: medical professionals are held responsible for the accuracy and consequences of their professional recommendations (Chen et al., 2023). However, when the communicator is a digital avatar, this chain becomes blurred. The avatar functions primarily as a message conduit but may not be held responsible for the accuracy or consequences of the information it conveys. In other words, even if the avatar delivers incorrect or misleading messages, it is not typically considered the party to blame. Based on this reasoning, the following hypothesis is proposed:

*H2: Compared with real human communicators, users will perceive their digital avatars as having lower communicative responsibility.*

Further, perceived communicative responsibility functions as a key factor in facilitating successful persuasion (Aune et al., 2005). When audiences believe that a communicator is obligated to provide accurate information and can be held responsible for the consequences of their statements, they infer that the speaker is more careful, honest, and committed to the recipients' wellbeing (Mahr and Csibra, 2021). This attribution signals that the communicator possesses both the motivation and the moral duty to ensure message accuracy, which reduces perceived risks and increases confidence in the recommended actions (Sperber et al., 2010). As a result, audiences become more willing to accept and act on the message, because responsibility enhances trust, reinforces perceptions of expertise and reliability, and establishes a social norm whereby following the communicator's guidance feels justified and safe. Studies suggest that speakers are evaluated for how much they "own" their words: when a communicator is perceived as accountable for potentially false or misleading information, they are more likely to earn the trust of their audience (Haugh, 2013). Moreover, reputation built through consistent accountability in past interaction can significantly boost persuasion, as more reputable communicators are perceived as more reliable (Manzoor et al., 2024). Similarly, in health behaviors such as cancer screening, when audiences believe that a communicator is capable and willing to take responsibility for their messages, they are more likely to accept and act upon the recommendations.

Thus, the following hypotheses are proposed:

*H3: Perceived communicative responsibility of the speaker will positively predict individuals' intentions to adopt the recommended health behavior.*

*H4: Perceived communicative responsibility of the speaker will mediate the negative effect of communicator type (digital avatar vs. human) on individuals' intentions to adopt the recommended health behavior.*

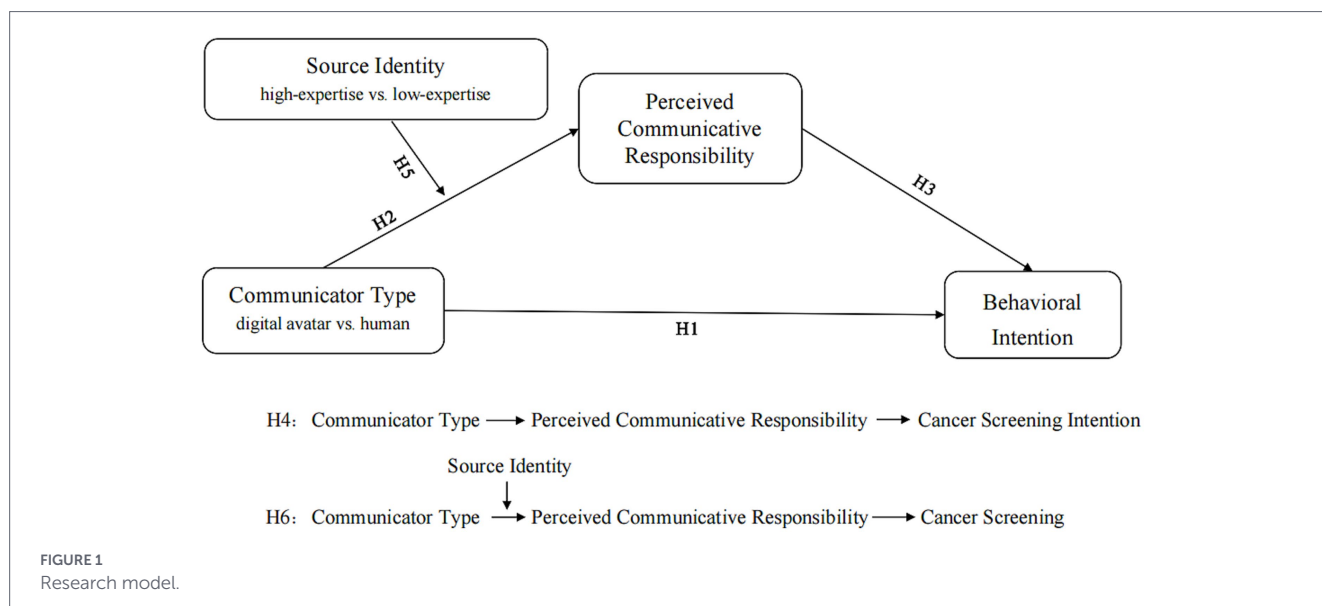
## The moderating role of source identity

Source identity plays a crucial role in shaping persuasive outcomes (Giffin, 1967). Research has long established that expert identities, as a critical dimension of source identity, enhance perceptions of message professionalism, credibility, and overall persuasiveness (Friedman and Friedman, 1979; Inglis and Mejia-Ramos, 2009). Building on this, we propose that source identity (high-expertise vs. low-expertise) moderates both the effect of communicator type (digital avatar vs. human) on perceived communicative responsibility and its mediating effect.

Expectancy violation theory (EVT) provides the theoretical logic for this moderation. EVT posits that individuals hold context-specific expectations about communicator behavior, and that violations of these expectations—whether positive or negative—produce evaluative consequences that shape subsequent attitudes and behaviors (Burgoon and Hale, 1988). The magnitude of an expectancy violation is determined by the gap between what audiences anticipate and what they actually perceive: the wider the gap, the more pronounced the violation and its downstream effects (Burgoon, 1993). Importantly, roles imbued with higher authority evoke stronger normative expectations of responsibility—physicians, for instance, are expected to prioritize patient welfare, just as police officers are trusted to prevent harm (Erlandsson et al., 2015; Haidt and Baron, 1996). A high-expertise identity thus conveys not only competence but also an elevated sense of role-based duty, setting a correspondingly high baseline of expected communicative responsibility.

According to the Computers Are Social Actors (CASA) paradigm, individuals apply social norms derived from human–human interaction to human–AI interaction (Reeves and Nass, 1996). This means that expertise cues common in human contexts—such as professional role labels, credential-signaling language, and role-consistent visual design—can function as powerful signals of competence and authority when present in AI agents (Liew and Tan, 2021). Consequently, users may extend the heightened responsibility expectations associated with expert roles to AI communicators whenever such cues are present (Rheu et al., 2024).

In health communication, medical experts such as physicians are typically viewed as high-credibility, high-trust, and high-responsibility sources (Falomir-Pichastor et al., 2006). Audiences thus approach such communicators with elevated expectations—anticipating not only accuracy and competence but also accountability, care, and role-consistent duty fulfillment. When this expert identity is embodied by a real human communicator, these expectations are generally congruent with the perceived agency and moral capacity of the source. However, when the same medical expert identity is represented by a digital avatar, a structural mismatch emerges: digital avatars are widely perceived as possessing reduced agency, diminished intentionality, and weaker accountability relative to human communicators (Gray and Wegner, 2012; Sundar and Kim, 2019). Because these perceived attributes fall short of the heightened normative standards associated with the medical expert role, the discrepancy between expected and perceived



communicative responsibility becomes pronounced. Under EVT, this substantial expectation–perception gap constitutes a negative expectancy violation, which tends to elicit unfavorable evaluations and diminish perceived communicative responsibility—producing a larger penalty for the avatar relative to the human communicator.

Conversely, when the communicator holds a low-expertise identity—such as a blogger—the baseline expectation of professional responsibility is substantially lower. In this context, audiences do not strongly anticipate high levels of accountability or role-based duty fulfillment from the source. Although a digital avatar may still be perceived as somewhat less responsible than an equivalent human communicator, the discrepancy between expectation and perception is comparatively smaller, because non-expert human communicators are themselves not strongly associated with professional responsibility norms. The magnitude of the negative expectancy violation is therefore attenuated, and the difference in perceived communicative responsibility between avatar and human communicators is expected to narrow. Based on this reasoning, we propose the following hypotheses:

*H5:* Source identity (high-expertise vs. low-expertise) moderates the effect of communicator type (digital avatar vs. human) on perceived communicative responsibility. Specifically, when the communicator holds a high-expertise identity (e.g., physician), the decrease in perceived communicative responsibility for digital avatars, relative to human communicators, will be stronger than when the source holds a low-expertise identity (e.g., blogger).

*H6:* Source identity (high-expertise vs. low-expertise) moderates the mediating role of perceived communicative responsibility between communicator type (digital avatar vs. human) and individuals' intentions to adopt the recommended health behavior. Specifically, when the communicator is identified as a medical expert, the negative mediating effect of perceived communicative responsibility will be stronger.

## Materials and methods

### Participants

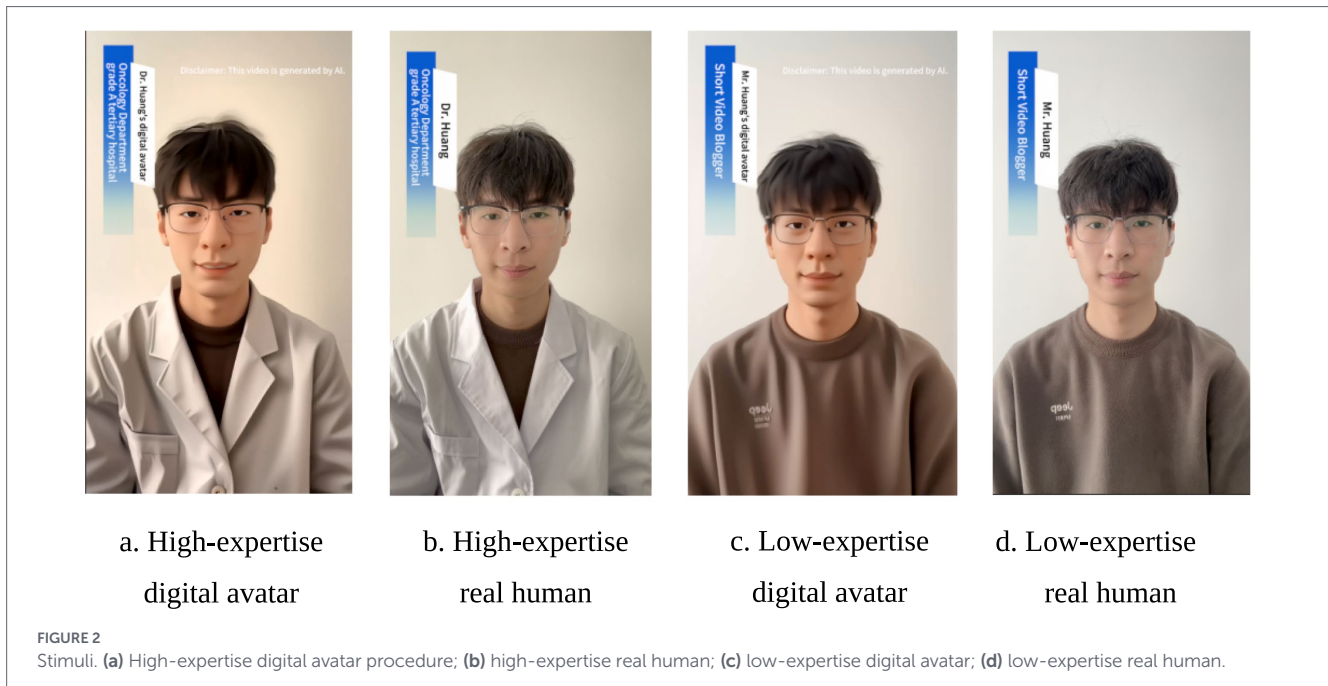
This study employed a 2 (communicator type: digital avatar vs. human) × 2 (source identity: high-expertise vs. low-expertise) between-subjects factorial design. The online experiment was conducted through the Chinese data collection platform Credamo.<sup>1</sup> A *priori* power analysis using G\*Power 3.1 (Faul et al., 2009) for a two-way between-subjects ANOVA indicated that a minimum of 128 participants were needed to detect a medium effect size ( $f = 0.25$ ) at  $\alpha = 0.05$  with a statistical power of 0.80. A total of 286 participants from various regions of China enrolled in the experiment. Because the experiment involved audiovisual stimuli, participants were instructed to ensure they were in a quiet environment suitable for clearly watching and listening to the materials (Figure 1).

Participants were excluded if they failed attention or manipulation checks, or if their completion times fell beyond two standard deviations from the mean ( $M = 575.71$  s,  $SD = 159.14$  s), resulting in a valid range of approximately 257–894 s. After applying these criteria, 211 valid questionnaires were retained for analysis. Among the valid respondents, 51.7% were female, with an average age of 32.14 years ( $SD = 8.10$ ). Regarding education, 2.8% had completed high school, 4.7% held an associate degree, 74.9% held a bachelor's degree, 17.1% held a master's degree, and 0.5% held a doctoral degree. Regarding monthly income, 0.9% reported earning 0–1,000 RMB, 9.5% earned 1,001–3,000 RMB, 39.0% earned 3,001–6,000 RMB, 34.1% earned 6,001–10,000 RMB, and 37.0% reported an income above 10,000 RMB.

### Stimuli

To enhance the ecological validity of the study, the experimental stimuli consisted of short health science videos focusing on cancer

<sup>1</sup> [www.credamo.com](http://www.credamo.com)



screening recommendations for pancreatic cancer. Short-form videos were selected as the presentation format because they have become one of the most influential channels for public health information—combining vivid expression, rich content, and accessible style, thereby lowering the cognitive threshold for understanding medical knowledge (Zhu et al., 2020). The topic of cancer screening was chosen for both public health and methodological considerations. According to the China National Cancer Center (Han et al., 2024), more than 4.5 million new cancer cases occur annually in China, far exceeding other countries, with a large proportion of patients diagnosed at middle or late stages. Among these, pancreatic cancer is particularly concerning due to its low detection rate and high fatality, making it a critical target for health communication. Meanwhile, compared with other cancers like breast or colorectal cancer, public awareness of pancreatic cancer remains remarkably low, which may further hinder early detection and underscores the need for effective health communication interventions. Methodologically, this low public familiarity reduces the likelihood of ceiling effects driven by entrenched prior attitudes, thereby enhancing experimental sensitivity to message-level variables such as source characteristics (Cacioppo et al., 1986).

Accordingly, this study designed persuasive messages centered on pancreatic cancer screening to reflect real-world health communication contexts while maintaining practical relevance. To ensure the scientific accuracy and credibility of the content, all video scripts were adapted from publicly available materials published by the Fudan University Cancer Hospital (2022).<sup>2</sup> The content included descriptions of pancreatic cancer risk, the benefits of early screening, and explicit encouragement to undergo screening (see Appendix for details).

For control and consistency, all videos were recorded by the same individual in the same setting. Figure 2 shows the screenshots of the stimuli. In the human condition, the original video of the person was used; in the avatar condition, a digital avatar version of the same person was generated using AI-powered avatar customization techniques. This

process involved customizing the avatar based on the real person's appearance, using advanced AI-driven algorithms to render a lifelike and accurate digital version, tailored specifically to reflect their unique features. Additionally, according to current practices, an AI disclaimer was added to the avatar videos, stating: "Disclaimer: This video was generated by AI." This reflects China's regulatory requirements for AI-generated content, as the Measures for Labeling of AI-Generated Synthetic Content (2025)<sup>3</sup> mandates explicit labeling of deep-synthesized content, ensuring the ecological validity of the experimental stimuli. The high- versus low-expertise identity was manipulated using multiple cues. Before viewing the stimulus, participants were informed that they would view an online video delivered either by an oncologist from a Grade A tertiary hospital (high-expertise condition) or by a short-video blogger (low-expertise condition). In the high-expertise condition, the communicator wore a white medical coat and was labeled on screen with his title as "Dr. Huang, Department of Oncology, Grade A Tertiary Hospital." In the low-expertise condition, the communicator wore casual clothing and was labeled "Mr. Huang, Short-Video Blogger." These manipulations align with the DASSCI model, which posits that expertise cues for artificial agents can be conveyed through dressing, occupational role, and social descriptors signaling domain specialization (Liew and Tan, 2021).

## Procedure

All participants provided informed consent and were then randomly assigned to one of four experimental conditions: (1) an oncologist represented by a digital avatar, (2) the oncologist himself, (3) a short video blogger represented by a digital avatar, or (4) the short video blogger himself. Participants were informed that they would watch a short video featuring a health message delivered by their assigned communicator. Importantly, regardless of whether the communicator appeared as an avatar or a human, all participants were

<sup>2</sup> From <https://mp.weixin.qq.com/s/115M6sRzqhYQ1BjG1tRWXw>.

<sup>3</sup> From [https://www.cac.gov.cn/2025-03/14/c\\_1743654684782215.htm](https://www.cac.gov.cn/2025-03/14/c_1743654684782215.htm).

explicitly told that the message content was written by the human originator.

After viewing the video, participants completed a questionnaire. This included an attention check item to confirm they watched the video carefully (“According to the video, approximately how many years does it take for pancreatic lesions to develop into pancreatic cancer? 11 years/12 years/21 years/22 years”), manipulation check items, perceived communicative responsibility of the communicator, alternative mechanisms, and behavioral intention to undergo pancreatic cancer screening. The demographic information was also collected.

## Measures

### Manipulation check

A manipulation check confirmed the perceived identity of the communicator in the video stimulus. Participants were asked, “Who was delivering the message to you?” with four response options: a real-human doctor, a doctor’s AI avatar, a real-human blogger, or a blogger’s AI avatar. Those who failed the manipulation check ( $N = 6$ ) were excluded from subsequent analysis.

Additionally, to assess the manipulation of participants’ perceptions regarding the expertise of source identity, participants were also asked, “To what extent do you think the speaker has expertise?” (1 = not at all, 7 = to a great extent;  $M = 5.34$ ,  $SD = 1.11$ ).

### Perceived communicative responsibility

Perceived communicative responsibility of the communicator was measured using a 7-point Likert scale (1 = not at all, 7 = to a great extent), adapted from Ahn et al. (2024). Three items were included: “To what extent do you think the communicator is responsible for the validity of this information?”, “To what extent do you think the communicator is responsible for the accuracy of this information?”, and “To what extent do you think the communicator will be responsible for the potential impact of this information?” (Cronbach’s  $\alpha = 0.87$ ,  $M = 5.46$ ,  $SD = 1.12$ ).

### Cancer screening intention

Participants’ intention to undergo cancer screening for pancreatic cancer was assessed using a 7-point Likert scale (1 = strongly disagree, 7 = strongly agree), adapted from Ajzen and Fishbein (1977). The construct was measured with three items: “At this moment, how strongly do you want to make an appointment for pancreatic cancer screening?”, “At this moment, to what extent are you willing to undergo a systematic pancreatic cancer screening (including blood tests, CT, MRI, and endoscopic ultrasonography)?”, and “Within the next week, how likely are you to search for pancreatic cancer screening options at nearby hospitals?” (Cronbach’s  $\alpha = 0.70$ ,  $M = 5.62$ ,  $SD = 0.82$ ).

### Alternative mechanisms

To control for potential confounding factors and isolate the effect of our proposed mediator, we also assessed several alternative mechanisms using a 7-point Likert scale (1 = strongly disagree, 7 = strongly agree). These included: perceived naturalness (“To what extent do you think this speaker is natural?”;  $M = 6.09$ ,  $SD = 0.92$ ), perceived objectivity (“To what extent do you think this speaker is objective?”;  $M = 6.13$ ,

$SD = 0.71$ ), perceived trustworthiness (“To what extent do you think this speaker is trustworthy?”;  $M = 6.13$ ,  $SD = 0.83$ ), perceived warmth (“To what extent do you think the speaker is warm?”;  $M = 4.94$ ,  $SD = 1.13$ ). Additionally, the manipulation check item—perceived expertise (“To what extent do you think the speaker has expertise?”;  $M = 5.34$ ,  $SD = 1.11$ )—was also examined as an alternative mechanism.

## Demographic variables and covariates

Demographic information included gender (1 = Male, 2 = Female), age, educational level (1 = Junior high school and below, 2 = High school, 3 = Associate college 4 = Bachelor, 5 = Master, 6 = Doctor), and income (1 = 0–1,000 RMB, 2 = 1,001–3,000 RMB, 3 = 3,001–6,000 RMB, 4 = 6,001–10,000 RMB, 5 = More than 10,000 RMB). In addition, health-related factors were also collected: subjective health (“How would you evaluate your own health condition? 1 = Very unhealthy, 7 = Very healthy”;  $M = 5.76$ ,  $SD = 0.81$ ), familiarity with pancreatic cancer (“How familiar are you with pancreatic cancer?”; 1 = Very unfamiliar, 7 = Very familiar,  $M = 3.08$ ,  $SD = 0.85$ ), and family cancer history (“Does your family have a history of cancer?”; No = 69.2%).

## Results

### Randomization check

To ensure the randomness of group assignment across experimental conditions, chi-square tests and one-way ANOVAs were conducted on all demographics and control variables. Results showed no significant differences among the four experimental groups in terms of gender ( $\chi^2(3) = 6.55$ ,  $p = 0.088$ ), age ( $F(3, 207) = 0.55$ ,  $p = 0.649$ ), education ( $F(3, 207) = 1.43$ ,  $p = 0.234$ ), or income ( $F(3, 207) = 0.60$ ,  $p = 0.618$ ).

Similarly, there were no significant differences in health-related factors, including subjective health ( $F(3, 207) = 1.08$ ,  $p = 0.360$ ), familiarity with pancreatic cancer ( $F(3, 207) = 1.13$ ,  $p = 0.338$ ), or family cancer history ( $\chi^2(3) = 0.45$ ,  $p = 0.930$ ). These results indicate that the random assignment was successful.

### Manipulation check

All participants, except for eight, successfully recognized the assigned communicator type and identity; these eight participants were excluded from the formal analyses. Additionally, to assess whether the experimental stimulus effectively manipulated participants’ expertise perception of the source identity, a one-way ANOVA was conducted. The results showed significant differences in expertise perception ( $F(1, 209) = 24.46$ ,  $p < 0.001$ ) between the high-expertise ( $M = 6.25$ ,  $SD = 0.78$ ) and the low-expertise ( $M = 5.58$ ,  $SD = 1.17$ ) group, confirming that the manipulation check was successful.

### Effect of communicator type

H1 predicted a main effect of communicator type (digital avatar vs. human). To test this hypothesis, we used an ANOVA to analyze the data. The results revealed that compared with human communicators ( $M = 5.80$ ,  $SD = 0.67$ ), digital avatars ( $M = 5.44$ ,  $SD = 0.93$ ) elicited weaker intentions for pancreatic cancer screening,  $F(1, 209) = 10.28$ ,  $p = 0.002$ ,  $\eta_p^2 = 0.047$ . Thus, H1 was supported.

## Mediating effect of perceived communicative responsibility

H2–H4 predicted the mediating role of perceived communicative responsibility. To test this hypothesis, we used PROCESS Model 4 to analyze the data. Results indicated that, compared with human communicators, digital avatars evoked significantly lower perceived communicative responsibility ( $\beta = -0.37, p < 0.001$ ). In turn, perceived communicative responsibility positively predicted cancer screening intention ( $\beta = 0.32, p < 0.001$ ). Moreover, perceived communicative responsibility significantly and fully mediated the relationship between communicator type (digital avatar vs. human) and participants' cancer screening intention (indirect effect:  $\beta = -0.12, CI_{95\%} = [-0.204, -0.051]$ ; direct effect:  $\beta = -0.10, p = 0.16, CI_{95\%} = [-0.234, 0.039]$ ). Therefore, H2, H3, and H4 were supported.

## Interaction effect between communicator type and source identity

H5 predicted an interaction effect between communicator type (digital avatar vs. human) and source identity (high-expertise vs. low-expertise) on perceived communicative responsibility. To test this hypothesis, we used PROCESS Model 7 to analyze the data. The results demonstrated a significant interaction effect ( $\beta = -0.17, p = 0.007$ ). A simple slope analysis (see Figure 3) further showed that the reduction in perceived communicative responsibility caused by using a digital avatar (vs. human) was stronger in the high-expertise identity condition ( $\beta = -0.54, p < 0.001$ ) than that in the low-expertise identity condition ( $\beta = -0.19, p = 0.034$ ). Thus, H5 was supported.

H6 predicted a moderated mediation effect of communicator type (digital avatar vs. human) and source identity (high-expertise vs. low-expertise) on the mediation pathway through perceived communicative responsibility. The analysis using PROCESS Model 7 revealed that source identity moderated the indirect effect of communicator type on cancer screening intention via perceived communicative

responsibility ( $\beta = -0.11, CI_{95\%} = [-0.223, -0.027]$ ). Specifically, the mediation effect was stronger ( $\beta = -0.17, CI_{95\%} = [-0.297, -0.079]$ ) in the high-expertise identity condition than that in the low-expertise identity condition ( $\beta = -0.06, CI_{95\%} = [-0.143, -0.006]$ ). Therefore, H6 was supported (Figure 4).

## Alternative mechanisms

When entering all alternative mechanisms as parallel mediators with perceived communicative responsibility, results showed that perceived warmth ( $\beta = -0.04, CI_{95\%} = [-0.090, -0.010]$ ) also mediated the relationship between communicator type (digital avatar vs. human) and cancer screening intention. However, no significant mediation effects were found for perceived objectivity ( $\beta = -0.02, CI_{95\%} = [-0.054, 0.000]$ ), perceived naturalness ( $\beta = -0.03, CI_{95\%} = [-0.092, 0.031]$ ), perceived trustworthiness ( $\beta = -0.03, CI_{95\%} = [-0.070, 0.003]$ ) and perceived expertise ( $\beta = -0.004, CI_{95\%} = [-0.032, 0.009]$ ). More importantly, perceived communicative responsibility remained a significant mediator and showed the strongest effect among all mediators ( $\beta = -0.06, CI_{95\%} = [-0.133, -0.010]$ ). Given that perceived warmth emerged as a significant parallel mediator, we conducted a follow-up mediation analysis with it as the sole mediator. We found that its mediating effect was partial (indirect effect:  $\beta = -0.08, CI_{95\%} = [-0.147, -0.031]$ ; direct effect:  $\beta = -0.14, p = 0.040, CI_{95\%} = [-0.268, -0.001]$ ). However, as reported earlier, perceived communicative responsibility fully mediated this effect. These findings highlight the added and crucial explanatory power of our proposed mediator.

## Discussion

This study examined how communicator type and source identity jointly influence audiences' behavioral intentions in the context of health persuasion. Using a 2 (communicator type: digital avatar vs.

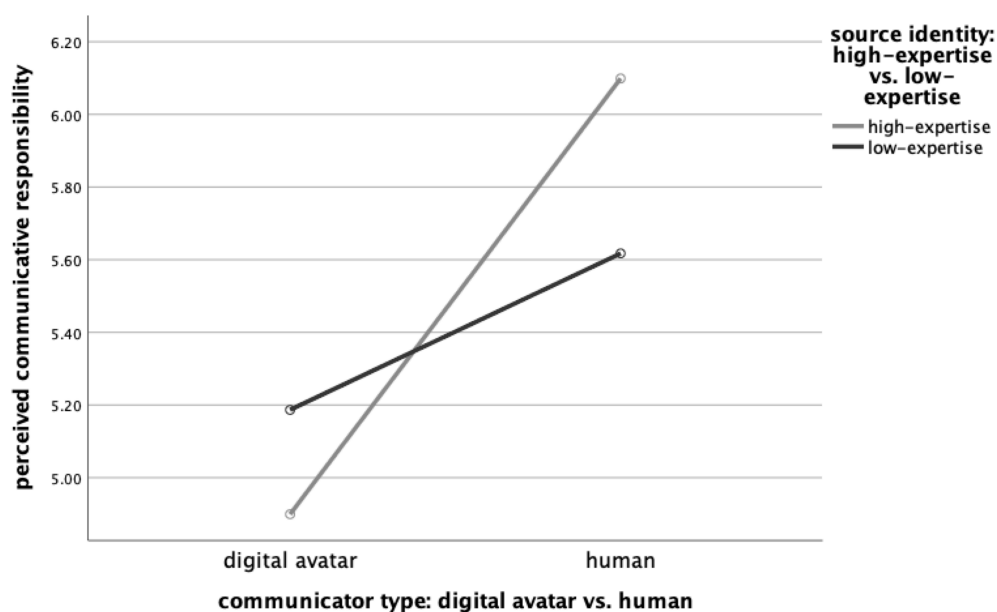
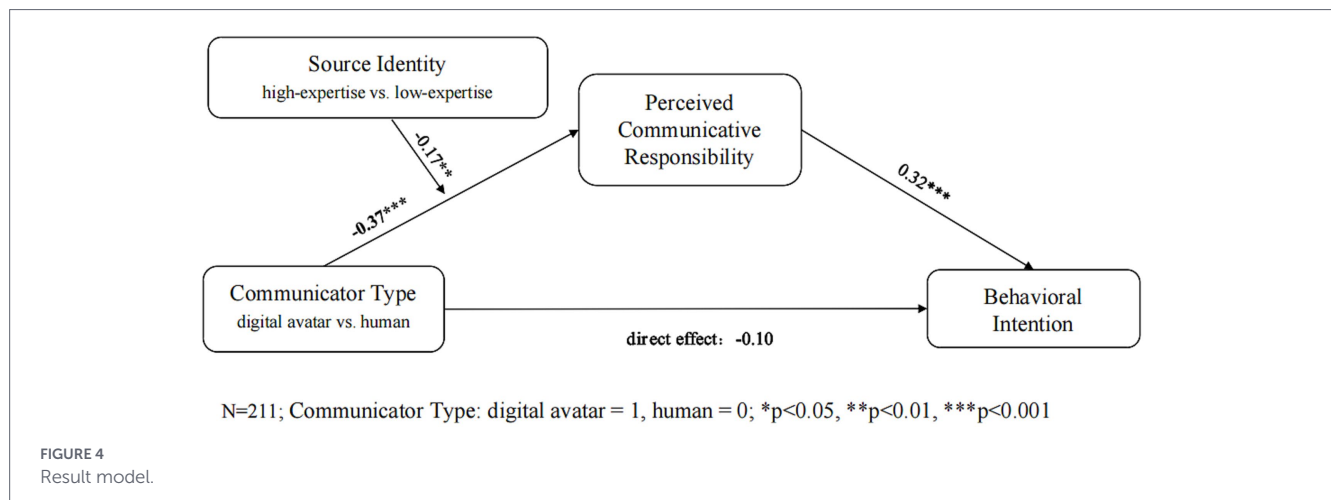


FIGURE 3  
Interaction effect of communicator type and source identity on perceived communicative responsibility.



human)  $\times$  2 (source identity: high-expertise vs. low-expertise) between-subjects experimental design, this study explored the effects of these variables on individuals' intention to comply with pancreatic cancer screening recommendations. The results showed that (1) participants exposed to digital avatar communicators reported significantly lower behavioral intention than those exposed to human communicators; (2) perceived communicative responsibility mediated the negative effect of communicator type on behavioral intention; and (3) source identity moderated this mediation effect. Specifically, the negative effect of digital avatars on perceived communicative responsibility was stronger in the high-expertise identity condition (i.e., a doctor) than in the low-expertise identity condition (i.e., a blogger), leading to lower behavioral intention.

## Theoretical implications

The main effect extends emerging research on the persuasive effectiveness of AI into the domain of health communication. Our findings indicate that, as a specific application of AI technologies, humanlike digital avatars were less persuasive than real humans in encouraging individuals to undergo pancreatic cancer screening. This result aligns with prior studies showing that individuals are less motivated to seek health recommendations from AI systems compared to human providers (Longoni et al., 2019; Li et al., 2025; Riedl et al., 2024), as well as research indicating that humanlike digital avatars become less effective in knowledge transfer once their artificial nature is detected (e.g., Lind, 2025). Although other studies have demonstrated that more humanlike interactions with avatars can enhance users' acceptance of medical services through heightened perceived anthropomorphism (Sestino and D'Angelo, 2023), the persuasive gap between humanlike avatars and real-human communicators may nevertheless persist in high-stakes medical decision-making contexts, as reflected in our findings. Nevertheless, our findings contrast with studies demonstrating that digital avatars can be as persuasive as humans (Zanbaka et al., 2006, 2007; Ogawa et al., 2019) and even outperform them in certain marketing and consumer-decision contexts (Sakuma et al., 2023). These discrepancies suggest that the persuasive advantage of AI communicators is not uniform across domains but is highly contingent on the situational demands and psychological expectations embedded in the persuasive setting.

Moreover, we demonstrate that perceived communicative responsibility plays a critical role in explaining why health recommendations

delivered by a humanlike digital avatar (vs. a real human) lead to lower behavioral intentions. Our findings show that, compared with human speakers, audiences perceive humanlike digital avatars as having a weaker obligation to ensure message accuracy and a reduced capacity to be held accountable for potential consequences. Notably, even though our experiment explicitly informed participants that the avatar merely conveyed a message created by its human originator, this clarification did not eliminate the avatar's persuasive disadvantage. This indicates that who delivers the message—independently of message content or authorship—matters in health communication.

Prior explanations for the limited persuasive effectiveness of AI-powered agents have primarily emphasized technological factors (e.g., perceptual unnaturalness), emotional cues (e.g., lack of warmth), relational factors (e.g., low trust), and cognitive judgments (e.g., perceived lack of expertise). However, even after controlling for these variables, perceived communicative responsibility remained a robust predictor and emerged as the strongest mediator in our analysis. This pattern points to an underlying lay assumption: AI-powered digital avatars are not regarded as socially accountable communicators, even when they faithfully speak on behalf of a real human. Ultimately, it is the human originator—rather than the avatar acting as their mouthpiece—who is intuitively seen as the one who can be held accountable for what is said. Thus, when an avatar replaces the human communicator as the visible speaker, audiences may struggle to establish a clear linkage between who is delivering the message and who can be held responsible for it.

In light of this, we propose that avatar-based health communication gives rise to a previously overlooked phenomenon, which we term “detachment of responsibility.” This reflects a breakdown in the perceived transmission of accountability from the human communicator to their embodied avatar, ultimately weakening persuasive impact. While the concept builds on prior research, it departs in important ways. Source credibility emphasizes that persuasion depends on evaluations of the communicator (Hovland and Weiss, 1951), but credibility pertains to expertise and trustworthiness, whereas responsibility concerns accountability. A communicator may be seen as competent or trustworthy without being perceived as the party who can be held accountable for the message. Likewise, agency attribution in human-AI interaction research (Ciardo et al., 2020) typically examines whether an entity is perceived as an intentional actor. However, we focus on responsibility attribution that emphasizes a distinct normative judgment: who can legitimately be held

accountable for the consequences of communication. That is, an avatar may be perceived as acting or speaking, yet still not considered responsible for the outcomes of its messages. Detachment of responsibility also relates to, but is distinct from, moral accountability. While moral accountability focuses on ethical or value-based judgments (Shoemaker, 2011), our concept emphasizes communicative responsibility, which concerns a broader perception of who is answerable for delivering and representing the message itself, independent of its moral or ethical content. Taken together, we define detachment of responsibility as the failure of perceived accountability to transfer from a human originator to their avatar. It reflects a structural misalignment between the visible speaker and the locus of responsibility. When audiences cannot clearly anchor accountability in the apparent communicator, the persuasive impact of the message is diminished.

Furthermore, this study identifies source identity (i.e., high-expertise vs. low-expertise) as a boundary condition shaping the effect of communicator type (digital avatar vs. human) on behavioral intention via perceived communicative responsibility. Prior research on persuasive AI has primarily emphasized a simple “real vs. virtual” distinction (e.g., Do et al., 2022). However, as AI design advances and avatars are increasingly deployed in contextualized domains, they now embody specific social identities—such as “medical expert”—raising the critical question of whether these identities can carry the same persuasive weight as their human counterparts.

Moving beyond this basic dichotomy, our study examines how communicator type interacts with social role to influence perceived communicative responsibility and persuasive effectiveness. Our findings reveal that the phenomenon of responsibility detachment is magnified—not mitigated—when the communicator is expected to be a medical expert, compared with when the communicator is a non-expert (e.g., blogger). This pattern emerges because audiences typically hold elevated expectations that medical experts will act responsibly and be accountable for the accuracy and implications of the health information they convey. Yet once this expert identity is instantiated through an avatar, these heightened expectations diminish sharply, effectively collapsing toward the level associated with a non-expert. This suggests that the responsibility expected of a medical expert is far more difficult to transfer to an avatar than the expectations associated with a non-expert identity, resulting in a disproportionately large accountability gap for expert avatars. This suggests an accountability gap that widens with the communicator’s presumed expertise—the higher the normative expectations, the more severe the responsibility detachment when a human is replaced by an AI surrogate.

## Practical implications

The findings of this research carry several practical implications for the design and deployment of AI-powered digital avatars in health communication. First, the results caution against relying on human-like digital avatars as frontline agents for delivering high-stakes medical recommendations, such as in the context of pancreatic cancer screening. Despite ongoing technological advances in AI embodiment and fluency, avatars remain less persuasive than real-human communicators in encouraging consequential health behaviors such as cancer screening. Health organizations and public agencies should therefore exercise additional scrutiny before substituting human health professionals with digital avatars in contexts that require strong behavioral compliance.

Second, communicative responsibility should be explicitly designed into avatar-based health communication systems. Because audiences perceive avatars as having weaker obligations to ensure message accuracy and as less accountable for potential consequences, system designers may need to embed clearer responsibility cues—such as transparent attribution of the human clinician behind the message, explicit statements of accountability, or interface elements that visually reinforce the human–avatar linkage. Merely stating that an avatar is transmitting a human-created message, as shown in our experiment, is insufficient. More robust designs may be required to ensure that responsibility is perceptibly anchored to a real human agent.

Third, the heightened vulnerability of “expert avatars” suggests that organizations should be especially cautious when assigning professional roles to AI embodiments. Our results indicate that avatars representing medical experts suffer from an even larger responsibility and persuasion deficit than avatars portraying non-experts. Healthcare providers should therefore refrain from positioning avatars as physicians or specialists unless their design can meaningfully meet the normative expectations associated with medical expertise. Otherwise, attempts to use expert-branded avatars may unintentionally undermine audiences perceived communicative responsibility and behavioral compliance.

Finally, these findings suggest opportunities for targeted training and hybrid models. Rather than replacing clinicians, digital avatars may be better suited to supporting roles—such as delivering procedural information, answering low-stakes queries, or assisting with follow-up reminders—while leaving recommendation-heavy or ethically laden communication to human providers. Hybrid designs that integrate avatar support with human follow-up may preserve efficiency gains while minimizing responsibility detachment.

## Limitations and future directions

Several limitations should be acknowledged. First, this study focused specifically on pancreatic cancer screening as the persuasive context due to its public health significance. However, it is important to note that the results may not directly generalize to all high-stakes medical decision-making contexts. Future studies are encouraged to test the observed effects in other contexts involving different disease profiles, treatment pathways, or varying levels of scientific uncertainty. Second, the AI disclaimer included in the digital avatar condition, while necessary for ecological validity under China’s regulatory requirements, means that the observed effects may reflect the combined influence of avatar embodiment and AI disclosure. Future research could independently manipulate AI labeling to disentangle these two factors. Third, as an initial exploration, our stimuli featured only male communicator identities; it remains unclear whether these findings extend to female avatars or whether communicator gender interacts with perceived communicative responsibility and persuasion. Future research should also consider intersectional factors, such as age, ethnicity, and cultural context, that may influence perceived communicative responsibility and persuasive outcomes. Moreover, although screening intention is a strong precursor to actual screening behavior, intentions do not always translate into action. Future work could incorporate behavioral outcomes (e.g., completing a screening request). Finally, this study assessed only immediate responses; future research should employ longitudinal designs or randomized controlled trials (Miri et al., 2024), to examine whether these effects persist over time and generalize to other health communication contexts, such as interventions addressing health misinformation.

## Conclusion

This study advances understanding of AI-mediated health communication by demonstrating that humanlike digital avatars are less effective than human communicators in promoting cancer screening behaviors, and by identifying perceived communicative responsibility as a key psychological mechanism underlying this persuasive gap. Even when avatars faithfully convey messages created by real humans, audiences struggle to attribute responsibility to them, leading to diminished accountability perceptions and reduced behavioral intentions. We further show that this “responsibility detachment” is amplified when avatars represent medical experts than short video bloggers, highlighting a widening accountability gap as normative expectations rise. These findings underscore the importance of reconsidering how responsibility, expertise, and social roles are communicated and perceived in emerging human-AI interactions.

## Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

## Ethics statement

The studies involving humans were approved by The Academic Committee of the School of Broadcasting and Anchoring, CUC. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

## Author contributions

YW: Writing – original draft, Writing – review & editing, Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Visualization. FW: Writing – review

& editing, Conceptualization, Investigation, Methodology. BN: Writing – review & editing, Conceptualization, Data curation, Formal analysis, Methodology, Validation.

## Funding

The author(s) declared that financial support was received for this work and/or its publication. This research was funded by the 2024 Central University Research Project of Communication University of China, “Effects of Intelligent Voice Technologies on Health Information Dissemination” (Grant numbers: CUC24BS32).

## Conflict of interest

The author(s) declared that this work was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Generative AI statement

The author(s) declared that Generative AI was not used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

## Publisher’s note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## References

- Ahmed, M. M., Okesanya, O. J., Olaleke, N. O., Adigun, O. A., Adebayo, U. O., Oso, T. A., et al. (2025). Integrating digital health innovations to achieve universal health coverage: promoting health outcomes and quality through global public health equity. *Healthcare* 13:1060. doi: 10.3390/healthcare13091060
- Ahn, J., Kim, J., and Sung, Y. (2024). The role of perceived freewill in crises of human-AI interaction: the mediating role of ethical responsibility of AI. *Int. J. Advert.* 43, 847–873. doi: 10.1080/02650487.2023.2299563
- Ajzen, I., and Fishbein, M. (1977). Attitude-behavior relations: a theoretical analysis and review of empirical research. *Psychol. Bull.* 84, 888–918. doi: 10.1037/0033-2909.84.5.888
- Aljaroodi, H. M., Adam, M. T., Chiong, R., and Teubner, T. (2019). Avatars and embodied agents in experimental information systems research: a systematic review and conceptual framework. *Australas. J. Inf. Syst.* 23, 1–37. doi: 10.3127/ajis.v23i0.1841
- Aune, R. K., Levine, T. R., Park, H. S., Asada, K. J. K., and Banas, J. A. (2005). Tests of a theory of communicative responsibility. *J. Lang. Soc. Psychol.* 24, 358–381. doi: 10.1177/0261927x05281425
- Ayad, R., and Plaks, J. E. (2025). Attributions of intent and moral responsibility to AI agents. *Comput. Hum. Behav. Artif. Hum.* 3:100107. doi: 10.1016/j.chbah.2024.100107
- Biocca, F. (1997). The cyborg’s dilemma: progressive embodiment in virtual environments. *J. Comput.-Mediat. Commun.* 3:JCMC324.
- Burgoon, J. K. (1993). Interpersonal expectations, expectancy violations, and emotional communication. *J. Lang. Soc. Psychol.* 12, 30–48. doi: 10.1177/0261927x93121003
- Burgoon, J. K., and Hale, J. L. (1988). Nonverbal expectancy violations: model elaboration and application to immediacy behaviors. *Commun. Monogr.* 55, 58–79. doi: 10.1080/03637758809376158

- Cacioppo, J. T., Petty, R. E., Kao, C. F., and Rodriguez, R. (1986). Central and peripheral routes to persuasion: an individual difference perspective. *J. Pers. Soc. Psychol.* 51:1032. doi: 10.1037/0022-3514.51.5.1032
- Chen, A., Wang, C., and Zhang, X. (2023). Reflection on the equitable attribution of responsibility for artificial intelligence-assisted diagnosis and treatment decisions. *Intell. Med.* 3, 139–143. doi: 10.1016/j.imed.2022.04.002
- Ciaro, F., Beyer, F., De Tommaso, D., and Wykowska, A. (2020). Attribution of intentional agency towards robots reduces one's own sense of agency. *Cognition* 194:104109. doi: 10.1016/j.cognition.2019.104109
- Cloudy, J. T., Banks, J. T., and Bowman, N. D. (2022). AI journalists and reduction of perceived hostile media bias: replication and extension considering news organization cues. *Technol. Mind Behav.* 3:364. doi: 10.1037/tmb0000083
- Coeckelbergh, M. (2020). Artificial intelligence, responsibility attribution, and a relational justification of explainability. *Sci. Eng. Ethics* 26, 2051–2068. doi: 10.1007/s11948-019-00146-8
- Do, T. D., McMahan, R. P., and Wisniewski, P. J. (2022). “A new uncanny valley? The effects of speech fidelity and human listener gender on social perceptions of a virtual-human speaker,” in Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems, (1–11).
- Erlandsson, A., Björklund, F., and Bäckström, M. (2015). Emotional reactions, perceived impact and perceived responsibility mediate the identifiable victim effect, proportion dominance effect and in-group effect respectively. *Organ. Behav. Hum. Decis. Process.* 127, 1–14. doi: 10.1016/j.obhdp.2014.11.003
- Falimir-Pichastor, J. M., Mugny, G., and Invernizzi, F. (2006). Smokers' (dis) satisfaction, persuasive constraint, and influence of expert and non-expert sources. *Swiss J. Psychol.* 65, 7–14. doi: 10.1024/1421-0185.65.1.7
- Faul, F., Erdfelder, E., Buchner, A., and Lang, A. G. (2009). Statistical power analyses using G\*Power 3.1: Tests for correlation and regression analyses. *Behav. Res. Methods.* 41, 1149–1160.
- Friedman, H. H., and Friedman, L. (1979). Endorser effectiveness by product type. *J. Advert. Res.* 19, 63–71. doi: 10.1080/00218499.1979.12518848
- Giffin, K. (1967). The contribution of studies of source credibility to a theory of interpersonal trust in the communication process. *Psychol. Bull.* 68:104. doi: 10.1037/h0024833
- Gray, K., and Wegner, D. M. (2012). Feeling robots and human zombies: mind perception and the uncanny valley. *Cognition* 125, 125–130. doi: 10.1016/j.cognition.2012.06.007
- Grieshofer, T. (2022). Lay advisers in family law settings: the role and quality of advice provided on social media. *Soc. Legal Stud.* 31, 941–960. doi: 10.1177/09646639221090132
- Gunkel, D. J. (2020). Mind the gap: responsible robotics and the problem of responsibility. *Ethics Inf. Technol.* 22, 307–320. doi: 10.1007/s10676-017-9428-2
- Habermas, J. (1985). *The Theory of Communicative Action: Volume 1: Reason and the Rationalization of Society*, vol. 1. Boston: Beacon Press.
- Haidt, J., and Baron, J. (1996). Social roles and the moral judgement of acts and omissions. *Eur. J. Soc. Psychol.* 26, 201–218. doi: 10.1002/(sici)1099-0992(199603)26:2<201::j-hajec>3.0.co;2-j
- Hajeer, A., Papp-Váry, Á., and Pólya, É. (2024). AI tutors vs. human instructors: perceptions of higher education students in Hungary and Spain. *Educat. Rev. Electron. Technol. Educ.* 89, 105–120. doi: 10.21556/edutec.2024.89.3523
- Han, B., Zheng, R., Zeng, H., Wang, S., Sun, K., Chen, R., et al. (2024). Cancer incidence and mortality in China, 2022. *J. Natl. Cancer Center* 4, 47–53. doi: 10.1016/j.jncc.2024.01.006
- Haugh, M. (2013). Speaker meaning and accountability in interaction. *J. Pragmat.* 48, 41–56. doi: 10.1016/j.pragma.2012.11.009
- Hovland, C. I., and Weiss, W. (1951). The influence of source credibility on communication effectiveness. *Public Opin. Q.* 15, 635–650. doi: 10.1086/266350
- Huang, G., and Wang, S. (2023). Is artificial intelligence more persuasive than humans? A meta-analysis. *J. Commun.* 73, 552–562. doi: 10.1093/joc/fjad024
- Inglis, M., and Mejia-Ramos, J. P. (2009). The effect of authority on the persuasiveness of mathematical arguments. *Cogn. Instr.* 27, 25–50. doi: 10.1080/07370000802584513
- Kaňková, J., Binder, A., and Matthes, J. (2024). Helpful or harmful? Navigating the impact of social media influencers' health advice: insights from health expert content creators. *BMC Public Health* 24:3511. doi: 10.1186/s12889-024-21095-3
- Kim, D. Y., Lee, H. K., and Chung, K. (2023). Avatar-mediated experience in the meta-verse: the impact of avatar realism on user-avatar relationship. *J. Retail. Consum. Serv.* 73:103382.
- Kim, J., Xu, K., and Merrill, K. Jr. (2025). Man vs. machine: human responses to an AI newscaster and the role of social presence. *Soc. Sci. J.* 62, 1–13. doi: 10.1016/j.jretconser.2023.103382
- Kuutila, M., Kili, C., Kupiainen, R., Huusko, E., Li, J., Hosio, S., et al. (2024). Revealing complexities when adult readers engage in the credibility evaluation of social media posts. *Comput. Hum. Behav.* 151:108017. doi: 10.1016/j.chb.2023.108017
- Li, S., Chen, M., Liu, P. L., and Xu, J. (2025). Following medical advice of an AI or a human doctor? Experimental evidence based on clinician-patient communication pathway model. *Health Commun.* 40, 1810–1822. doi: 10.1080/10410236.2024.2423114
- Li, H., Lei, Y., Zhou, Q., and Yuan, H. (2023). Can you sense without being human? Comparing virtual and human influencers endorsement effectiveness. *J. Retail. Consum. Serv.* 75:103456. doi: 10.1016/j.jretconser.2023.103456
- Liew, T. W., and Tan, S. M. (2021). Social cues and implications for designing expert and competent artificial agents: a systematic review. *Telemat. Inform.* 65:101721. doi: 10.1016/j.tele.2021.101721
- Lind, S. J. (2025). Can AI-powered avatars replace human trainers? An empirical test of synthetic humanlike spokesperson applications. *J. Workplace Learn.* 37, 19–40. doi: 10.1108/JWL-04-2024-0075
- Longoni, C., Bonezzi, A., and Morewedge, C. K. (2019). Resistance to medical artificial intelligence. *J. Consum. Res.* 46, 629–650. doi: 10.1093/jcr/ucz013
- Mahr, J. B., and Csibra, G. (2021). The effect of source claims on statement believability and speaker accountability. *Mem. Cogn.* 49, 1505–1525. doi: 10.3758/s13421-021-01186-x
- Malle, B. F., Guglielmo, S., and Monroe, A. E. (2014). A theory of blame. *Psychol. Inq.* 25, 147–186. doi: 10.1080/1047840x.2014.877340
- Manzoor, E., Chen, G. H., Lee, D., and Smith, M. D. (2024). Influence via ethos: on the persuasive power of reputation in deliberation online. *Manag. Sci.* 70, 1613–1634. doi: 10.1287/mnsc.2023.4762
- Masic, I. (2022). Medical decision making—an overview. *Acta Inform. Med.* 30:230. doi: 10.5455/aim.2022.30.230-235
- Miri, A., Karimi-Shahanjari, A., Afshari, M., Bashirian, S., and Tapak, L. (2024). Understanding the features and effectiveness of randomized controlled trials in reducing COVID-19 misinformation: a systematic review. *Health Educ. Res.* 39, 495–506. doi: 10.1093/her/cyae036
- Misselhorn, C. (2020). Artificial systems with moral capacities? A research design and its implementation in a geriatric care system. *Artif. Intell.* 278:103179. doi: 10.1016/j.artint.2019.103179
- Nalivaiké, J., and Miliukaitė, G. (2024). “Influence of AI-generated avatars on consumer trust in the brand,” in New Trends in Contemporary Economics, Business and Management. Selected Proceedings of the 14th International Scientific Conference “Business and Management 2024”, (601–610).
- Ni, B., Wu, F., and Huang, Q. (2023). When artificial intelligence voices human concerns: the paradoxical effects of AI voice on climate risk perception and pro-environmental behavioral intention. *Int. J. Environ. Res. Public Health* 20:3772. doi: 10.3390/ijerph20043772
- Ogawa, N., Narumi, T., and Hirose, M. (2019). “Virtual hand realism affects object size perception in body-based scaling,” in 2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR), (Osaka, Japan: 2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)IEEE), 519–528.
- Reeves, B., and Nass, C. (1996). The media equation: how people treat computers, television, and new media like real people. *Cambridge, UK* 10, 19–36.
- Rheu, M., Dai, Y., Meng, J., and Peng, W. (2024). When a chatbot disappoints you: expectancy violation in human-chatbot interaction in a social support context. *Commun. Res.* 51, 782–814. doi: 10.1177/00936502231221669
- Riedl, R., Hogeterp, S. A., and Reuter, M. (2024). Do patients prefer a human doctor, artificial intelligence, or a blend, and is this preference dependent on medical discipline? Empirical evidence and implications for medical practice. *Front. Psychol.* 15:1422177. doi: 10.3389/fpsyg.2024.1422177
- Rodero, E., and Lucas, I. (2023). Synthetic versus human voices in audiobooks: the human emotional intimacy effect. *New Media Soc.* 25, 1746–1764. doi: 10.1177/14614448211024142
- Sakuma, H., Hori, A., Murashita, M., Kondo, C., and Hijikata, Y. (2023). YouTubers vs. VTubers: persuasiveness of human and virtual presenters in promotional videos. *Front. Comput. Sci.* 5:1043342. doi: 10.3389/fcomp.2023.1043342
- Searle, J. R. (1969). *Speech Acts: An Essay in the Philosophy of Language*. New York, United States: Cambridge University Press.
- Seitz, L. (2024). Artificial empathy in healthcare chatbots: does it feel authentic? *Comput. Hum. Behav. Artif. Hum.* 2:100067. doi: 10.1016/j.chbah.2024.100067
- Sestino, A., and D'Angelo, A. (2023). My doctor is an avatar! The effect of anthropomorphism and emotional receptivity on individuals' intention to use digital-based healthcare services. *Technol. Forecast. Soc. Change* 191:122505. doi: 10.1016/j.techfore.2023.122505
- Shoemaker, D. (2011). Attributability, answerability, and accountability: toward a wider theory of moral responsibility. *Ethics* 121, 602–632. doi: 10.1086/659003
- Sperber, D., Clément, F., Heintz, C., Mascaro, O., Mercier, H., Origg, G., et al. (2010). Epistemic vigilance. *Mind Lang.* 25, 359–393. doi: 10.1111/j.1468-0017.2010.01394.x
- Stivers, T., and Timmermans, S. (2020). Medical authority under siege: how clinicians transform patient resistance into acceptance. *J. Health Soc. Behav.* 61, 60–78. doi: 10.1177/0022146520902740
- Sundar, S. S. (2008). *The MAIN Model: A Heuristic Approach to Understanding Technology Effects on Credibility*. Cambridge, MA: MacArthur Foundation Digital Media and Learning Initiative, 73–100.
- Sundar, S. S. (2020). Rise of machine agency: a framework for studying the psychology of human-AI interaction (HAII). *J. Comput.-Mediat. Commun.* 25, 74–88. doi: 10.1093/jcmc/zmz026

- Sundar, S. S., Jia, H., Waddell, T. F., and Huang, Y. (2015). "Toward a theory of interactive media effects (TIME) four models for explaining how interface features affect user psychology," in *The Handbook of the Psychology of Communication Technology*, New York, NY, United States: Association for Computing Machinery, 47–86.
- Sundar, S. S., and Kim, J. (2019). "Machine heuristic: when we trust computers more than humans with our personal information," in *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, (1–9).
- Sundar, S. S., Xu, Q., and Dou, X. (2019). "The role of technology in online persuasion: a MAIN model perspective," in *Advertising Theory*, (Routledge), 70–88.
- Xia, D., Song, M., and Zhu, T. (2025). A comparison of the persuasiveness of human and ChatGPT generated pro-vaccine messages for HPV. *Front. Public Health* 12:1515871. doi: 10.3389/fpubh.2024.1515871
- Youk, S., and Park, H. S. (2023). Why scapegoating can ruin an apology: the mediated-moderation model of appropriate crisis response messages in the context of South Korea. *Front. Psychol.* 13:1082152. doi: 10.3389/fpsyg.2022.1082152
- Zanbaka, C., Goolkasian, P., and Hodges, L. (2006). "Can a virtual cat persuade you? The role of gender and realism in speaker persuasiveness," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, New York, NY, United States: Association for Computing Machinery, 1153–1162.
- Zanbaka, C. A., Ulinski, A. C., Goolkasian, P., and Hodges, L. F. (2007). "Social responses to virtual humans: implications for future interface design," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, New York, NY, United States: Association for Computing Machinery, 1561–1570.
- Zhang, J., Conway, J., and Hidalgo, C. A. (2023). "Why people judge humans differently from machines: the role of perceived agency and experience," in *2023 14th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)*, (Budapest, Hungary: 14th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)), 000159–000166.
- Zhu, C., Xu, X., Zhang, W., Chen, J., and Evans, R. (2020). How health communication via Tik Tok makes a difference: a content analysis of Tik Tok accounts run by Chinese provincial health committees. *Int. J. Environ. Res. Public Health* 17:192. doi: 10.3390/ijerph17010192

## Appendix

Pancreatic cancer is an extremely malignant digestive system tumor, with a five-year survival rate of approximately 10%. In China, pancreatic cancer ranks 9th in terms of incidence among all malignant tumors, and 6th in terms of mortality. The primary reason for this is its very low early diagnosis rate, which is less than 5%. In fact, it takes nearly 21 years for pancreatic cancer to develop from the initial pancreatic changes to metastatic pancreatic cancer, providing a sufficient window of time for early screening.

Early screening for pancreatic cancer can be achieved through fasting blood glucose and glycated hemoglobin tests, combined with serum CA19-9 testing, as well as imaging methods like CT, MRI, and endoscopic ultrasound. Statistics show that the surgical resection rate for early-stage pancreatic cancer is 90–100%, and the five-year survival rate can reach 70–100%. Therefore, the early detection, early diagnosis, and early treatment of pancreatic cancer are of significant importance in improving patients' survival and quality of life.