Social Science & Medicine 347 (2024) 116717

Contents lists available at ScienceDirect

Social Science & Medicine

journal homepage: www.elsevier.com/locate/socscimed

Justice at the Forefront: Cultivating felt accountability towards Artificial Intelligence among healthcare professionals

Weisha Wang [a], Yichuan Wang [b,*], Long Chen [c], Rui Ma [d], Minhao Zhang [e]

[a] Research Center for Smarter Supply Chain, Business School, Soochow University, 50 Donghuan Road, Suzhou, 215006, China
[b] Sheffield University Management School, University of Sheffield, Conduit Rd, Sheffield, S10 1FL, United Kingdom
[c] Brunel University London, United Kingdom
[d] Greenwich Business School, University of Greenwich, United Kingdom
[e] University of Bristol School of Management, University of Bristol, United Kingdom

ARTICLE INFO

Handling Editor: Susan J. Elliott

Keywords:
Felt accountability
Artificial Intelligence (AI)
Ethical principles
Trusting belief in AI
fsQCA
Healthcare

ABSTRACT

The advent of AI has ushered in a new era of patient care, but with it emerges a contentious debate surrounding accountability for algorithmic medical decisions. Within this discourse, a spectrum of views prevails, ranging from placing accountability on AI solution providers to laying it squarely on the shoulders of healthcare professionals. In response to this debate, this study, grounded in the mutualistic partner choice (MPC) model of the evolution of morality, seeks to establish a configurational framework for cultivating felt accountability towards AI among healthcare professionals. This framework underscores two pivotal conditions: AI ethics enactment and trusting belief in AI and considers the influence of organizational complexity in the implementation of this framework. Drawing on Fuzzy-set Qualitative Comparative Analysis (fsQCA) of a sample of 401 healthcare professionals, this study reveals that a) focusing justice and autonomy in AI ethics enactment along with building trusting belief in AI reliability and functionality reinforces healthcare professionals' sense of felt accountability towards AI, b) fostering felt accountability towards AI necessitates ensuring the establishment of trust in its functionality for high complexity hospitals, and c) prioritizing justice in AI ethics enactment and trust in AI reliability is essential for low complexity hospitals.

1. Introduction

Consider the hypothetical yet plausible scenario where an AI system, adopted by Innovative Care Hospital to interpret chest x-ray images for efficiency and cost reduction, tragically fails to detect an obvious case of pneumonia, leading to a patient's death from septic shock (Jha, 2020). This incident not only highlights the perilous consequences of AI errors in healthcare but also ignites a complex debate over liability when AI falls short. The question of who bears the responsibility for such errors—be it the healthcare provider that implemented the AI, the developer of the algorithm, or the regulatory bodies overseeing these technologies—remains mired in legal ambiguity. The case of Innovative Care Hospital underscores the pressing need for clear accountability mechanisms in the deployment of AI within healthcare settings. As AI systems become more deeply integrated into medical practice, the imperative grows to develop robust frameworks for AI accountability among healthcare professionals (Martin and Waldman, 2022; Morley and Floridi, 2020; Price et al., 2019).

Current legal frameworks are ill-equipped to navigate the intricacies posed by AI technologies, leaving a significant gap in addressing emerging challenges (Porter et al., 2022). The absence of established case law on medical AI liability (Price et al., 2019) further complicates the landscape. This complexity is exemplified by UnitedHealthcare's class-action lawsuit, spotlighting the repercussions of an AI algorithm's systematic denial of extended care claims—a situation that, upon review, revealed a staggering overturn rate of approximately 90%, raising serious concerns about the algorithm's reliability and the imperative for ethical oversight (Laney, 2023). The discourse on AI accountability in healthcare is fraught with divergent viewpoints (Saenz et al., 2023). Some argue for placing primary responsibility on AI designers and IT executives (Floridi, 2021; Martin, 2019), advocating for a stringent adherence to ethical standards and obligations. Others contend that

* Corresponding author.
E-mail addresses: wswang1004@suda.edu.cn (W. Wang), yichuan.wang@sheffield.ac.uk (Y. Wang), long.chen@brunel.ac.uk (L. Chen), rui.ma@greenwich.ac.uk (R. Ma), minhao.zhang@bristol.ac.uk (M. Zhang).

https://doi.org/10.1016/j.socscimed.2024.116717
Received 25 September 2023; Received in revised form 10 February 2024; Accepted 20 February 2024
Available online 6 March 2024

healthcare professionals should navigate AI's inherent opacity, treating technology not as a panacea but as a tool demanding meticulous scrutiny (Lebovitz et al., 2022; Miller, 2020).

In response to these challenges, several governance mechanisms aim to bolster AI accountability in healthcare. These include advocating for holistic AI design and deployment (Rana et al., 2022), assessing AI through a professional knowledge work lens (Lebovitz et al., 2021), and developing tactics to interrogate AI-driven decisions (Lebovitz et al., 2022). Yet, the role of healthcare professionals in this ecosystem remains underexplored. Moreover, the "black box" nature of AI technologies, characterized by algorithms that are neither understandable nor visible to healthcare professionals (Ananny and Crawford, 2018), exacerbates the challenge. When AI-driven decisions lead to patient harm, the propensity to shift blame to these opaque systems or external entities, such as AI solution providers, is notably problematic (DeCamp and Tilburt, 2019; Munoko et al., 2020). This scenario underlines the critical need for healthcare professionals to clearly comprehend their role in AI accountability. Therefore, our study seeks to illuminate this aspect, focusing on healthcare professionals' perceptions and responsibilities regarding AI accountability.

This study introduces a configurational framework to foster felt accountability towards AI among healthcare professionals, focusing on AI ethics enactment (beneficence, privacy, security, autonomy, justice) and trusting belief in AI (functionality, helpfulness, reliability). We build on the mutualistic partner choice (MPC) model of the evolution of morality (Baumard et al., 2013) to answer the following research question: *How AI ethics enactment and trust belief in AI simultaneously combine to elicit healthcare professionals' sense of felt accountability towards AI*?

Our research identifies justice as a pivotal component, asserting its role as a cornerstone in developing accountability perceptions among healthcare practitioners. Furthermore, we offer novel insights by demonstrating that trust in AI's reliability and functionality, particularly when aligned with justice and autonomy in AI ethics, enhances felt accountability. This contrasts with the divided perspectives on accountability in existing literature and suggests a more healthcare professional-centric approach to AI ethics in medical practice. Our findings not only advance theoretical understanding but also offer practical guidance for healthcare organizations of various complexities in cultivating AI accountability, underscoring the need for context-specific accountability building strategies that accommodate the ethical and functional demands of AI in healthcare.

2. Theoretical background and framework

2.1. Why does felt accountability towards AI matter?

Felt accountability emerges from the anticipation that one's decisions or actions will undergo evaluation by significant stakeholders, carrying the possibility of receiving rewards or sanctions based on this scrutiny (Hall and Ferris, 2011). This concept shifts the focus from external judgments of accountability to the internal perceptions held by decision-makers themselves (Hall et al., 2017). According to Hall et al. (2009), felt accountability reflects an individual's internal compass, guiding their understanding of expected duties. In healthcare, professionals navigate a complex web of accountability, answering not only to patients but also to regulatory entities for actions taken, including those influenced by AI technologies.

Defining AI accountability entails the ethical inception and creation of AI systems, designed to attribute outcomes of their actions clearly (Mikalef et al., 2022). The discourse on who bears responsibility for AI's decisions—whether the AI itself in cases of high algorithmic transparency (Ananny and Crawford, 2018; Mittelstadt et al., 2016; Thiebes et al., 2021) or humans when AI design is transparent and controllable, or when organizational reputation is at stake (Johnson, 2015; Buhmann et al., 2020)—remains unsettled. Martin (2019) suggests that IT executives and developers should assume full accountability for AI's actions, echoing the sentiment that, given AI's lack of agency, healthcare professionals should ultimately be accountable (DeCamp and Tilburt, 2019).

While existing literature frequently addresses the impacts of felt accountability (Wikhamn and Hall, 2014) and occasionally its antecedents (Mero et al., 2014), a gap persists in understanding the specific conditions that foster a high level of felt accountability towards AI in healthcare settings. This gap signifies an uncharted domain ripe for investigation. The debate often orbits around external perspectives of imposing accountability rather than delving into the experiential reality of professionals themselves (DeCamp and Tilburt, 2019; Munoko et al., 2020; Siala and Wang, 2022), thus leaving a critical examination of the internalization of accountability in the context of AI use within healthcare organizations as an area in dire need of exploration.

2.2. Exploring configurations of the particular set of the conditions for cultivating felt accountability

We introduce a configurational model to investigate how healthcare professionals develop felt accountability towards AI, drawing upon the Mutualistic Partner Choice (MPC) model from evolutionary morality. The MPC model suggests that mutually beneficial relationships stem from choosing partners who exhibit fairness and reliability, which are key to enhancing cooperative fitness (Baumard et al., 2013; Everett et al., 2016). In translating this to AI, we acknowledge a fundamental difference: unlike humans, AI lacks innate morality and operationalizes ethical principles as programmed by humans. Despite this, AI designed within ethical frameworks can still be construed as "digital partners," earning trust through consistent ethical behavior. The evolving landscape of AI in healthcare demonstrates that these systems are transitioning from passive tools to proactive entities capable of autonomous action and adaptation (Baird and Maruping, 2021; Fügener et al., 2022; Tarafdar et al., 2023), paralleling human partners in the MPC model.

Our study examines healthcare professionals' development of felt accountability towards AI, recognizing that AI's adherence to ethical standards is contingent on human design and governance. However, AI's ethical actions are not self-derived but are instead reflections of ethical guidelines implemented by developers, posing a risk of "ethics washing" if not accompanied by genuine ethical practice and oversight (Hao, 2019). In the rapidly evolving AI landscape, where AI transition from passive tools to proactive entities, the alignment with ethical standards becomes critical for establishing digital trustworthiness. The delegation of tasks to these agentic IS artifacts, as Baird and Maruping (2021) noted, involves a deliberate transfer of rights and responsibilities, highlighting a conscious act of trust from healthcare professionals who must navigate the complexities of AI use. The pressing need for ethical oversight, underscored by the consequences of premature and unchecked AI deployment noted by Hao (2019), calls for an integration of ethical considerations that go beyond mere compliance. It requires a proactive institutionalization of ethics, akin to Tseng (2019), which can ensure that AI systems are not only designed to act ethically but are also deployed and monitored to uphold those standards continuously.

A pertinent example of this dynamic is the EyeArt AI system developed by Eyenuk. With robust testing and FDA approvals, EyeArt has demonstrated the MPC model's principles of fairness and reliability in practice. By autonomously detecting diabetic retinopathy and transparently processing diverse patient data, it has gained trust and exemplified how ethical alignment in AI design can underpin the trust and felt accountability of healthcare professionals. This shows that when AI tools provide transparent, unbiased, and confirmable outputs, they mirror the ethical judgment expected of human agents, thereby fostering a reciprocal sense of ethical responsibility and accountability (Pallardy, 2023).

Given these perspectives, we postulate that if healthcare organizations embed AI ethics effectively and healthcare professionals trust the AI's reliability and adherence to these ethical standards, they are more inclined to express heightened felt accountability. It underscores the potential symbiotic relationship where the ethical implementation of AI and the trust it garners could serve as foundational conditions for cultivating a profound sense of felt accountability among healthcare professionals. The MPC model, explaining the development of healthcare professionals' felt accountability towards AI through AI ethics enactment and trusting belief in AI, is illustrated in Fig. 1. In Fig. 2, we present our configurational model with a set of conditions. A summary of the conditions and outcomes considered in the model, including the definitions, justification, and supporting sources of each condition, is presented in Web Appendix 1.

2.3. The enactment of AI ethics

The ethical deployment of AI in healthcare necessitates a nuanced understanding of ethical principles that guide both the development and application of AI technologies. According to De Togni et al. (2021), Floridi et al. (2018), and Morley et al. (2020), the ethical enactment of AI is crucial for aligning healthcare professionals' actions with ethical norms, ensuring AI's integration into clinical workflows enhances patient care while adhering to ethical standards. This enactment assists in identifying and mitigating systematic biases in AI decision-making, outlining a comprehensive regulatory framework for its clinical use. Yet, empirical exploration into which ethical aspects should be prioritized to cultivate healthcare professionals' felt accountability towards AI remains sparse, a gap acknowledged by Greene et al. (2019), Jobin et al. (2019), and Munoko et al. (2020).

Floridi et al. (2018) propose an AI ethics framework, in alignment with the four principles of biomedical ethics proposed by Beauchamp and Childress (2001), highlighting the interaction between healthcare and AI technologies: beneficence, non-maleficence (comprising privacy and security), autonomy, and justice. These five key principles of ethical AI are believed to address the ethical challenges in the AI development in the healthcare sector (Floridi et al., 2018).

Beneficence transcends the traditional confines of well-being and encompasses a holistic approach towards the social, environmental, and common good. This principle emphasizes the development of AI to ultimately promote the well-being not only of humans but of all sentient creatures (Thiebes et al., 2021). This principle extends beyond protecting and benefiting patients and professionals, as suggested by Reddy et al. (2020), to ensuring that AI technologies are designed, deployed, and utilized in a manner that prioritizes the common good. Therefore, beneficence in AI, particularly in healthcare, encompasses a commitment to creating technologies that serve a broader spectrum of societal and environmental well-being, upholding Floridi et al.'s (2018) vision of promoting the well-being of all stakeholders involved.

Non-maleficence in AI ethics, as described by Jobin et al. (2019), extends beyond mere caution against negative outcomes; it encompasses a proactive stance against both the unintentional and intentional misuse of AI technologies that could lead to harm. This approach to non-maleficence is reflected in the Montreal Declaration's call for developers to actively work against risks from their technological innovations and the European Group on Ethics (EGE)'s emphasis on responsibility. This necessitates a dual focus: preventing harms arising from human intentions and those stemming from the unpredicted behavior of AI, including its unintended influences on human behavior as noted by Floridi et al. (2018). Building on this foundation, the principle of non-maleficence also profoundly emphasizes the importance of ***security*** and ***privacy***. Privacy concerns are not merely about data protection; they are deeply interwoven with patients' autonomy, giving them control over their patient data and its usage (Floridi et al., 2018). This aspect of non-maleficence thus ensures that AI operates within ethical boundaries respecting patient rights. Meanwhile, security extends this principle further, underscoring the need for medical AI systems to function within secure and well-defined limits (Floridi et al., 2018).

Autonomy in the context of AI implicates the patients' right to informed decision-making regarding their treatment, considering the delegation of certain decisions to AI (Floridi et al., 2018). This principle is explored through the lens of "meta-autonomy," examining the balance between human oversight and AI's decision-making capabilities, ensuring AI's reliability and integrity are maintained (Dalton-Brown, 2020).

Justice, within the sphere of AI implementation, encompasses a broader obligation than merely distributing benefits equitably, as highlighted by Newman et al. (2020). AI-based systems in healthcare are expected to operate with integrity, honesty, and sincerity, consistently adhering to these expanded ethical AI principles (Floridi et al., 2018). Moreover, there is a crucial need to ensure that AI does not exacerbate existing disparities or create new forms of inequity. This involves a vigilant approach to algorithmic design, ensuring fairness in data handling and algorithmic decision-making, and safeguarding against biases that might compromise the equitable treatment of patients. Thus, in the healthcare context, the principle of justice in AI goes beyond non-discrimination, embedding a proactive approach to fostering equity and solidarity within and beyond the healthcare ecosystem (Floridi et al., 2018).

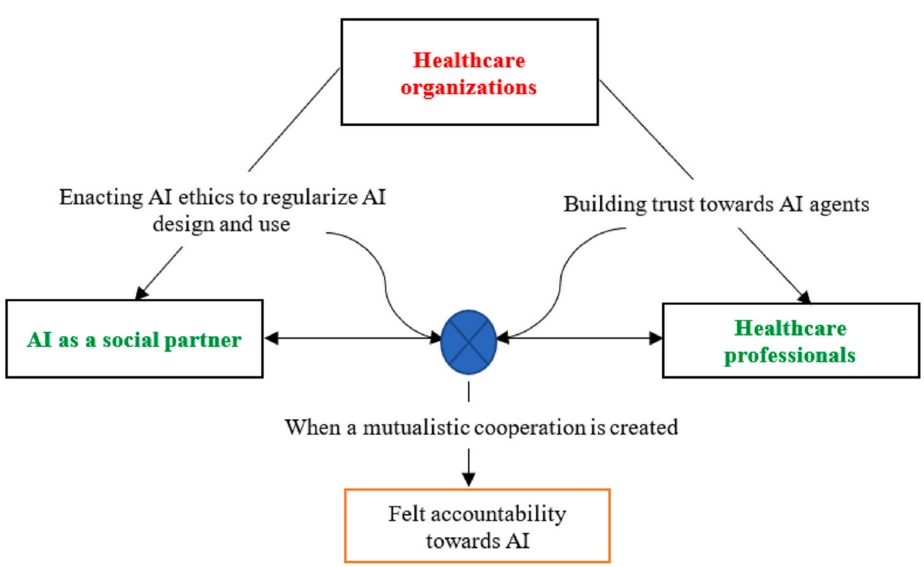

Fig. 1. The mutualistic partner choice model explaining the development of healthcare professionals' felt accountability towards AI.

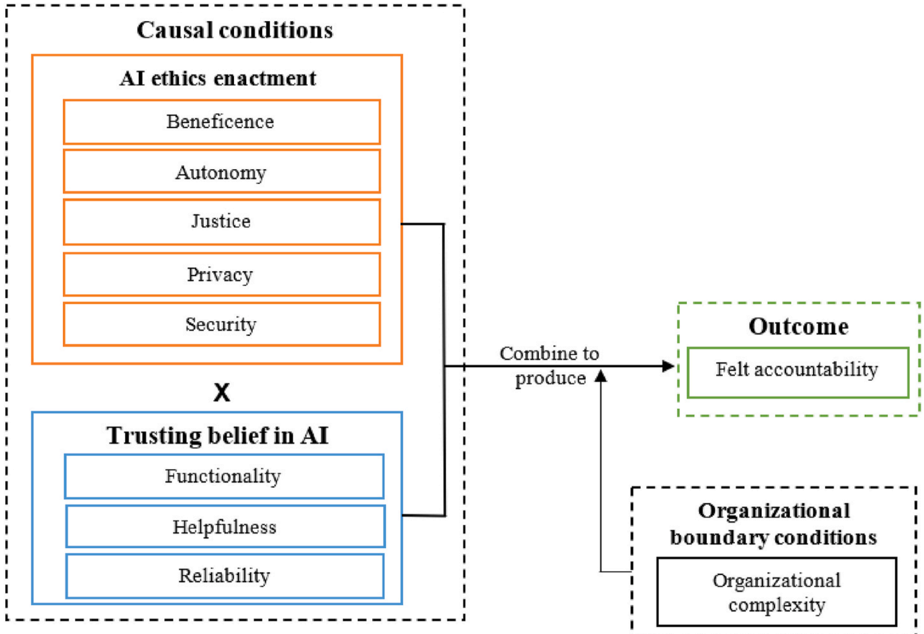

Fig. 2. Research model: A confgurational perspective.

2.4. Trusting belief in AI

The deployment of AI within healthcare is predicated not only on its technological prowess but also on its perceived trustworthiness (Thiebes et al., 2021). Trust in AI centralizes on technological merits, sidelining personal agency and focusing on informed trust following user experience with the system (McKnight et al., 2011; Thiebes et al., 2021). Discussions surrounding trustworthy AI primarily concentrate on its necessity, conceptualization, and potential benefits at a conceptual level (e.g., De Togni et al., 2021; Thiebes et al., 2021). Empirical evidence linking these dimensions of trusting belief with healthcare professionals' felt accountability is scant. This research focuses on trust that pertains to the information technology artifacts - trusting belief (or trust) in AI. Trust is generally defined as an individual's willingness to depend on another party due to the characteristics of the latter (Mayer et al., 1995). Trusting belief in AI represents the knowledge users have acquired during post-adoption interaction with the technology in various contexts, reflected in three key beliefs: functionality, helpfulness, and reliability (Mcknight et al., 2011).

Mcknight et al. (2011) propose that functionality, helpfulness, and reliability are the three aspects that reflect people's trust in information technology artifacts. Functionality refers to the expectation of the technology's capacity or capability to complete required tasks. In AI context, functionality is the belief in AI's capability to perform tasks effectively, a notion particularly resonant in healthcare where AI's predictive analytics can significantly affect patient outcomes. AI must not only be advanced in its capabilities but must also align with the complex workflows and the dynamic nature of clinical environments (Rajpurkar et al., 2022). Helpfulness pertains to the supportive function of the technology, excluding its agency and volition. Helpfulness reflects the supportive role of AI in enhancing healthcare delivery. It underscores the expectation that AI systems assist professionals without usurping their critical decision-making authority. This supportive function must be contextualized within healthcare settings, ensuring AI augments rather than diminishes the clinician's role. Reliability, on the other hand, underscores the expectation that the technology will consistently operate properly (Mcknight et al., 2011). Reliability stresses the necessity for AI systems to function dependably and predictably (Shneiderman, 2020). In a domain where decisions can be life-altering, the consistent performance of AI is non-negotiable. Healthcare professionals may rely on AI not only in routine scenarios but also in high-stakes situations demanding impeccable accuracy.

3. Method

3.1. Data collection

To gather data for our study, we reached out to healthcare professionals working in hospitals located in the southwestern region of China. We specifically chose the southwestern region of China for our study due to its unique convergence of traditional healthcare practices and rapid adoption of AI technologies, influenced by progressive local policies. This region presents a microcosm of the broader challenges and opportunities in implementing AI across diverse healthcare systems. After obtaining approval from the top management teams, a survey questionnaire was distributed to 1452 active healthcare professionals. To provide study context, all respondents received a summary of medical AI applications and exemplars from a white book published by the Chinese Innovative Alliance of Industry, Education, Research, and Application of Artificial Intelligence for Medical, as presented in Web Appendix 2. Only healthcare professionals who reported having experience using AI technologies in their workplaces were included in the study, which was indicated by a screening question ("Have you used AI technologies in your workplace before?"). Those who reported no experience with AI were excluded. All participants were informed of the research purpose and ensured that their responses would be confidential and voluntary.

In our study, we differentiated organizational complexity into two categories: tertiary public hospitals and secondary and primary hospitals, based on the classification of Chinese hospitals, proposed by the Ministry of Health of the People's Republic of China overseeing all hospitals in China (https://en.wikipedia.org/wiki/Classification_of_Chinese_hospitals). This comprehensive classification incorporates multiple dimensions, including the level of service provision, size, medical technology, medical equipment, and management and medical quality. This approach has been adopted by Park et al. (2017). In our context, it allows for an assessment of how healthcare professionals interact with AI technologies, which may vary depending on the hospital classification.

The study received a total of 413 responses, but 12 were removed due to failing the screening questions, leaving 401 valid responses for analysis. The responses were anonymized and aggregated for further analysis. Among the participants, 217 were female (54.1%) and 184 were male (45.9%), with a mean age of 31 years old (SD = 5.13). There were 194 participants from high complexity hospitals and 207 participants from low complexity hospitals. To achieve better response rate, we implemented several strategies, including personalized email invitations, incentives for participation, and the timely reminder emails sent throughout the data collection period. A breakdown of the demographics is presented in Table 1.

3.2. Fuzzy-set qualitative comparative analysis

The primary objective of this study is to investigate the combined influence of AI ethics enactment and trusting belief in AI within varying organizational complexities. fsQCA is inherently suited for such analysis as it excels in exploring configurational effects and identifying combinations of conditions leading to a specific outcome (Ragin, 2008). Such a method complements regression analysis (e.g., structural equation modeling) by highlighting that not all the conditions are necessary to explain the outcomes, as some of the combinations of the conditions can be sufficient to explain the outcome (Furnari et al., 2021). We developed the evaluation criteria (see Table 2) to justify the use of fsQCA for this study.

3.3. Measurement and validation

Measurement items of the key constructs were adapted from previous studies and were modified to fit our research context. Several steps have been taken to validate all the measurement items, including 1) AI experts and healthcare professionals involved in finalizing the measurement scales, 2) back-to-back translation was adopted, 3) reliability assessments and 4) convergent validity, and discriminant validity assessments (see Web Appendix 3).

We have implemented several of Podsakoff et al.'s (2003) procedural remedies to minimize the risk of common method bias (CMB). This includes guaranteeing respondents' anonymity, which reduces evaluation apprehension and social desirability bias, and ensuring that the survey questions were clearly worded to avoid ambiguity and reduce measurement error. In addition to procedural steps, we performed Harman's single-factor test and tested the unmeasured latent method construct technique recommended by Podsakoff et al. (2003). These tests did not indicate a significant presence of CMB in our dataset.

3.4. Calibration of set memberships

To prepare for the fsQCA analysis, continuous variables need to be calibrated into set-membership scores based on their degree of membership in sets of cases. Ragin (2008) proposed three qualitative anchors for calibration: full membership, full non-membership, and the crossover point (Rihoux and Ragin, 2008). In fsQCA, variables are calibrated into fuzzy membership scores ranging from 0 to 1, where 0 indicates full non-membership, 1 indicates full membership, and 0.5 indicates the point of maximum ambiguity on whether a case belongs to this group or not. To calibrate the variables, we use Ragin's (2008) direct method and the fsQCA software (Ragin et al., 2008), which is recommended by Pappas and Woodside (2021) due to its rigorousness and ease of replication and validation.

To calibrate the continuous variables into set-membership scores, three qualitative anchors were used: full membership, full non-membership, and the crossover point. The values were calibrated into fuzzy sets with values ranging from 0 to 1, where 0 indicates full non-membership, 1 indicates full membership, and 0.5 indicates the point of maximum ambiguity. Three percentiles were set as the thresholds for full membership, crossover points, and full non-membership. A value of 0.001 was added to the calibrated value of exactly 0.50 to ensure that no cases were dropped from the analysis. Three separate fsQCA analyses were performed for all hospitals and two main hospital groups in the dataset, in line with considerations regarding hospital complexity. This method was based on the works of Pappas and Woodside (2021) and Ragin (2008).

4. Results

4.1. Necessity analysis

A necessity analysis was conducted to determine whether any individual condition is necessary for achieving high levels of felt accountability. The analysis compared each condition against the outcome of felt accountability (felt accountability and ~felt accountability) using Ragin's (2008) method. The results indicate that no condition has a consistency value above the threshold of 0.9, as shown in Web Appendix 4 (Schneider and Wagemann, 2012). Therefore, it can be concluded that no single condition is necessary for achieving high levels of felt accountability. A sufficiency analysis was conducted to identify which configurations are sufficient for achieving high levels of felt accountability.

Table 1
Demographic breakdown (n = 401).

Variables	Items	High Complexity Hospitals (n = 194)		Low Complexity Hospitals (n = 207)		Total (n = 401)	
		Frequency	Percentage	Frequency	Percentage	Frequency	Percentage
Gender	Male	87	44.8%	97	46.9%	184	45.9%
	Female	107	55.2%	110	53.1%	217	54.1%
Age	18–25	26	13.4%	1	0.5%	27	6.7%
	26–33	160	82.5%	92	44.4%	252	62.8%
	34–41	6	3.1%	103	49.8%	109	27.2%
	42 and above	2	1.0%	11	5.3%	13	3.3%
Degree of Education	Bachelor's degree	16	9.3%	2	0.9%	18	4.5%
	Master's degree	119	61.3%	161	77.8%	280	69.8%
	Doctoral degree and above	59	30.4%	44	21.3%	103	25.7%
Job title	Intern	13	6.7%	7	3.4%	20	5.0%
	Resident physician	132	68.0%	44	21.3%	176	43.9%
	Doctor-in-charge	8	4.1%	98	47.3%	106	26.4%
	Associate senior doctor	2	1.0%	48	23.2%	50	12.5%
	Senior doctor	1	0.5%	10	4.8%	11	2.7%
	Others	38	19.6%	0	0.0%	38	9.5%
Length of Working (year)	Less than one year	3	1.5%	0	0.0%	3	0.8%
	1–4	174	89.7%	17	8.2%	191	47.6%
	5–8	11	5.7%	116	56.0%	127	31.6%
	9–12	3	1.5%	61	29.5%	64	16.0%
	13 and above	3	1.5%	13	6.3%	16	4.0%

Table 2
The evaluation criteria to justify the use of fsQCA.

Criteria	Our approach to address
Justification for fsQCA	• We compared fsQCAwith with crisp-set QCA (csQCA) and multi-value QCA (mvQCA). Our rationale for selecting fsQCA lies in its nuanced treatment of partial membership, allowing for a more fine-grained analysis of the conditions leading to the perception of accountability in AI use among healthcare professionals. This method excels in handling the grey areas of membership, reflecting the complex reality of ethical considerations in AI deployment (Ragin, 2008; Woodside, 2013; Pappas and Woodside, 2021). • fsQCA's accommodation of causal asymmetry aligns perfectly with the objectives of our research. It enables us to discern between conditions that are merely associated with the outcome from those that are determinants of the outcome. This distinction is vital in our calibration of set memberships, where we aim to understand not just the presence of conditions but their potency in influencing felt accountability.
Theoretical grounding for conditions and outcomes	Anchored conditions and outcomes in the MPC model, ensuring a strong theoretical basis for our fsQCA, detailed in Web Appendix 1.
Complementarity with other methods	fsQCA complements the regression analysis provided by Partial Least Squares Structural Equation Modeling (PLS-SEM, as shown in Web Appendix 6) by underscoring that all conditions are not essential to account for the outcomes, a finding that is corroborated by the PLS-SEM analysis.
Familiarity with cases	• To achieve a more robust understanding of our cases, we initially identified relevant conditions by reviewing China's national AI policy for healthcare and scrutinizing AI adoption trends across hospitals. • Our exploration uncovered the variability in outcomes across different hospital complexities, revealing how distinct combinations of conditions could lead to similar outcomes. fsQCA's adaptability in identifying these variances provided invaluable insights, elucidating the diverse accountability landscapes within high and low complexity hospitals.
Number of conditions	Maintained a balanced condition-to-case ratio (9 conditions for 401 cases) to ensure fsQCA reliability, adhering to recommended methodological guidelines.

4.2. Truth table analysis

To further understand the sufficiency analysis in fsQCA, a truth table analysis is used to identify possible configurations of AI ethical enactment and trusting belief that can result in high levels of felt accountability. This process generates a truth table with 2k rows, where each row represents a unique configuration of conditions. The truth table also calculates and sorts the frequency of cases for each combination, the raw consistency of each combination, and the proportional reduction in inconsistency (PRI) consistency of each combination, allowing for the identification of the most frequent and consistent configurations of conditions that lead to high levels of felt accountability (Ragin, 2014).

The frequency of a configuration represents how many cases in the sample are accounted for by that specific combination of conditions. When a configuration has a frequency of zero, it means that none of the cases in the sample can be explained by that combination. To ensure meaningful and valid results, we set the frequency threshold at 3 (Fiss, 2011; Ragin, 2014) and removed all combinations with smaller frequencies from further analysis. After filtering out infrequent configurations, the truth table is sorted by raw consistency and PRI consistency to identify the combinations with high scores on the outcome variable. Raw consistency measures the extent to which a configuration is a subset of cases with high scores in the outcome variable. PRI consistency is an alternative measure of consistency that is based on a quasi-proportional reduction in error calculation (Ragin, 2014). We set a cut-off of 0.8 for raw consistency and 0.7 for PRI consistency, meaning that we only considered combinations with a raw consistency of at least 0.8 and a PRI consistency of at least 0.7 as reliably resulting in felt accountability (Pappas and Woodside, 2021; Rihoux and Ragin, 2008). Next, the outcome variable was assigned a value of 1 when a configuration had a raw consistency above 0.8 and a PRI consistency above 0.7, indicating that the combination explains the outcome, and 0 otherwise, indicating that it does not.

After applying the truth table algorithm, fsQCA generated three solutions: intermediate, parsimonious, and complex. The complex solution includes all possible combinations of conditions, making it impractical for interpretation. On the other hand, the parsimonious solution presents the most important conditions that must be included in any solution, also known as the core condition. However, it may miss some theoretically plausible counterfactuals. The intermediate solution includes both parsimonious and additional conditions known as complementary conditions, which make a relatively weaker contribution to the outcome (Fiss, 2011). To identify the core and complementary causal conditions contributing to high levels of felt accountability in hospitals, we followed the approach of Fiss (2011), considering both the parsimonious and intermediate solutions. We summarized the causal recipes for all hospitals in Table 3, and Tables 4 and 5 present the causal recipes for achieving high levels of felt accountability in high and low complexity hospitals, respectively.

Consistency and coverage scores are important measures for validating the solutions in fsQCA. A high consistency score above the threshold of 0.80 indicates that the configurations are sufficient in predicting the outcomes of interest in both hospital groups (Ragin, 2008). Additionally, coverage scores describe the proportion of memberships in the outcome explained by each solution, while the overall coverage describes the extent to which the outcome of interest can be explained by the configurations (Ragin et al., 2008). Although a higher coverage does not necessarily imply theoretical importance, it denotes the empirical relevance and effectiveness of the solution for the outcome (Ragin, 2008).

4.3. Rigorousness and robustness check for fsQCA findings

We employed 10 evaluation criteria to ensure the rigor and robustness of the fsQCA performed in our study (Schneider and Wagemann, 2010; Park et al., 2020). Web Appendix 5 outlines our approach to addressing these criteria.

4.4. Theorizing the cultivation of felt accountability towards AI among healthcare professionals in diverse organizational contexts

The configurations revealed through our analysis consistently demonstrate various parallel solutions for achieving a high level of felt accountability towards AI. Table 3 reports seven configurations of AI ethics enactment and trusting belief in AI associated with high levels of felt accountability towards AI for all hospitals. The overall coverage score was 50.0%, indicating that the configurations explained half of the variance in felt accountability.

Our findings indicate that to foster a deep sense of accountability towards AI among healthcare professionals, it is crucial to emphasize justice and autonomy within AI ethics. This aligns with Solomonides et al. (2022), who emphasized the importance of these principles in ethically implementing AI in healthcare settings. Building on this, Gillner (2024) contends that effectively integrating justice and autonomy into AI, particularly within the complex subsystems of healthcare,

Table 3
Configurations for achieving high levels of felt Accountability in all hospitals (n = 401).

	Solutions						
	S1	**S2**	**S3**	**S4**	**S5**	**S6**	**S7**
AI ethics enactment							
Beneficence		•	•	●		●	●
Autonomy	●		●	●	●	●	●
Security	•	●	•		●		●
Privacy		•					•
Justice	●	●	●	●	●	●	●
Trusting belief in AI							
Reliability	●	•	●	●	●		
Functionality		●		●	●	●	●
Helpfulness	•		•	•	•	•	•
Consistency	0.919	0.939	0.942	0.954	0.948	0.936	0.936
Raw Coverage	0.218	0.269	0.321	0.312	0.332	0.102	0.214
Unique Coverage	0.016	0.066	0.007	0.009	0.015	0.015	0.007
Overall solution consistency	0.9147						
Overall solution coverage	0.5001						

Note: Black circles (●) denote the presence of a condition, big circles = core conditions; small circles = complimentary conditions; blank spaces denote a "do not care" condition.

Table 4
Configurations for achieving high levels of felt accountability for the high complexity hospitals (n = 194).

	Solutions					
	S1a	**S2a**	**S3a**	**S4a**	**S5a**	**S6a**
AI ethics enactment						
Beneficence	•	•		•	•	•
Autonomy	•		•	●	●	●
Security	•	•	•	•		
Privacy	⊗	⊗	⊗	●	●	●
Justice	●	●	●	•		•
Trusting belief in AI						
Reliability		•	•	•		
Functionality	●	●	●	●	●	●
Helpfulness	•	•	•			•
Consistency	0.969	0.954	0.964	0.870	0.865	0.890
Raw Coverage	0.182	0.179	0.188	0.144	0.083	0.080
Unique Coverage	0.015	0.017	0.027	0.034	0.031	0.015
Overall solution consistency	0.9045					
Overall solution coverage	0.3303					

Note: Black circles (●) denote the presence of a condition, crossed-out circles (⊗) denote the absence of a condition; big circles = core conditions; small circles = complimentary conditions; blank spaces denote a "do not care" condition.

requires a dynamic approach. This approach must ensure equitable AI resource distribution and maintain healthcare professionals' autonomy in decision-making.

However, previous studies may not have fully acknowledged the vital role of healthcare professionals' trust in AI. Our fsQCA results offer a new perspective, showing that the principles of justice and autonomy gain significantly more impact when merged with the establishment of trust in AI's reliability and functionality. This is evidenced in configurations S1, S3, S4, and S5, and S4, S5, S6, and S7. It implies that intertwining healthcare professionals' trust in AI's reliability and functionality with a strong ethical framework is key to cultivating their sense of accountability towards AI technologies. In this regard, AI should not be viewed as an autonomous entity but rather as a tool that augments healthcare professionals' commitment to their patients, adhering to the ethical principles of justice and autonomy (Hatherley, 2020). Concurrently, the implementation of AI in healthcare demands careful management to maintain the essential elements of trust that healthcare professionals perceive. Based on the foregoing discussion, we suggest

Table 5
Configurations for achieving high levels of felt accountability for low complexity hospitals (n = 207).

	Solutions			
	S1b	**S2b**	**S3b**	**S4b**
AI ethics enactment				
Beneficence		●	●	●
Autonomy	●		●	●
Security	●	●		
Privacy		●	⊗	⊗
Justice	⬤	⬤	⬤	⬤
Trusting belief in AI				
Reliability	⬤	⬤	⬤	
Functionality	●	●	●	●
Helpfulness	●			●
Consistency	0.902	0.934	0.945	0.922
Raw Coverage	0.403	0.169	0.071	0.076
Unique Coverage	0.286	0.071	0.018	0.018
Overall solution consistency	0.9036			
Overall solution coverage	0.5175			

Note: Black circles (●) denote the presence of a condition, crossed-out circles (⊗) denote the absence of a condition; big circles = core conditions; small circles = complimentary conditions; blank spaces denote a "do not care" condition.

the proposition:

Proposition 1. *Enhancing healthcare professionals' felt accountability towards AI in clinical settings requires an integrated approach that emphasizes justice and autonomy within AI ethics enactment while also fostering their trusting belief in AI's reliability and functionality.*

Within high-complexity hospital environments, our fsQCA analysis has delineated six distinct configurations (S1a to S6a) that contribute to nurturing a high level of felt accountability towards AI among healthcare professionals. As illustrated in Table 4, the overall coverage score of 33.0% denotes that these configurations account for a substantial, though not exhaustive, share of the variance in felt accountability. Notably, configuration S3a, with the highest coverage of 18.8%, underscores that integrating justice within AI ethics, bolstered by a secure and autonomy framework and trust in AI's functional performance, stands out as the most effective strategy for elevating felt accountability.

This approach is corroborated by Buhmann et al. (2020) and Cath (2018), who argue that genuine accountability transcends mere procedural compliance, demanding that AI needs to deliver tangible, functional outcomes. The challenges of implementing AI in practice, particularly in complex healthcare settings, compound the ethical decisions and dilemmas faced by healthcare professionals. The recurrent focus on justice, as Solomonides et al. (2022) have observed, reflects a commitment to using AI equitably and fairly, thus adhering to the ethical standards of medical practice. Therefore, it is posited that a meaningful integration of AI in healthcare necessitates the alignment of functional, reliable, and helpful AI tools within a framework rooted in justice. Such an integration is vital to foster an environment where healthcare professionals can not only trust but also feel a sense of responsibility for the AI systems they employ. Hence, we propose the following refined statement:

Proposition 2. *To engender a substantial sense of felt accountability towards AI within high-complexity hospital settings, it is imperative to prioritize justice as a core tenet of AI ethics enactment. This priority must be supported by a trusting belief in AI's functionality and further reinforced by additional ethical principles such as security and autonomy.*

The analysis of low complexity hospitals identified four solutions (S1b to S4b) that can be adopted to achieve high levels of felt accountability, as presented in Table 5. The overall coverage score was 51.8%, indicating that the configurations explain about half of felt accountability. In configurations S1b (with the highest coverage score of 0.403) and S3b, justice appears as a condition associated with high levels of felt accountability, while reliability is present in three configurations. This suggests that justice as an ethical principle and reliability as a trusting belief are considered necessary for fostering felt accountability towards AI among healthcare professionals in less complex hospital environments. In less complex settings, where healthcare professionals might engage with AI on a more routine or less critical basis, the perceived fairness and reliability of AI could be more influential than the autonomous decision-making capacity of AI systems (Shneiderman, 2020). Furthermore, reliability serves as a foundational trust factor that can directly impact the user's trust in technology (Mcknight et al., 2011), which is especially relevant in healthcare where trust is paramount due to the potential consequences of AI decision-making on patient outcomes (Gillner, 2024). Thus, we propose:

Proposition 3. *In low complexity hospital environments, emphasizing justice within AI ethics enactment and building trust in AI reliability are essential components within configurations that lead to high levels of felt accountability towards AI among healthcare professionals.*

5. Discussion

5.1. Theoretical contributions

This study advances the dialogue on AI accountability within healthcare by adopting a configurational perspective, examining how AI ethics enactment and trusting belief in AI intertwine to influence healthcare professionals' felt accountability towards AI. We address the call by Hall et al. (2017) for research into the antecedents of perceived accountability and respond to the pleas made by Greene et al. (2019), Jobin et al. (2019), and Munoko et al. (2020), who urged for empirical investigations into the enactment of ethical AI. Ethical AI enactment has been widely recognized as an important consideration in the AI implementation process in the healthcare sector (Floridi et al., 2018). While prior conceptual research and policies acknowledge the importance of ethical AI principles (e.g., Floridi et al., 2018; Jobin et al., 2019; Thiebes et al., 2021), it remains unknown as to facets of AI ethics should be emphasized to foster a sense of felt accountability among healthcare professionals towards medical AI.

Our fsQCA results discover that justice within AI ethics is the linchpin for cultivating strong accountability perceptions among healthcare professionals. Notably, while multiple pathways can lead to high levels of felt accountability, four out of the seven identified solutions denote that a combination of ensuring justice and autonomy in AI ethical enactment and fostering trust in AI reliability results in heightened levels of felt accountability. Contrasting with the existing literature's divided stance on AI accountability—which oscillates between assigning responsibility to AI solution developers, organizations, or healthcare professionals—, our findings contribute empirical insights into prioritizing AI ethics components to bolster AI accountability in medical practice from healthcare professional perspective. Furthermore, complementary PLS-SEM results also provide additional support and confirm that justice is the most valued principle (Web Appendix 6).

Intriguingly, privacy was absent in the most effective solution for all hospitals, echoed by the supplementary PLS-SEM analysis. This could be reflective of the unique challenges in managing patient privacy within the socio-cultural and legal milieu of China. AI technologies necessitate vast, continually updated patient datasets for optimal diagnostics, entailing data sharing among hospitals and stakeholders (Hathaliya and Tanwar, 2020). Given this requirement, patient privacy is often pragmatically balanced against the imperatives of medical AI development. Furthermore, the prevailing collectivist ethos in China, which prioritizes societal harmony and authority obedience over individual privacy rights (Tam, 2018), may contribute to a distinctive attitude towards personal privacy in healthcare settings. Kui (2021) suggests that while the cultural context may influence perceptions of privacy, the absence of robust legal protections for personal data in China compounds the issue, casting personal privacy as a distant ideal rather than an immediate reality. Despite this, the principle of security is valued and seen as vital by healthcare professionals, as it underpins the safeguarding of patient data against breaches (Hathaliya and Tanwar, 2020). This distinction between the concepts of privacy and security suggests that while privacy in the context of AI may not yet be a salient concern, the protection of patient information from potential misuse remains crucial.

Finally, we underscore a gap in the literature regarding strategic trust-building in AI across healthcare organizations of varying complexity. While trust in AI is acknowledged as pivotal, strategies to cultivate this trust must be context-specific, adaptable to the organizational scale and the nature of healthcare delivery. Our findings highlight that in high-complexity hospitals, which are often early adopters of AI, trust is primarily derived from the functionality of AI systems. These hospitals, being more advanced in their use of technology, also appreciate the value of autonomy in the ethical deployment of AI, although to a lesser extent. In contrast, low-complexity hospitals place greater emphasis on justice within AI ethics, regarding it as essential for cultivating felt accountability among healthcare professionals. These hospitals value how AI systems are programmed to treat all patients and their data with equity and impartiality. The difference between high and low complexity hospitals suggests that as healthcare organizations become more acquainted with AI, the emphasis may shift from AI ethical principles to practical functionalities of AI systems. In essence, for healthcare professionals in varying institutional contexts, the salience of AI's ethical and functional attributes in building trust—and by extension, felt accountability—differs, reflecting the complexity of integrating AI into diverse healthcare environments.

We have synthesized our key findings and theoretical implications into a visual manner, demonstrating the evolution of our understanding of healthcare professionals' felt accountability towards AI before and after our investigation as shown in Fig. 3.

5.2. Practical implications

The implementation of AI technology in the healthcare sector has tremendously changed the clinical routines of the healthcare sector in recent years. When many private and public hospitals are rushing to embed machine learning into medical decision-making (Char et al., 2018), shaping healthcare professionals' sense of accountability for their AI-driven clinical decisions has become an AI deployment priority. To operationalize this accountability, we propose targeted strategies for healthcare administrators and policymakers.

First, healthcare organizations should assess the integration of AI within their operations to identify new roles or skill sets required. This step involves a thorough analysis of how AI complements existing roles and the specific competencies needed for effective AI utilization. Tailoring professional development to these findings ensures that healthcare professionals are equipped to harness AI's potential responsibly. Healthcare organizations can also establish partnerships with technology companies, academic institutions, and regulatory bodies to gain access to the latest AI research, tools, and training resources. These

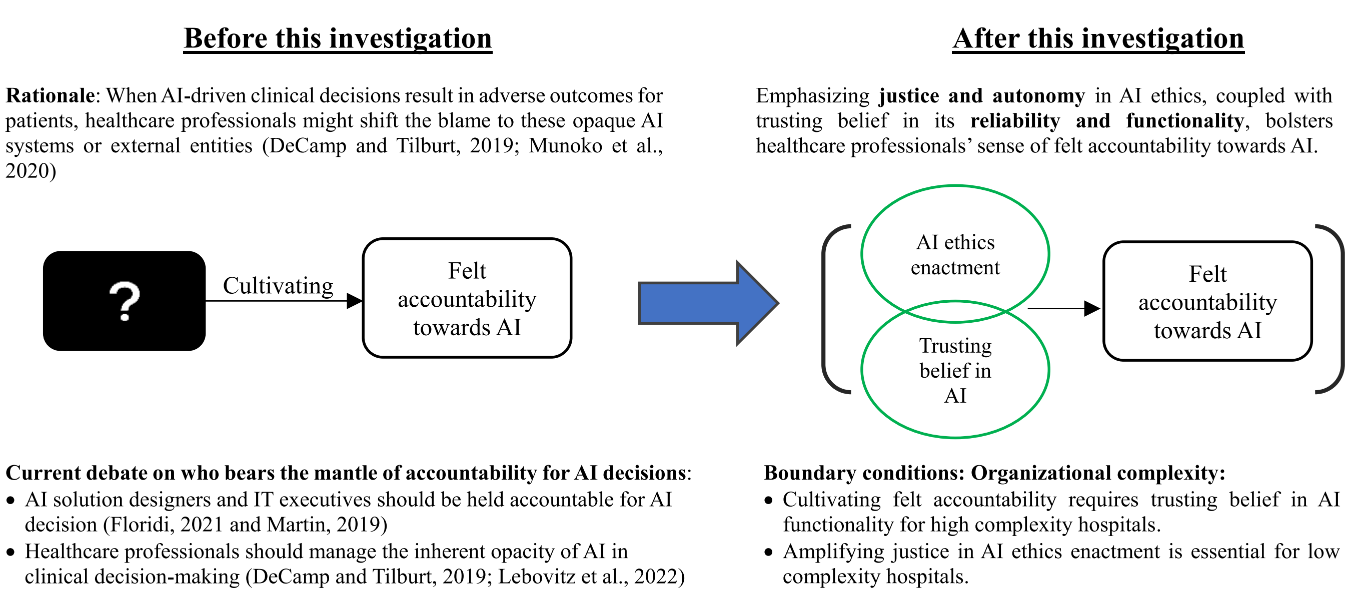

Fig. 3. Main value added of this investigation.

collaborations can also help healthcare organizations stay abreast of ethical guidelines and compliance requirements.

Second, continuous education on the ethical dimensions of AI, particularly concerning algorithmic justice, is essential. Healthcare institutions should prioritize training that addresses the risks of bias in AI algorithms and explores practical measures to counteract such biases. This training should emphasize the ethical implications of AI decisions and foster a culture of critical engagement with AI technologies. The implementation plan can include 1) schedule regular training sessions as part of continuing medical education credits, ensuring healthcare professionals remain current on ethical AI use, and 2) utilize e-learning platforms to offer flexible training options, enabling staff to learn at their own pace and according to their schedules.

Third, developing trust in AI requires transparent communication about the reliability and functionality of AI tools. Healthcare organizations should invest in processes that rigorously evaluate and demonstrate AI systems' effectiveness, thereby building confidence among practitioners and patients alike. The implementation plan can include 1) create interdisciplinary teams comprising clinicians, IT professionals, and ethicists to review and validate AI tools before their deployment., 2) organize open forums and discussion panels where healthcare professionals can share experiences, challenges, and successes related to AI use, promoting an open dialogue about AI's role in healthcare.

Moreover, the approach to AI training and governance should be carefully adapted to the specific context of each hospital. For instance, in hospitals with a high degree of organizational complexity, it may be beneficial to focus training efforts on understanding the functional aspects of AI technologies. Conversely, in settings with less complexity, a stronger emphasis on ethical considerations, such as AI justice, may be more pertinent. It is crucial that AI training programs are customized to meet the unique challenges and needs of different healthcare environments—be they tertiary or primary care facilities. A universal approach to AI integration is unlikely to be effective; instead, nuanced, context-aware strategies that consider the diverse AI challenges and requirements across hospital settings are essential for fostering ethical, accountable, and effective AI use in healthcare.

5.3. Limitations and future research

The current research has several limitations. ***First***, since Chinese hospitals served as the research context, it is unclear whether our results hold elsewhere. For instance, our findings challenge prevailing assumptions by showing that a heightened emphasis on privacy does not necessarily augment felt accountability in public Chinese hospitals. This counterintuitive result prompts a reevaluation of privacy's role in AI frameworks, suggesting that perceptions of accountability may be shaped by a complex interplay of factors including the degree of familiarity with AI and the specific ethical concerns of healthcare professionals in diverse settings. As different ideologies may have influenced healthcare professionals' felt accountability, further investigation can explore *how cultural, regulatory, and institutional dynamics in other healthcare systems impact the relationship between AI ethics and accountability*. Future studies could also compare the effects of privacy and security across healthcare settings with varying legal and ethical standards, potentially offering a global perspective on the generalizability of our findings. This invites further exploration through a research question: *How does the emphasis on privacy and security within AI ethical frameworks affect felt accountability among healthcare professionals across different healthcare environments globally?* This question seeks to explore the broader applicability of our findings beyond the Chinese healthcare context, considering the influence of different ideologies on healthcare professionals' perceptions of accountability towards AI.

Second, our reliance on self-reported data to investigate perceptions of AI accountability introduces subjectivity inherent to such methodologies. This consideration leads to the formulation of another research question: *How do on-site observations of healthcare professionals' interactions with AI and the tracking of AI-driven decisions, along with healthcare professionals' justifications for their choices, compare with self-reported perceptions of AI accountability?* Implementing systems that monitor AI decisions and healthcare professionals' responses could offer a more nuanced understanding of how ethical considerations are integrated into AI interactions in real-time.

Lastly, considering the wide range of AI applications in healthcare, from administrative tasks to complex clinical decisions, the degree of decision complexity is likely a significant factor influencing felt accountability. This observation prompts the development of a possible research question: *How does the complexity of AI-assisted decisions affect healthcare professionals' felt accountability, and how do these perceptions change in simulated scenarios of varying decision complexity?* Designing controlled experiments to simulate AI interactions in scenarios of differing complexity could reveal valuable insights into the dynamics of accountability in healthcare settings.

Compliance with ethical standards

- This article does not contain any studies involving animals performed by any of the authors.
- All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional research committee. Informed consent was obtained from all individual participants involved in the study.

Funding

This work was supported by the National Natural Science Foundation of China (Grants number 72372111).

CRediT authorship contribution statement

Weisha Wang: Writing – original draft, Resources, Data curation, Conceptualization. **Yichuan Wang:** Writing – original draft, Supervision, Methodology, Conceptualization. **Long Chen:** Writing – original draft, Visualization, Formal analysis. **Rui Ma:** Writing – review & editing, Validation. **Minhao Zhang:** Writing – review & editing, Visualization, Supervision, Resources.

Data availability

Data will be made available on request.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.socscimed.2024.116717.

References

Ananny, M., Crawford, K., 2018. Seeing without knowing: limitations of the transparency ideal and its application to algorithmic accountability. New Media Soc. 20 (3), 973–989.

Baird, A., Maruping, L.M., 2021. The next generation of research on IS use: a theoretical framework of delegation to and from agentic IS artifacts. MIS Q. 45 (1), 315–341.

Baumard, N., André, J.B., Sperber, D., 2013. A mutualistic approach to morality: the evolution of fairness by partner choice. Behav. Brain Sci. 36 (1), 59–78.

Beauchamp, T.L., Childress, J.F., 2001. Principles of Biomedical Ethics. Oxford University Press, USA.

Buhmann, A., Paßmann, J., Fieseler, C., 2020. Managing algorithmic accountability: balancing reputational concerns, engagement strategies, and the potential of rational discourse. J. Bus. Ethics 163 (2), 265–280.

Cath, C., 2018. Governing artificial intelligence: ethical, legal and technical opportunities and challenges. Phil. Trans. Math. Phys. Eng. Sci. 376 (2133), 20180080.

Char, D.S., Shah, N.H., Magnus, D., 2018. Implementing machine learning in health care—addressing ethical challenges. N. Engl. J. Med. 378 (11), 981.

Dalton-Brown, S., 2020. The ethics of medical AI and the physician-patient relationship. Camb. Q. Healthc. Ethics 29 (1), 115–121.

De Togni, G., Erikainen, S., Chan, S., Cunningham-Burley, S., 2021. What makes AI 'intelligent'and 'caring'? Exploring affect and relationality across three sites of intelligence and care. Soc. Sci. Med. 277, 113874.
DeCamp, M., Tilburt, J.C., 2019. Why we cannot trust artificial intelligence in medicine. The Lancet Digital Health 1 (8), e390.
Everett, J.A., Pizarro, D.A., Crockett, M.J., 2016. Inference of trustworthiness from intuitive moral judgments. J. Exp. Psychol. Gen. 145 (6), 772.
Fiss, P.C., 2011. Building better causal theories: a fuzzy set approach to typologies in organization research. Acad. Manag. J. 54 (2), 393–420.
Floridi, L., 2021. The European Legislation on AI: a brief analysis of its philosophical approach. Philosophy & Technology 34 (2), 215–222.
Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., Luetge, C., Madelin, R., Pagallo, U., Rossi, F., 2018. AI4People—an ethical framework for a good AI society: opportunities, risks, principles, and recommendations. Minds Mach. 28 (4), 689–707.
Fügener, A., Grahl, J., Gupta, A., Ketter, W., 2022. Cognitive challenges in human–artificial intelligence collaboration: investigating the path toward productive delegation. Inf. Syst. Res. 33 (2), 678–696.
Furnari, S., Crilly, D., Misangyi, V.F., Greckhamer, T., Fiss, P.C., Aguilera, R.V., 2021. Capturing causal complexity: heuristics for configurational theorizing. Acad. Manag. Rev. 46 (4), 778–799.
Gillner, S., 2024. We're implementing AI now, so why not ask us what to do?–How AI providers perceive and navigate the spread of diagnostic AI in complex healthcare systems. Soc. Sci. Med. 340, 116442.
Greene, D., Hoffmann, A.L., Stark, L., 2019. Better, nicer, clearer, fairer: a critical assessment of the movement for ethical artificial intelligence and machine learning. In: Proceedings of the 52nd Hawaii International Conference on System Sciences.
Hall, A.T., Ferris, G.R., 2011. Accountability and extra-role behavior. Empl. Responsib. Rights J. 23 (2), 131–144.
Hall, A.T., Frink, D.D., Buckley, M.R., 2017. An accountability account: a review and synthesis of the theoretical and empirical research on felt accountability. J. Organ. Behav. 38 (2), 204–224.
Hall, A.T., Zinko, R., Perryman, A.A., Ferris, G.R., 2009. Organizational citizenship behavior and reputation: mediators in the relationships between accountability and job performance and satisfaction. J. Leader. Organ Stud. 15 (4), 381–392.
Hao, K., 2019. 2020, let's stop AI ethics-washing and actually do something. MIT Technoloyg Review. Avaialble at: https://www.technologyreview.com/2019/12/27/57/ai-ethics-washing-time-to-act/#:~:text=We%27re%20falling%20into%20a,members%20whose%20inclusion%20provoked%20controversy.
Hathaliya, J.J., Tanwar, S., 2020. An exhaustive survey on security and privacy issues in Healthcare 4.0. Comput. Commun. 153, 311–335.
Hatherley, J.J., 2020. Limits of trust in medical AI. J. Med. Ethics 46 (7), 478–481.
Jha, S., 2020. Can you sue an algorithm for malpractice? It depends. Stat News. Avaiable at: https://www.statnews.com/2020/03/09/can-you-sue-artificial-intelligence-algorithm-for-malpractice/.
Jobin, A., Ienca, M., Vayena, E., 2019. The global landscape of AI ethics guidelines. Nat. Mach. Intell. 1 (9), 389–399.
Johnson, D.G., 2015. Technology with no human responsibility? J. Bus. Ethics 127 (4), 707–715.
Kui, S., 2021. The stumbling balance between public health and privacy amid the pandemic in China. The Chinese Journal of Comparative Law 9 (1), 25–50.
Laney, D.B., 2023. AI ethics essentials: lawsuit over AI denial of healthcare. Forbes. Available at: https://www.forbes.com/sites/douglaslaney/2023/11/16/ai-ethics-essentials-lawsuit-over-ai-denial-of-healthcare/?sh=415b7ff13ac6.
Lebovitz, S., Levina, N., Lifshitz-Assaf, H., 2021. Is AI ground truth really true? The dangers of training and evaluating AI tools based on experts' know-what. MIS Q. 45 (3), 1501–1525.
Lebovitz, S., Lifshitz-Assaf, H., Levina, N., 2022. To engage or not to engage with AI for critical judgments: how professionals deal with opacity when using AI for medical diagnosis. Organ. Sci. 33 (1), 126–148.
Martin, K., 2019. Designing ethical algorithms. MIS Q. Exec. 18 (2), 129–142.
Martin, K., Waldman, A., 2022. Are algorithmic decisions legitimate? The effect of process and outcomes on perceptions of legitimacy of AI decisions. J. Bus. Ethics. https://doi.org/10.1007/s10551-021-05032-7.
Mayer, R.C., Davis, J.H., Schoorman, F.D., 1995. An integrative model of organizational trust. Acad. Manag. Rev. 20 (3), 709–734.
Mcknight, D.H., Carter, M., Thatcher, J.B., Clay, P.F., 2011. Trust in a specific technology: an investigation of its components and measures. ACM Transactions on Management Information Systems 2 (2), 1–25.
Mero, N.P., Guidice, R.M., Werner, S., 2014. A field study of the antecedents and performance consequences of perceived accountability. J. Manag. 40 (6), 1627–1652.
Mikalef, P., Conboy, K., Lundström, J.E., Popovič, A., 2022. Thinking responsibly about responsible AI and 'the dark side'of AI. Eur. J. Inf. Syst. 31 (3), 257–268.
Miller, D.D., 2020. Machine intelligence in cardiovascular medicine. Cardiol. Rev. 28 (2), 53–64.
Mittelstadt, B.D., Allo, P., Taddeo, M., Wachter, S., Floridi, L., 2016. The ethics of algorithms: mapping the debate. Big Data & Society 3 (2), 2053951716679679.
Morley, J., Floridi, L., 2020. An ethically mindful approach to AI for health care. Lancet 395 (10220), 254–255.
Morley, J., Machado, C.C., Burr, C., Cowls, J., Joshi, I., Taddeo, M., Floridi, L., 2020. The ethics of AI in health care: a mapping review. Soc. Sci. Med., 113172
Munoko, I., Brown-Liburd, H.L., Vasarhelyi, M., 2020. The ethical implications of using artificial intelligence in auditing. J. Bus. Ethics 167 (2), 209–234.
Newman, D.T., Fast, N.J., Harmon, D.J., 2020. When eliminating bias isn't fair: algorithmic reductionism and procedural justice in human resource decisions. Organ. Behav. Hum. Decis. Process. 160, 149–167.
Pallardy, C., 2023. How AI ethics are being shaped in health care today. Information week. Available at: https://www.informationweek.com/machine-learning-ai/how-ai-ethics-are-being-shaped-in-health-care-today-#close-modal.
Pappas, I.O., Woodside, A.G., 2021. Fuzzy-set qualitative comparative analysis (fsQCA): guidelines for research practice in information systems and marketing. Int. J. Inf. Manag. 58, 102310.
Park, Y., El Sawy, O.A., Fiss, P., 2017. The role of business intelligence and communication technologies in organizational agility: a configurational approach. J. Assoc. Inf. Syst. Online 18 (9), 1.
Park, Y., Fiss, P.C., El Sawy, O.A., 2020. Theorizing the multiplicity of digital phenomena: the ecology of configurations, causal recipes, and guidelines for applying QCA. MIS Q. 44 (4), 1493–1520.
Podsakoff, P.M., MacKenzie, S.B., Lee, J.Y., Podsakoff, N.P., 2003. Common method biases in behavioral research: a critical review of the literature and recommended remedies. J. Appl. Psychol. 88 (5), 879–903.
Porter, Z., Zimmermann, A., Morgan, P., McDermid, J., Lawton, T., Habli, I., 2022. Distinguishing two features of accountability for AI technologies. Nat. Mach. Intell. 4 (9), 734–736.
Price, W.N., Gerke, S., Cohen, I.G., 2019. Potential liability for physicians using artificial intelligence. JAMA 322 (18), 1765–1766.
Ragin, C.C., 2008. Redesigning Social Inquiry. University of Chicago Press.
Ragin, C.C., 2014. The Comparative Method. University of California Press.
Ragin, C.C., Strand, S.I., Rubinson, C., 2008. User's Guide to Fuzzy-Set/qualitative Comparative Analysis, vol. 87. University of Arizona, pp. 1–87.
Rajpurkar, P., Chen, E., Banerjee, O., Topol, E.J., 2022. AI in health and medicine. Nat. Med. 28 (1), 31–38.
Rana, N.P., Chatterjee, S., Dwivedi, Y.K., Akter, S., 2022. Understanding dark side of artificial intelligence (AI) integrated business analytics: assessing firm's operational inefficiency and competitiveness. Eur. J. Inf. Syst. 31 (3), 364–387.
Reddy, S., Allan, S., Coghlan, S., Cooper, P., 2020. A governance model for the application of AI in health care. J. Am. Med. Inf. Assoc. 27 (3), 491–497.
Rihoux, B., Ragin, C.C., 2008. Configurational Comparative Methods: Qualitative Comparative Analysis (QCA) and Related Techniques. Sage Publications.
Saenz, A.D., Harned, Z., Banerjee, O., Abràmoff, M.D., Rajpurkar, P., 2023. Autonomous AI systems in the face of liability, regulations and costs. NPJ Digital Medicine 6 (1), 185.
Schneider, C.Q., Wagemann, C., 2012. Set-Theoretic Methods for the Social Sciences: A Guide to Qualitative Comparative Analysis. Cambridge University Press, New York.
Schneider, C.Q., Wagemann, C., 2010. Standards of good practice in qualitative comparative analysis (QCA) and fuzzy-sets. Comp. Sociol. 9 (3), 397–418.
Shneiderman, B., 2020. Bridging the gap between ethics and practice: guidelines for reliable, safe, and trustworthy human-centered AI systems. ACM Transactions on Interactive Intelligent Systems 10 (4), 1–31.
Siala, H., Wang, Y., 2022. SHIFTing artificial intelligence to be responsible in healthcare: a systematic review. Soc. Sci. Med. 296, 114782.
Solomonides, A.E., Koski, E., Atabaki, S.M., Weinberg, S., McGreevey III, J.D., Kannry, J. L., et al., 2022. Defining AMIA's artificial intelligence principles. J. Am. Med. Inf. Assoc. 29 (4), 585–591.
Tam, L., 2018. Why Privacy Is an Alien Concept in Chinese Culture. https://www.scmp.com/news/hong-kong/article/2139946/why-privacy-alien-concept-chinese-culture. (Accessed 11 October 2021).
Tarafdar, M., Page, X., Marabelli, M., 2023. Algorithms as co-workers: human algorithm role interactions in algorithmic work. Inf. Syst. J. 33 (2), 232–267.
Thiebes, S., Lins, S., Sunyaev, A., 2021. Trustworthy artificial intelligence. Electron. Mark. 31 (2), 447–464.
Tseng, L.M., 2019. How implicit ethics institutionalization affects ethical selling intention: the case of Taiwan's life insurance salespeople. J. Bus. Ethics 158 (3), 727–742.
Wikhamn, W., Hall, A.T., 2014. Accountability and satisfaction: organizational support as a moderator. J. Manag. Psychol. 29 (5), 458–471.
Woodside, A.G., 2013. Moving beyond multiple regression analysis to algorithms: Calling for adoption of a paradigm shift from symmetric to asymmetric thinking in data analysis and crafting theory. J. Business Res. 66 (4), 463–472.