LINK MEDICAL JOURNAL
of Health and Community Research | LMJHCR
Volume 4, Issue 1 | June 2026 | ISSN 3007-3448

Indexed in DOAJ | CC Bq4jc7c36Y 4.0
DOI: 10.61919/j1617066

Original Article

Conditional Trust and Shared Responsibility in AI-Assisted Brain Tumour Surgery: A Qualitative Study of Patients, Family Surrogates, and Neurosurgeons

Hua Feng[1], Linda Chiuman[1], Dewi Sartika[1]

[1] Universitas Prima Indonesia, Indonesia

***Corresponding author: Linda Chiuman, lindachiuman@unprimdn.ac.id**

"Cite this Article" | **Received**: 21 March 2026; **Accepted**: 17 June 2026; **Published**: 30 June 2026.

ABSTRACT

Background: Artificial intelligence (AI) is increasingly applied in neuro-oncology for imaging, prognostic modelling, functional-risk estimation, and surgical planning. Its clinical acceptability depends not only on predictive performance but also on how stakeholders interpret reliability, uncertainty, communication, and responsibility. Objective: To explore how patients, family surrogates, and neurosurgeons understood AI reliability, communication, uncertainty, responsibility, family involvement, and decision-making confidence in AI-assisted brain tumour surgery. Methods: A qualitative interview study was conducted in tertiary neurosurgical and neuro-oncology settings in Indonesia. Maximum-variation purposive sampling included 25 participants: 10 patients, 8 family surrogates, and 7 neurosurgeons. Semi-structured interviews explored perceived AI reliability, clinician–AI agreement or disagreement, uncertainty, family involvement, confidence, and acceptable safeguards. Interviews lasted 45–75 minutes, were audio-recorded, transcribed verbatim, checked, and de-identified. Data were analysed using reflexive thematic analysis with within-group, cross-group, and negative-case comparison. Results: Five themes emerged: AI as a second opinion rather than an autonomous authority; communication as translation of probabilities into meaningful consequences; confidence despite acknowledged uncertainty; relational responsibility involving patients, families, and clinicians; and governance conditions including local applicability, fairness, accountability, auditability, and clinician override. Trust was conditional rather than absolute, and transparently explained clinician disagreement with AI could strengthen confidence. Conclusion: AI-assisted brain tumour surgery was considered most acceptable when AI supplemented rather than replaced human judgement. Trust depended on understandable communication, acknowledged uncertainty, patient values, negotiated family involvement, clinician scrutiny, and visible accountability. Keywords: artificial intelligence; brain tumour surgery; qualitative research; shared decision-making; uncertainty; trust; family surrogates

EDITORIAL INFORMATION

Author Contributions: HF: Concept, Design; LC: Data Collection, Analysis, Interpretation; DS: Drafting, Critical Revision, Supervision.
Ethical Approval: Universitas Prima Indonesia
Informed Consent: Written informed consent was obtained from all participants
Conflict of Interest: The authors declare no conflict of interest; **Funding**: No external funding; **Data Availability**: Available from the corresponding author on reasonable request; **Acknowledgments**: N/A.

INTRODUCTION

Surgical decision-making for brain tumours involves complex trade-offs among disease control, neurological function, cognition, communication, independence, and quality of life. In selected gliomas, greater extent of resection may improve disease control; however, aggressive resection near eloquent neural structures can increase the risk of irreversible functional impairment (1). Consequently, the most technically extensive surgical option may not necessarily represent the most acceptable option for an individual patient. Patients may prioritize preservation of speech, mobility, autonomy, employment, caregiving responsibilities, or avoidance of severe disability differently from clinicians or family members. These decisions are further complicated by prognostic uncertainty, time pressure, emotional distress, and, in some patients, cognitive or neurological impairment that may affect participation in decision-making.

Artificial intelligence (AI) is increasingly being investigated in neuro-oncology for applications including image segmentation, molecular classification, prognostic modelling, functional-risk estimation, and operative planning. By integrating imaging, clinical, genomic, and functional information, AI-based systems

have the potential to complement clinical assessment and provide additional decision-support information (2). Predictive performance alone, however, does not establish clinical usefulness or justify reliance on an AI-generated output. Models developed retrospectively or from restricted populations may perform differently when applied to patients whose demographic, clinical, imaging, or institutional characteristics differ from those represented in the development data. Dataset shift, incomplete external validation, data-quality limitations, workflow mismatch, and poor transportability may therefore reduce the reliability of apparently accurate systems when deployed in clinical practice (3,4). In high-stakes neurosurgical decisions, an algorithmic prediction may consequently be informative without being determinative, and the presentation of an output as authoritative may encourage reliance beyond the system's intended role (3).

Available neurosurgical evidence suggests that patients, relatives, and surgical professionals generally view AI more favourably when it supports rather than replaces clinical judgement (5–7). Concerns described in this literature include diagnostic or prognostic error, bias, experimental use, loss of hope, inappropriate dependence on technology, and the possibility that clinicians may become less critical when algorithmic recommendations appear convincing. These concerns are consistent with broader theories of trust in automation, which conceptualize safe use as appropriately calibrated reliance rather than unconditional acceptance of an automated system (8,9). In this context, trust in clinical AI depends not only on whether an output appears accurate, but also on whether its limitations can be recognized, its recommendations can be questioned, and human clinicians retain the capacity and responsibility to override it.

Communication is particularly important because numerical predictions do not become clinically meaningful merely by being disclosed. Probabilities must be interpreted in relation to the severity, duration, reversibility, and everyday consequences of possible outcomes. Risk estimates may be misunderstood when uncertainty is concealed, probabilities are presented without appropriate context, or distinctions between an algorithmic prediction and a clinician's recommendation are blurred (10,11). Shared decision-making therefore requires more than transmission of technical information. It involves explanation of available options and trade-offs, elicitation of patient values and preferences, acknowledgement of uncertainty, emotional responsiveness, and support for meaningful participation in the final decision (12,13). These processes may be particularly complex in brain tumour care, where family members can help patients retain information, interpret technical language, articulate established values, anticipate caregiving demands, and participate in decisions when decisional capacity is impaired. At the same time, family involvement may create tension when relatives' preferences differ from those of the patient (9).

A further conceptual challenge is the relationship between uncertainty and decision-making confidence. In brain tumour surgery, complete certainty regarding neurological or oncological outcomes is rarely attainable. Decision-support frameworks therefore distinguish certainty about an eventual outcome from the quality of the decision-making process itself: a decision may remain uncertain while still being informed, transparent, deliberative, and aligned with what matters to the patient (14,15). This distinction is especially relevant when AI-generated probabilities are introduced into emotionally and clinically complex discussions. Confidence may depend less on elimination of uncertainty than on whether uncertainty is explicitly acknowledged, translated into understandable consequences, and incorporated into a defensible, shared decision (10).

Despite increasing interest in AI-supported neurosurgical care, limited qualitative evidence has examined how patients, family surrogates, and neurosurgeons jointly interpret reliability, uncertainty, communication, responsibility, and confidence when AI-derived information enters brain tumour surgical decision-making. Understanding these perspectives is important because the legitimacy of clinical AI may depend as much on how its outputs are interpreted, contested, and incorporated into human relationships as on technical performance. This study therefore explored how patients, family surrogates, and neurosurgeons understood AI reliability, communication, uncertainty, responsibility, family involvement, and decision-making confidence in the context of AI-assisted brain tumour surgical decisions.

MATERIALS AND METHODS

Study Design and Setting

A qualitative interview study was conducted to explore how patients, family surrogates, and neurosurgeons understood the use of artificial intelligence in decision-making for adult brain tumour surgery. The study was situated within tertiary neurosurgical and neuro-oncology care in Indonesia. AI was considered as a clinical decision-support resource rather than an autonomous decision-maker, with particular attention to how AI-generated information was interpreted within interactions among patients, family members, and neurosurgeons. Reporting was guided by the Consolidated Criteria for Reporting Qualitative Research (COREQ) (16).

Participants and Sampling

Maximum-variation purposive sampling was used to recruit 25 participants representing three interdependent stakeholder groups: 10 adult patients, 8 family surrogates or primary caregivers, and 7 neurosurgeons. Sampling sought variation in stakeholder role, age, gender, hospital context, perceived reliability of AI, tolerance of uncertainty, confidence in decision-making, and direct or scenario-based exposure to AI-assisted surgical decision-making. The purpose of the sampling strategy was to obtain information-rich accounts that allowed examination of convergence, divergence, and negative cases across stakeholder perspectives rather than to estimate the frequency of particular views.

Eligible patients were adults who had either participated in a recent decision concerning surgery for a brain tumour or considered a standardized scenario involving AI-supported surgical advice. Family surrogates were individuals who had participated in surgical decision-making for an adult patient. Neurosurgeons were clinicians involved in brain tumour resection or neuro-oncological surgical decision-making. Because the study included both direct experience and scenario-based consideration of AI-assisted decision-making, these forms of exposure were treated as relevant contextual characteristics during interpretation rather than assumed to represent identical experiences.

Data Collection

Three role-adapted semi-structured interview guides were used for patients, family surrogates, and neurosurgeons. The guides explored participants' understanding of AI, perceptions of its reliability, responses to agreement or disagreement between AI outputs and neurosurgical judgement, experiences of uncertainty, family involvement, meanings of decision-making confidence, and views regarding safeguards required for acceptable clinical use. Questions were designed to allow participants to describe both supportive and critical perspectives on AI rather than presupposing that AI was either beneficial or harmful.

Interviews lasted approximately 45–75 minutes and were conducted in private or otherwise secure settings. With participants' permission, interviews were audio-recorded. The interviewer did not advocate for or against the use of AI during the interviews. Recordings were transcribed verbatim, transcripts were checked against the corresponding audio recordings, and identifying information was removed before analysis. Participants were identified in the analytic material by stakeholder group and participant identification code.

Data Analysis and Reflexivity

Data were analysed using reflexive thematic analysis, following processes of familiarisation, coding, development of candidate themes, review and refinement of themes, theme definition, and reporting (17,18). Analysis combined deductive attention to concepts central to the research question—including perceived reliability, communication, decisional uncertainty, and confidence—with openness to additional meanings arising from participants' accounts, including responsibility, hope, family authority, local validation, fairness, accountability, and clinician override. Both semantic and latent interpretations were considered in developing the final thematic structure.

Analysis proceeded at both within-group and cross-group levels. Within-group analysis examined how patients, family surrogates, and neurosurgeons understood AI-assisted decision-making from their

respective positions. Cross-group analysis was then used to identify areas of convergence, complementarity, and disagreement among the three stakeholder groups. Particular attention was given to negative or disconfirming cases that complicated dominant interpretations, including accounts in which disagreement between a clinician and an AI output increased rather than decreased trust, and accounts in which participants expressed confidence in a decision despite continuing uncertainty regarding its outcome. Direct and scenario-based accounts were considered in relation to their different experiential contexts during interpretation.

A second researcher scrutinised the interview material and developing analysis. Differences in interpretation were used to promote reflexive examination and refinement of the analysis rather than to calculate mechanical intercoder agreement or impose consensus as a criterion of validity. Reflexive memos, a coding log, de-identification procedures, attention to negative cases, use of participant quotations, and maintenance of an audit trail were used to support transparency and trustworthiness of the analytic process.

Ethical Considerations

Ethical and institutional permissions were obtained through Universitas Prima Indonesia and the participating clinical settings. Participants provided written or approved electronic informed consent before participation. Interview recordings and transcripts were handled as confidential research material, identifying information was removed from transcripts and quotations, and participant identifiers were used when reporting illustrative quotations.

RESULTS

Participant Overview and Thematic Structure

The analysis included 25 participants comprising 10 patients, 8 family surrogates or primary caregivers, and 7 neurosurgeons. The three stakeholder groups occupied different but interdependent positions within brain tumour surgical decision-making. Patients primarily considered the personal consequences of surgery, including neurological function, independence, quality of life, and willingness to accept risk. Family surrogates frequently interpreted information, recalled patient preferences, anticipated caregiving demands, and considered the practical consequences of possible disability. Neurosurgeons focused more strongly on technical validity, clinical applicability, workflow integration, professional accountability, and the need to retain authority to question or override AI-generated outputs.

Within- and cross-group analysis generated five interconnected themes: (1) conditional legitimacy of AI as a second opinion rather than an autonomous authority; (2) communication as the translation of numerical outputs into clinically meaningful information; (3) decision-making confidence despite acknowledged uncertainty; (4) relational responsibility involving patients, families, and clinicians; and (5) governance conditions required for acceptable reliance on AI. Across these themes, participants generally supported the use of AI when it supplemented rather than displaced human judgement, although the basis and meaning of trust differed across stakeholder groups (11-17).

Theme 1: AI Gained Legitimacy as a Second Opinion Rather Than an Autonomous Decision-Maker

Participants did not describe AI reliability as an absolute property of the technology. Instead, legitimacy depended on whether AI contributed additional information while remaining open to clinical scrutiny. Patients and family surrogates commonly valued AI when it reduced dependence on a single unaided opinion but did not replace the clinician responsible for the decision. One patient explained:

> *"The computer result made me less afraid that one doctor was guessing alone. It felt like another opinion, but not the person who should decide." (P04, patient)*

Trust was strengthened when AI outputs could be challenged rather than simply accepted. A family surrogate described clinician disagreement with the system as reassuring rather than destabilising:

> *"I trusted it more when the neurosurgeon disagreed with one part and explained the reason. That showed the machine was being checked, not obeyed." (S02, family surrogate)*

For neurosurgeons, reliability depended particularly on applicability to the individual patient. They questioned whether model performance derived from one population, imaging environment, or data structure could be assumed to apply directly to another clinical context. One neurosurgeon stated:

> *"A model can be accurate and still be wrong for the person in front of us. I need to know the training population, the missing variables, and whether the output changes when the scan quality changes." (N03, neurosurgeon)*

Across groups, therefore, perceived reliability depended on transparent limitations, clinical verification, individual relevance, and the continued possibility of human override. A negative case within this theme was that clinician disagreement with AI did not necessarily weaken trust; when the disagreement was explained, it could demonstrate active supervision and strengthen confidence in the decision-making process. Participants also recognised that institutional prestige or a convincing interface could create inappropriate over-reliance even when the limitations of the underlying model were poorly understood.

Theme 2: Communication Transformed Numerical Outputs Into Clinically and Emotionally Usable Knowledge

Participants consistently indicated that AI-generated probabilities were not sufficient in themselves to support a surgical decision. Numerical outputs became meaningful when clinicians translated them into functional consequences, possible alternatives, evidence strength, and implications for daily life. Patients particularly valued explanations that connected risk estimates with outcomes such as speech, movement, independence, employment, self-care, and the likelihood or duration of recovery (18).

One patient described why a percentage alone was insufficient:

> *"The percentage only became useful when the doctor explained what weakness could mean in daily life, whether it might improve, and what the alternative operation would involve." (P02, patient)*

Communication was therefore valued not because it necessarily increased reassurance, but because it improved understanding. In some accounts, explanation of uncertainty reduced the apparent certainty of an AI prediction while simultaneously increasing trust in the consultation. One patient stated:

> *"At first I heard seventy percent and thought it was almost guaranteed. When the surgeon explained the uncertainty around that number, I was more cautious but also more confident in the discussion." (P07, patient)*

Neurosurgeons emphasised the importance of distinguishing three elements during communication: the model's prediction, the strength and limitations of the evidence supporting that prediction, and the clinician's own recommendation. One neurosurgeon explained:

> *"I have to explain three things separately: what the model predicts, how strong the evidence is, and what I recommend. If those are blended together, patients can think the algorithm has already decided." (N05, neurosurgeon)*

The groups therefore converged on the importance of understandable communication, but their emphasis differed. Patients focused on personal and functional meaning, family surrogates on interpretation and practical consequences, and neurosurgeons on maintaining a clear distinction between prediction, evidence, and professional recommendation.

Theme 3: Decision-Making Confidence Could Coexist With Acknowledged Uncertainty

Participants distinguished uncertainty about what would ultimately happen from confidence in how a decision had been reached. Brain tumour surgery was understood as an inherently uncertain context in which neither clinicians nor predictive systems could provide guarantees. Confidence was therefore associated with whether available evidence had been considered, important unknowns had been disclosed, patient priorities had been discussed, and responsibility for the final plan remained clear.

One patient described continuing uncertainty regarding the outcome while expressing confidence in the transparency of the decision process:

> *"The model gave a range, not an answer. I remained worried, but knowing the range helped me accept that no one was hiding the difficult part from me." (P06, patient)*

Neurosurgeons similarly described professional confidence as preparedness for uncertainty rather than certainty of outcome. One stated:

> *"A confident surgical plan is not a promise. It is a plan with known assumptions, known limits, and a clear response if the intraoperative situation differs from the prediction." (N02, neurosurgeon)*

This theme also contained an important negative interpretation. High stated confidence did not always indicate that uncertainty had been appropriately understood. Participants' apparent confidence could potentially reflect deference to medical authority, resignation, optimism, or misunderstanding. The distinction that emerged from the analysis was therefore not between uncertainty and confidence as opposites, but between confidence grounded in an informed and transparent decision process and confidence that might arise without adequate understanding of the trade-offs involved.

Theme 4: Family Surrogates Interpreted Information, Protected Patient Values, and Shared Responsibility

Family participation formed a relational component of decision-making rather than a simple extension of individual patient choice. Surrogates frequently helped interpret technical information, recall questions, manage emotional responses, anticipate caregiving requirements, and consider how postoperative disability might affect everyday life. Their involvement could support deliberation, but it could also create pressure when family priorities differed from those of the patient.

One patient described a conflict between survival-oriented family preferences and personal concerns about functional loss:

> *"My children wanted the operation that offered the longest survival. I was more afraid of losing speech. The useful discussion was when those two goals were placed side by side." (P09, patient)*

Family surrogates also introduced practical considerations that were not necessarily represented in predictive outputs. One surrogate explained:

> *"I was not only choosing an operation. I was also thinking about whether we could manage rehabilitation, travel, and care at home if he lost movement." (S04, family surrogate)*

Neurosurgeons emphasised that family involvement required attention to the patient's own preferred level of participation and authority. One neurosurgeon stated:

> *"When the family spoke more than the patient, I had to check whether that was the patient's preference or whether fear and hierarchy had taken over the consultation." (N04, neurosurgeon)*

The meaning of confidence consequently varied by role. For patients, it was closely related to willingness to proceed with a decision they understood and could accept. For family surrogates, confidence involved protecting patient values while considering caregiving realities. For neurosurgeons, confidence was more closely associated with clinical defensibility and preservation of patient preference. The findings therefore suggested that family-centred decision-making was most acceptable when family participation was negotiated rather than assumed.

Theme 5: Trust Depended on Accountability, Local Validation, Fairness, and Clinician Override

The final theme extended beyond the consultation itself to the organisational and governance conditions under which participants considered AI trustworthy. Participants questioned whether models had been validated in populations similar to the patients for whom they were being used, how incomplete or poor-quality data would affect predictions, whether clinicians could reject an output, and who would remain responsible when AI-informed advice contributed to harm.

A patient expressed concern about population relevance:

> *"I would ask whether people like me were included in the data. If the model was trained somewhere very different, the number might look exact but not belong to my case." (P08, patient)*

Institutional reputation increased initial confidence for some participants but was not considered sufficient evidence of safety. A family surrogate stated:

> *"The hospital logo made the system seem safe, but I still wanted to know whether it had been tested with Indonesian patients and whether the surgeon could ignore it." (S02, family surrogate)*

Neurosurgeons placed particular emphasis on traceability and post-implementation learning. One explained:

> *"The model should record when I accept or reject its recommendation and why. Otherwise, we cannot learn whether it is helping or simply shaping practice without being noticed." (N03, neurosurgeon)*

Across stakeholder groups, acceptable reliance on AI was therefore associated with local applicability, explicit communication of uncertainty, clinical override, accountability, and the ability to review how AI influenced decisions over time. These governance conditions were interconnected: weaknesses in one

area could reduce confidence in otherwise technically useful systems or encourage inappropriate dependence on their outputs.

DISCUSSION

This qualitative study explored how patients, family surrogates, and neurosurgeons understood reliability, communication, uncertainty, responsibility, and decision-making confidence when AI-generated information was introduced into brain tumour surgical decisions. Across the three stakeholder groups, AI was regarded as most legitimate when it functioned as a contestable source of additional evidence rather than an autonomous authority. Trust was not described as unconditional confidence in technological performance; instead, it depended on the perceived relevance of the output to the individual patient, transparent communication of limitations, clinician scrutiny, and preservation of human responsibility. The findings further indicated that uncertainty and decision-making confidence were not necessarily opposing states. Participants could remain uncertain about surgical outcomes while expressing confidence in a decision process that they considered transparent, values-sensitive, and defensible.

The finding that participants preferred AI in a supportive rather than autonomous role is consistent with existing neurosurgical literature describing greater acceptance of systems that augment clinical judgement than systems perceived to replace it (5–7). The present findings extend this observation by suggesting that active clinician scrutiny may itself contribute to trust. In particular, clinician disagreement with an AI output was not uniformly interpreted as evidence that the technology had failed. When the reason for disagreement was explained, participants could interpret the clinician's response as evidence that the system remained subject to human evaluation. This suggests that the perceived legitimacy of AI may depend not only on agreement between human and algorithmic judgement but also on whether disagreement can be made visible, explained, and resolved within a clinically accountable process. In this sense, contestability emerged as an important component of calibrated trust, consistent with broader accounts of appropriate reliance on automation (8,9).

Communication was central to how participants evaluated AI-supported information. Numerical predictions were rarely regarded as sufficient without interpretation of their likely consequences for neurological function, independence, rehabilitation, work, family roles, and quality of life. This finding is consistent with evidence showing that probability information may be misunderstood when it is presented without adequate context or explanation of uncertainty (10,11). In the present study, clearer explanation sometimes reduced the apparent certainty of a prediction while increasing confidence in the consultation. This distinction is clinically important because effective communication should not be evaluated by whether it reassures a patient or increases acceptance of a proposed intervention. Its value lies in whether it enables participants to understand uncertainty, compare alternatives, consider trade-offs, and relate available evidence to personal priorities.

The separation of model prediction, evidential strength, and clinician recommendation was particularly important in neurosurgeons' accounts. When these elements are presented as if they constitute a single recommendation, patients may attribute decision-making authority to the algorithm that the system was not intended to possess. The findings therefore support a communication approach in which AI-derived predictions are explicitly identified as one component of a broader clinical assessment. This interpretation is consistent with shared decision-making models that emphasize explanation of options, discussion of benefits and harms, elicitation of patient preferences, and deliberation before a final recommendation is made (12,13). The present study does not establish that any specific communication strategy improves clinical outcomes; rather, participants' accounts indicate that these distinctions contributed to whether AI-assisted discussions were perceived as understandable and trustworthy.

A central interpretive contribution of the study was the distinction between outcome uncertainty and confidence in the decision-making process. Decision-support literature has often associated unresolved uncertainty with decisional conflict (14,15), yet the current findings suggest that acknowledged uncertainty does not necessarily prevent confidence. Participants described confidence when important unknowns were visible, the assumptions underlying a recommendation were explained, alternatives were considered,

and the final plan could be justified in relation to patient priorities. Confidence therefore appeared to concern the quality and defensibility of the decision process rather than certainty about the eventual surgical outcome. At the same time, the analysis cautions against treating expressed confidence as inherently desirable. Apparent confidence may also arise from deference, resignation, optimism, or incomplete understanding. Clinicians may therefore need to assess not only whether a patient reports being confident, but whether the patient can explain the principal trade-offs, uncertainties, and reasons underlying the preferred option.

Family involvement shaped each of the major themes. Surrogates contributed to interpretation of information, preservation of patient values, anticipation of postoperative caregiving demands, and consideration of whether potential outcomes would be manageable in the patient's social context. Previous cancer research has similarly described family participation in treatment decisions as multidimensional and requiring negotiation rather than assumption (19). The present study adds that AI-derived information may intensify this relational dimension because numerical outputs can be interpreted differently by patients, relatives, and clinicians. Families may increase the practical relevance of a decision by identifying rehabilitation, travel, financial, or caregiving implications that are not incorporated into predictive models. Conversely, family participation may become problematic when relatives' preferences displace the patient's own priorities. The findings therefore support explicitly asking patients how they wish family members to participate and distinguishing supportive family involvement from substitution of family preferences for patient preferences.

The governance theme indicates that participants' concerns extended beyond explainability at the level of an individual consultation. Patients, relatives, and neurosurgeons also questioned whether models had been evaluated in populations comparable with those in which they were being used, whether data-quality limitations would be visible, whether clinicians could override the output, and whether the influence of AI on clinical decisions could subsequently be audited. These concerns are consistent with broader literature describing the effects of dataset shift, limited transportability, workflow mismatch, biased proxies, and inappropriate reliance on clinical AI systems (3,4,20). The explainable-AI literature also indicates that explanations do not uniformly increase appropriate trust; their effect depends on whether they are relevant, credible, and interpreted correctly (21,22). The present findings therefore suggest that perceived trustworthiness involves a combination of technical performance, local applicability, transparent uncertainty, human oversight, and organisational accountability.

These findings have several implications for implementation, although they should be interpreted as qualitative insights rather than evidence that particular governance interventions improve safety. When AI-supported information is introduced into brain tumour surgical discussions, clinicians may benefit from distinguishing clearly between what the model predicts, how applicable and uncertain that prediction is, and what the treating clinician recommends after considering the broader clinical context. Probability estimates may be more useful when translated into patient-relevant functional consequences and compared with realistic alternatives. Family participation may be strengthened by explicitly clarifying the role that the patient wishes relatives to play. At an institutional level, participants' concerns indicate the importance of demonstrating population relevance, preserving clinician override, identifying data-quality limitations, and maintaining mechanisms through which the influence of AI on decisions can be reviewed. These recommendations should be evaluated prospectively rather than assumed to produce safer or more acceptable care.

The study has several strengths. Inclusion of patients, family surrogates, and neurosurgeons enabled examination of the same decision-making process from three different but interdependent positions. Within-group and cross-group analysis allowed areas of convergence and divergence to be considered, while deliberate attention to negative cases prevented the analysis from assuming that greater agreement with AI necessarily produced greater trust. The use of reflexive thematic analysis also allowed examination of both explicit participant accounts and underlying meanings related to responsibility, authority, uncertainty, and legitimacy.

Several limitations should also be considered. Participants were drawn from tertiary neurosurgical contexts in Indonesia, and the findings may not transfer directly to rural services, lower-resource settings, or populations with different levels of digital access and familiarity with AI. The three stakeholder groups were unequal in size, and the meaning of terms such as trust, confidence, and responsibility varied between patients, surrogates, and neurosurgeons. Importantly, the study included both participants with direct experience of brain tumour surgical decision-making and participants considering a standardised AI-assisted scenario. Experiential and scenario-based accounts should not be regarded as equivalent, and this heterogeneity may have influenced how participants described trust and uncertainty. The study was designed to generate information-rich qualitative understanding rather than estimate the prevalence of particular views or establish universal thematic saturation. In addition, the findings represent participant perceptions and interpretations; they do not establish the objective accuracy, safety, fairness, or clinical effectiveness of any AI system. Transferability should therefore be judged in relation to the clinical, cultural, technological, and organisational context in which AI-supported neurosurgical decisions are implemented.

CONCLUSION

Patients, family surrogates, and neurosurgeons described AI-assisted brain tumour surgical decision-making as most acceptable when AI supplemented rather than displaced human judgement. Trust was conditional on the perceived relevance of the output to the individual patient, understandable communication of probabilities and uncertainty, preservation of patient values, negotiated family involvement, clinician scrutiny and override, and visible accountability for the final decision. Participants also distinguished certainty about an outcome from confidence in the decision-making process: uncertainty could remain while confidence was maintained when the available evidence, limitations, alternatives, and patient priorities were discussed transparently. These findings suggest that the clinical legitimacy of AI in brain tumour surgery may depend not only on predictive performance but also on how AI-generated information is communicated, contested, and incorporated into shared human decision-making. Prospective evaluation is needed to determine whether such approaches improve the safety, acceptability, or quality of AI-supported neurosurgical care.

REFERENCES

1. Sanai N, Berger MS. Glioma extent of resection and its impact on patient outcome. Neurosurgery. 2008;62(4):753-766. doi:10.1227/01.NEU.0000318159.21731.CA.

2. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. Nat Med. 2019;25:44-56. doi:10.1038/s41591-018-0300-7.

3. Kelly CJ, Karthikesalingam A, Suleyman M, Corrado G, King D. Key challenges for delivering clinical impact with artificial intelligence. BMC Med. 2019;17:195. doi:10.1186/s12916-019-1426-2.

4. Wiens J, Saria S, Sendak M, et al. Do no harm: a roadmap for responsible machine learning for health care. Nat Med. 2019;25:1337-1340. doi:10.1038/s41591-019-0548-6.

5. Palmisciano P, Jamjoom AAB, Taylor D, Stoyanov D, Marcus HJ. Attitudes of patients and their relatives toward artificial intelligence in neurosurgery. World Neurosurg. 2020;138:e627-e633. doi:10.1016/j.wneu.2020.03.029.

6. Layard Horsfall H, Palmisciano P, Khan DZ, et al. Attitudes of the surgical team toward artificial intelligence in neurosurgery: international two-stage cross-sectional survey. World Neurosurg. 2021;146:e724-e730. doi:10.1016/j.wneu.2020.10.171.

7. Parsons MV, Buckley O, Ebadi H, Leuthardt E, McIntosh T. Machine learning-assisted prognosis prediction and surgical decision-making for glioblastoma: perceived benefits and concerns of patients, caregivers, and neurosurgeons. Front Neurol. 2026;17:1818670. doi:10.3389/fneur.2026.1818670.

8. Lee JD, See KA. Trust in automation: designing for appropriate reliance. Hum Factors. 2004;46(1):50-80. doi:10.1518/hfes.46.1.50_30392.

9. Hoff KA, Bashir M. Trust in automation: integrating empirical evidence on factors that influence trust. Hum Factors. 2015;57(3):407-434. doi:10.1177/0018720814547570.

10. Fagerlin A, Zikmund-Fisher BJ, Ubel PA. Helping patients decide: ten steps to better risk communication. J Natl Cancer Inst. 2011;103(19):1436-1443. doi:10.1093/jnci/djr318.

11. Zipkin DA, Umscheid CA, Keating NL, et al. Evidence-based risk communication: a systematic review. Ann Intern Med. 2014;161(4):270-280. doi:10.7326/M14-0295.

12. Elwyn G, Durand MA, Song J, et al. A three-talk model for shared decision making: multistage consultation process. BMJ. 2017;359:j4891. doi:10.1136/bmj.j4891.

13. Street RL, Makoul G, Arora NK, Epstein RM. How does communication heal? Pathways linking clinician-patient communication to health outcomes. Patient Educ Couns. 2009;74(3):295-301. doi:10.1016/j.pec.2008.11.015.

14. O'Connor AM. Validation of a decisional conflict scale. Med Decis Making. 1995;15(1):25-30. doi:10.1177/0272989X9501500105.

15. Stacey D, Legare F, Lewis K, et al. Decision aids for people facing health treatment or screening decisions. Cochrane Database Syst Rev. 2017;(4):CD001431. doi:10.1002/14651858.CD001431.pub5.

16. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care. 2007;19(6):349-357. doi:10.1093/intqhc/mzm042.

17. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol. 2006;3(2):77-101. doi:10.1191/1478088706qp063oa.

18. Braun V, Clarke V. One size fits all? What counts as quality practice in reflexive thematic analysis? Qual Res Psychol. 2021;18(3):328-352. doi:10.1080/14780887.2020.1769238.

19. Laidsaar-Powell R, Butow P, Bu S, et al. Family involvement in cancer treatment decision-making: a qualitative study of patient, family, and clinician attitudes and experiences. Patient Educ Couns. 2016;99(7):1146-1155. doi:10.1016/j.pec.2016.01.014.

20. Obermeyer Z, Powers B, Vogeli C, Mullainathan S. Dissecting racial bias in an algorithm used to manage the health of populations. Science. 2019;366(6464):447-453. doi:10.1126/science.aax2342.

21. Ghassemi M, Oakden-Rayner L, Beam AL. The false hope of current approaches to explainable artificial intelligence in health care. Lancet Digit Health. 2021;3(11):e745-e750. doi:10.1016/S2589-7500(21)00208-9.

22. Rosenbacke R, Melhus A, McKee M, Stuckler D. How explainable artificial intelligence can increase or decrease clinicians' trust in AI applications in health care: systematic review. JMIR AI. 2024;3:e53207. doi:10.2196/53207.