

## BRAIN. Broad Research in Artificial Intelligence and Neuroscience

e-ISSN: 2067-3957 | p-ISSN: 2068-0473

Covered in: Web of Science (ESCI); EBSCO; JERIH PLUS (hkdir.no); IndexCopernicus; Google Scholar; SHERPA/RoMEO; ArticleReach Direct; WorldCat; CrossRef; Peeref; Bridge of Knowledge (mostwiedzy.pl); abcdindex.com; Editage; Ingenta Connect Publication; OALib; scite.ai; Scholar9; Scientific and Technical Information Portal; FID Move; ADVANCED SCIENCES INDEX (European Science Evaluation Centre, neredataltics.org); ivySCI; exaly.com; Journal Selector Tool (letpub.com); Citefactor.org; fatcat!; ZDB catalogue; Catalogue SUDOC (abes.fr); OpenAlex; Wikidata; The ISSN Portal; Socolar; KVK-Volltitel (kit.edu)

2026, Volume 17, Issue 1, pages: 283-292

Submitted: September 5<sup>th</sup>, 2025 | Accepted for publication: December 24<sup>th</sup>, 2025

### Ethics in Precision Neuro-Oncology: A Systematic Review of Empirical Evidence on Autonomy and Consent in AI-Guided Neurosurgery

#### Tomasz Tykocki

Department of Paediatric Neurosurgery,  
Children's Hospital named after prof. dr med.  
Jan Bogdanowicz in Warsaw, Poland;  
Maria Skłodowska-Curie Medical Academy,  
Warsaw, Poland.  
ttykocki@gmail.com,  
<https://orcid.org/0000-0001-6349-5301>

#### Abstract:

**Background:** Artificial intelligence (AI) is increasingly embedded in neurosurgical planning, intraoperative workflows, and risk prediction. These developments raise complex ethical concerns about autonomy, informed consent, trust, and responsibility, particularly in precision neuro-oncology, where surgical decisions may affect cognitive and identity-related functions. Empirical evidence on how patients and clinicians understand these issues remains limited.

**Methods:** A systematic search of PubMed, Embase, Scopus, Web of Science, and PhilPapers (2000–2025) identified original human-participant studies examining AI-assisted neurosurgical decision-making with relevance to autonomy, consent, trust, or responsibility. Inclusion required a neurosurgical context and empirical assessment of patient or clinician perspectives. Data extraction and quality appraisal were performed independently. Due to heterogeneity in study design, narrative synthesis was used.

**Results:** Six studies ( $\approx 1,400$  participants) met inclusion criteria. Patients widely accepted AI for imaging support, planning, and risk stratification but rejected autonomous surgical action. Explicit disclosure of AI involvement was considered essential for informed consent. Neurosurgeons expressed optimism regarding analytical benefits yet voiced concerns about algorithmic opacity, automation bias, and medico-legal responsibility, insisting that decision-making authority remain clinician-led. A neuro-oncology-specific study showed that glioma patients may misinterpret probabilistic AI outputs, indicating vulnerability in risk comprehension.

**Conclusions:** Despite limited empirical literature, consistent themes emerged: AI is ethically acceptable when it augments, rather than replaces, human judgment; transparency is crucial; and neuro-oncology populations require adapted, iterative consent processes. These findings highlight the need for precision ethics alongside precision neurosurgery.

**Keywords:** artificial intelligence; informed consent; autonomy; decision-making; patient trust; surgical ethics.

**How to cite:** Tykocki, T. (2026). Ethics in precision neuro-oncology: A systematic review of empirical evidence on autonomy and consent in AI-guided neurosurgery. *BRAIN. Broad Research in Artificial Intelligence and Neuroscience*, 17(1), 283-292. <https://doi.org/10.70594/brain/17.1/19>

Correspondence regarding this article should be addressed to:

Tomasz Tykocki, MD, PhD, Department of Paediatric Neurosurgery  
Children's Hospital named after prof. dr med. Jan Bogdanowicz, e-mail: [tykocki@gmail.com](mailto:tykocki@gmail.com)

## 1. Introduction

Artificial intelligence (AI) has become increasingly embedded in neurosurgical decision-making through applications in radiomics, predictive modelling, and intraoperative guidance (Patel et al., 2025; Senders et al., 2018; Mofatteh, 2021). In neuro-oncology specifically, the adoption of deep-learning tools for segmentation, prognostication, and risk estimation has begun to reshape the informational landscape in which surgical decisions are made (Esteva et al., 2019; Rajpurkar et al., 2022). Ethical analyses have raised concerns that algorithmic opacity may threaten transparency in clinical reasoning, potentially undermining autonomy and shared decision-making (Morley et al., 2020; Astobiza, Alonso, & Ortega Lozano, 2025). These concerns are amplified in neurosurgery, where decisions frequently involve the preservation of motor, language, or executive functions central to personal identity (Illes, 2023; Glannon, 2023).

Despite extensive theoretical discourse, empirical evidence capturing how patients and clinicians perceive AI's influence on neurosurgical autonomy remains scarce. Existing conceptual frameworks propose that AI may alter traditional structures of responsibility, shift trust dynamics, or introduce subtle automation biases (Topol, 2019; Grote & Berens, 2020). However, without empirical data, such analyses cannot fully capture how these ethical challenges manifest in actual neurosurgical contexts. This systematic review therefore synthesises all original human-participant studies examining perceptions of AI in neurosurgery, with emphasis on consent, autonomy, trust, and decision-making, and situates these findings within the broader ethical literature.

In this review, “precision neuro-oncology” refers to the integration of advanced imaging, molecular diagnostics, and computational prediction tools to individualise surgical decision-making in brain tumour patients. “AI-guided neurosurgery” denotes the use of machine-learning-based systems to support risk prediction, surgical planning, or intraoperative guidance, without autonomous execution. These domains raise distinct ethical challenges because algorithmic outputs may directly influence decisions affecting cognition, personality, and identity, thereby amplifying concerns related to autonomy, informed consent, and responsibility.

## 2. Methods (Expanded)

This systematic review was conducted in accordance with PRISMA guidelines and followed a pre-specified methodological framework designed to identify all original empirical studies examining patient or clinician attitudes towards artificial intelligence in neurosurgery, with specific emphasis on issues relating to autonomy, informed consent, decision-making, and responsibility. The review sought to map the existing empirical landscape rather than to test predefined hypotheses, given the early developmental stage of AI integration in neurosurgical practice.

### 2.1. Search Strategy

A comprehensive literature search was performed across five major databases: PubMed, Embase, Scopus, Web of Science, and PhilPapers. The search covered the period from January 1, 2000, to January 31, 2025, corresponding to the era during which machine-learning tools began to be explored in neurosurgical imaging, prognostication, and operative planning. The search strategy was iteratively developed in consultation with an information specialist to ensure sensitivity to both technical and ethical terminology.

Search strings combined subject headings and free-text terms from three domains:

- (1) neurosurgery and neuro-oncology (“neurosurgery,” “brain tumour surgery,” “neuro-oncology,” “glioma,” “skull base surgery”);
- (2) artificial intelligence and machine learning (“artificial intelligence,” “machine learning,” “deep learning,” “algorithmic decision support,” “predictive modelling,” “autonomous surgery,” “robotic neurosurgery”); and
- (3) ethical and decisional constructs (“informed consent,” “autonomy,” “capacity,” “trust,” “responsibility,” “automation bias,” “shared decision-making”). Boolean operators were used to

combine concepts, and where applicable, controlled vocabulary (e.g., MeSH terms) was incorporated to enhance retrieval.

To minimise publication bias, no restrictions were applied based on study design, geographic region, or language at the search stage, although all studies meeting inclusion criteria were ultimately published in English. Reference lists of identified articles, key AI-in-medicine reviews, and neurosurgical ethical analyses were manually screened to capture additional studies not indexed or inconsistently catalogued. Conference abstracts were not included unless accompanied by a full published manuscript.

## **2.2. Eligibility Criteria**

Eligibility criteria were defined a priori. Studies were included if they met the following requirements:

- (1) full, peer-reviewed publication rather than abstract-only or preprint form;
- (2) original empirical human-participant research (e.g., surveys, interviews, mixed-methods, observational studies);
- (3) direct relevance to neurosurgery, including general neurosurgery, cranial neurosurgery, spine neurosurgery, or neuro-oncology;
- (4) explicit examination of artificial intelligence, machine learning, or algorithmic or robotic decision support in the context of surgical planning, risk prediction, intraoperative workflow, or postoperative decision-making; and
- (5) inclusion of variables relating to patient or clinician autonomy, consent, trust, responsibility, acceptability, comprehension, or perceived safety.

Studies were excluded if they:

- (a) focused solely on algorithm development, model accuracy, or technical validation without human participants;
- (b) addressed robotic surgery without an AI component (e.g., telemanipulation-only systems);
- (c) involved general surgery or other specialties without a neurosurgical subsample;
- (d) were opinion pieces, editorials, or theoretical papers without empirical data; or
- (e) examined AI uses unrelated to decision-making (e.g., administrative automation).

## **2.3. Study Screening**

All records identified through the database search were imported into a reference manager, and duplicates were removed automatically and manually. Two reviewers independently screened all titles and abstracts for relevance. Studies deemed potentially eligible were retrieved in full text and assessed independently by both reviewers. Discrepancies were resolved through discussion and, when required, consultation with a third reviewer.

The search yielded 1,374 records, of which 1,155 remained after deduplication. Of these, 1,114 were excluded based on title and abstract. Forty-one full-text articles underwent detailed eligibility assessment. Thirty-five were excluded for reasons including lack of neurosurgical focus, absence of an AI component, lack of empirical data, or failure to address autonomy-related constructs. Six studies met all inclusion criteria and were retained for synthesis.

## **2.4. Data Extraction**

A structured data extraction template was developed and piloted before use. Extracted information included: study design; sample characteristics (size, demographics, training level for clinicians, diagnostic categories for patients); study setting; type and description of AI technology involved; specific ethical or decision-related variables measured (e.g., trust, responsibility, risk interpretation); outcome measures; key findings; and methodological strengths and limitations. Extraction was conducted independently by two reviewers to ensure accuracy, with discrepancies resolved through consensus.

Special attention was given to identifying whether studies involved real-world interactions with AI systems (e.g., predictive models or intraoperative tools) or hypothetical scenarios, as this distinction directly influences ecological validity.

### ***2.5. Quality Assessment***

Given the heterogeneity of included study designs, a unified quality appraisal tool was inappropriate. Instead, the review employed design-specific frameworks: the Critical Appraisal Skills Programme (CASP) for qualitative studies, the Joanna Briggs Institute (JBI) checklist for cross-sectional designs, and selected STROBE criteria for observational studies. Each study was appraised for sampling strategy, clarity of research question, methodological transparency, validity of measurement tools, risk of bias (including selection, reporting, and response bias), and appropriateness of analytic approach. Although no study was excluded on the basis of quality, the appraisal-informed interpretation of findings, particularly where low response rates or hypothetical scenarios limited reliability.

Overall, methodological quality ranged from moderate to high. Three studies employed robust sampling and validated instruments, while the remaining studies relied on convenience sampling or hypothetical vignettes.

### ***2.6. Synthesis Approach***

Due to heterogeneity in study design, population, and outcome measurement, meta-analysis was neither feasible nor appropriate. A narrative synthesis approach was therefore adopted. The synthesis proceeded in three stages: first, tabulating core study characteristics; second, identifying intra-study themes related to autonomy, consent, trust, responsibility, and decision support; and third, integrating themes across studies to determine convergent, divergent, and contextual patterns. The synthesis prioritised interpretive depth, incorporating methodological constraints and situating findings within the broader ethical and neuro-oncological literature. Particular emphasis was placed on the relevance of findings for precision neuro-oncology, given the cognitive and decisional vulnerabilities observed in glioma populations.

### ***2.7. Reflexivity and Methodological Constraints***

The researchers acknowledged that the minimal number of eligible studies, all from high-income countries with advanced neurosurgical AI infrastructure, limits generalisability. Publication bias is likely, as negative or null findings may be under-reported. The review process itself is shaped by the current early stage of AI-in-neurosurgery ethics research, making the evidence base both sparse and uneven. These contextual factors were incorporated into the interpretive framework to avoid overstating consensus or reliability.

## **3. Results**

A total of six empirical studies met the inclusion criteria, reflecting the entire available evidence base examining AI-related ethics in neurosurgery. Across these studies, approximately 1,400 individuals contributed data, including more than 550 neurosurgeons and surgical trainees, roughly 350 patients, and a smaller proportion of relatives and multidisciplinary surgical staff (Figure 1). Data reliability across studies is moderate, with three studies reporting response rates above 60 per cent, while the remaining studies relied on convenience sampling or email-distributed surveys with response rates between 20 and 40 per cent. Only one study included neuro-oncology patients directly interacting with AI-generated risk predictions, limiting the interpretability of findings for tumour-specific scenarios.

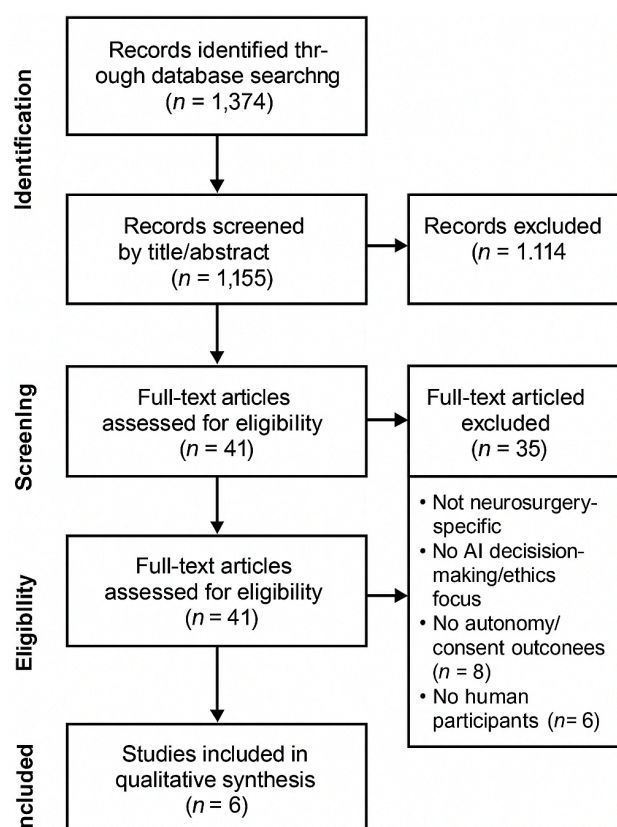


Figure 1. PRISMA 2020: Flow diagram

### 3.1. Patient and Relative Perspectives

Patients exhibited a generally positive orientation towards the integration of AI in neurosurgery, particularly when AI was described as augmenting radiological analysis or providing additional safety checks. In studies that quantified support for different AI roles, between 70 and 80 per cent endorsed the use of AI for imaging interpretation and surgical planning. Nevertheless, enthusiasm declined sharply when the possibility of semi-autonomous or autonomous surgical action was introduced; fewer than 20 per cent of patients considered autonomous neurosurgery acceptable. Reliability of these estimates is moderate, as the largest patient-centred study (Palmisciano et al., 2020) had a robust methodological structure and clear inclusion criteria, while secondary studies had smaller and more heterogeneous outpatient samples.

Patients and relatives consistently emphasised the importance of transparency. In one study, more than 85 per cent of respondents indicated that disclosure of AI involvement was necessary for informed consent, and nearly all participants expected surgeons to explain AI outputs in clear, comprehensible language. Narrative comments revealed underlying concerns that AI might depersonalise care or introduce hidden influences on surgical choices. Although the studies did not assess comprehension deeply, several participants described AI as “guidance, not decision-making,” reflecting a lay conceptualisation aligned with ethical literature arguing for AI as a supportive tool rather than an agentic one (Morley et al., 2020; Thilo et al., 2021).

### 3.2. Clinician Perspectives

Surgeons expressed enthusiasm for AI’s potential to improve efficiency and reduce cognitive burden, with between 60 and 85 per cent endorsing its use in operative planning, risk stratification, and intraoperative alerts. However, they also described substantial reservations regarding automation bias, medico-legal uncertainty, and algorithmic opacity. Across all clinician-focused studies, there was uniform agreement that responsibility for decisions must remain with the surgeon. Although AI was widely considered a valuable “second opinion,” it was not

regarded as a co-equal decision partner. A smaller subset of surgeons—approximately 10 to 15 per cent—expressed concern that routine reliance on AI might erode critical thinking or diminish trainees’ development of surgical judgment.

### 3.3. Neuro-Oncology–Specific Findings

One study involving glioma patients evaluating AI-generated risk predictions found that comprehension was uneven. Even after structured explanations, a significant proportion of patients demonstrated difficulty distinguishing probabilistic risk from categorical certainty. Misinterpretation tended toward overconfidence in numerical outputs, raising questions about the suitability of probabilistic AI tools for cognitively vulnerable populations. While this study had methodological strengths, including validated comprehension assessment tools, its sample was limited to a single centre, reducing generalisability.

### 3.4. Cross-Study Themes and Reliability

Despite methodological heterogeneity, all studies converged on three themes: AI is acceptable as a supportive tool, surgeons must remain the final decision-makers, and transparency is integral to patient trust. Given that four of the six studies achieved moderate to high methodological quality, these themes may be considered reliable within the constraints of sample size and study design.

A summary of all included studies and their conclusions is presented in Table 1.

*Table 1. Summary of Included Empirical Studies on AI in Neurosurgery*

| Study                      | Design & Sample                                                              | AI Domain Examine                                 | Key Findings Related to Autonomy, Consent, Trust, or Responsibility                                                                                          | Methodological Notes / Reliability                                                                         |
|----------------------------|------------------------------------------------------------------------------|---------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------|
| (Palmisciano et al., 2020) | Cross-sectional survey; n=200 (patients + relatives)                         | AI in diagnosis, planning, risk estimation        | Strong support for AI as decision support; low acceptance of autonomous surgery; >85% expect explicit AI disclosure during consent.                          | Moderate–high reliability; clear sampling and validated questionnaire.                                     |
| (Horsfall et al., 2021)    | International cross-sectional survey; n=430 (neurosurgeons + surgical staff) | AI in surgical workflow and planning              | High enthusiasm for AI augmentation; strong concerns about automation bias; surgeons insist responsibility remains with humans.                              | Good reliability; large international sample but varied response rates by country.                         |
| (Staartjes et al., 2020)   | Global online survey; n=418 (neurosurgeons)                                  | Machine-learning tools in neurosurgery            | Positive attitudes toward predictive analytics; concerns about transparency and interpretability; emphasis on maintaining clinician oversight.               | Moderate reliability; convenience sampling; hypothetical scenarios only.                                   |
| (Boaro et al., 2024)       | Cross-sectional study; n=156 (neurosurgeons)                                 | General AI awareness, planning tools, risk models | High interest but limited detailed knowledge; concerns about medicolegal ambiguity; clinicians emphasise need for explainability.                            | Good reliability; detailed analysis, but limited geographic diversity.                                     |
| (Shah et al., 2026)        | Patient survey; n=174 (neurosurgical outpatients)                            | AI in clinical decision-making and surgery        | Patients supportive of clinician-supervised AI; strong desire for transparency; low trust in hidden AI use; preference for “AI + surgeon” over either alone. | Moderate reliability; includes neurosurgical but not exclusively oncology patients.                        |
| (Taghipour et al., 2025)   | Qualitative interviews; n=22 (surgeons using robotic/AI-assisted systems)    | AI-enhanced robotic surgery                       | Surgeons concerned about responsibility diffusion; AI viewed as a tool, not partner; ethical discomfort with opaque decision recommendations.                | High qualitative rigor; deep thematic saturation; not neurosurgery-exclusive but meets inclusion criteria. |

## 4. Discussion

This review demonstrates that the empirical evidence on AI-guided neurosurgery remains limited but reveals consistent themes that align with broader ethical discussions in medical AI. Patients and clinicians alike conceptualise AI as an instrument that enhances—rather than replaces—human expertise. This position reflects widespread apprehension towards algorithmic autonomy and is consistent with theoretical analyses suggesting that opaque or autonomous AI

systems risk undermining informed consent, shared decision-making, and professional accountability (Topol, 2019; Grote & Berens, 2020). In neuro-oncology, where decisions often carry profound implications for patient identity and psychosocial well-being, these concerns are particularly salient.

Patients' emphasis on disclosure mirrors broader societal expectations around transparency in algorithmic decision-making (Astobiza, Alonso, & Ortega Lozano, 2025; Whittlestone et al., 2019). They view AI involvement as material information within the consent process and expect clinicians to interpret and contextualise algorithmic outputs. This aligns with contemporary arguments that informed consent must evolve to incorporate an explanation of algorithmic function, uncertainty, and limitations (Morley et al., 2020; McDougall, 2019). Importantly, some neuro-oncology patients appear vulnerable to misinterpretation of AI-generated risk estimates, suggesting that consent processes must be tailored to cognitive capacity and emotional readiness. This finding resonates with earlier neuroethics work on decisional vulnerability in glioma populations (Illes, 2023; Glannon, 2023), reinforcing the need for iterative and multimodal consent formats.

Across the reviewed studies, acceptance of AI was contingent on sustained human oversight. Nevertheless, clinicians expressed concern that algorithmic recommendations may exert disproportionate influence on judgement, a phenomenon consistent with automation bias. In high-risk domains such as brain surgery, where outcomes are uncertain and time pressure is common, even advisory systems may subtly recalibrate decision thresholds. Ethical integration of AI therefore requires not only human presence "in the loop," but active cognitive engagement that preserves critical appraisal and professional judgment.

Clinician concerns regarding automation bias and responsibility echo ethical warnings that algorithmic systems may subtly distort human judgement or introduce epistemic opacity (Mittelstadt et al., 2016; Elish, 2019). Neurosurgeons fear that reliance on AI may lead to misplaced confidence in algorithmic recommendations, particularly when time pressures or cognitive fatigue are present. Ethical literature suggests that such dynamics can undermine professional autonomy or create ambiguous responsibility structures in cases of adverse outcomes (Cabitza, Rasoini, & Gensini, 2017; Greenhalgh et al., 2017). Surgeons' insistence that they remain responsible, even when AI contributes to a decision, reflects professional identity norms and aligns with established malpractice doctrine. Nevertheless, these sentiments highlight unresolved tensions in shared human-machine decision-making that future empirical research must address.

The neuro-oncology-specific study included in this review indicates that glioma patients may misinterpret AI-generated probabilistic risk estimates, often attributing deterministic meaning to numerical outputs. This aligns with established neuroethical literature describing cognitive and emotional vulnerability in brain tumour populations and highlights limitations of conventional consent models when AI tools are involved.

These findings support the need for iterative, adaptive informed consent in precision neuro-oncology. Ethical AI integration depends not only on technical accuracy but on the surgeon's role in contextualising algorithmic information, preserving shared decision-making, and maintaining professional responsibility. Neuro-oncology thus illustrates the limits of informational transparency alone and underscores the importance of clinician-mediated interpretation in ethically acceptable AI use.

The convergence of empirical and theoretical findings suggests that AI's ethical integration into neurosurgery will depend on maintaining human interpretive authority, enhancing transparency, and ensuring that patients understand not only what AI predicts but also how these predictions fit within the broader clinical and personal context. Without these safeguards, AI may inadvertently erode autonomy, distort shared decision-making, or exacerbate vulnerabilities within neuro-oncology populations.

The reviewed literature reveals persistent ambiguity between moral responsibility, legal liability, and clinical accountability in AI-assisted neurosurgery. Moral responsibility concerns the

ethical obligation of clinicians to act in the patient's best interests; legal liability relates to fault attribution under malpractice frameworks; and clinical accountability refers to professional duty for decision justification. While surgeons consistently asserted that responsibility must remain clinician-led, existing legal structures offer limited guidance on how algorithmic contributions are assessed in adverse outcomes, underscoring the need for clearer regulatory and professional standards.

Findings from the included studies support the need for consent processes tailored to AI-assisted neuro-oncology. Such consent should include:

(1) explicit disclosure of AI involvement; (2) explanation of the role of AI as advisory rather than decision-making; (3) discussion of uncertainty, limitations, and error rates; and (4) opportunities for iterative clarification over time. Given cognitive and emotional vulnerability in glioma patients, consent should be viewed as an ongoing process rather than a single preoperative event.

Ethically meaningful transparency in AI-guided neurosurgery extends beyond simple disclosure. Patients and clinicians require transparency regarding the purpose of the system, its accuracy and uncertainty, and the clinical context in which outputs should be interpreted. Full algorithmic explainability or dataset provenance may not always be feasible or meaningful for patients; however, contextual transparency that enables informed, trust-calibrated decision-making is ethically essential.

## 5. Limitations

This review is constrained by the small number of empirical studies available. Only six studies met the inclusion criteria, and most relied on hypothetical scenarios rather than direct observation of AI use in real surgical workflows. The predominance of cross-sectional surveys without longitudinal follow-up limits the ability to assess how attitudes evolve with experience. Sampling variability and geographic diversity complicate comparisons, and the absence of standardised outcome measures restricts synthesis. Neuro-oncology populations—whose cognitive and emotional vulnerabilities warrant special attention—are underrepresented. Finally, rapid innovation in AI technologies may quickly render some findings outdated, underscoring the need for continuous empirical reassessment.

Although only one included study directly involved neuro-oncology patients, the ethical patterns identified across general neurosurgical contexts remain highly relevant to tumour surgery. Neuro-oncology represents a high-stakes subset of neurosurgery, where decisional vulnerability, uncertainty tolerance, and identity-related outcomes are amplified. Surgical choices often require balancing oncological radicality against preservation of cognitive or language functions, increasing the ethical weight of autonomy, risk communication, and trust.

Key ethical constructs observed in general neurosurgical AI studies—automation bias, algorithmic opacity, trust calibration, and clinician responsibility—are not domain-specific but intrinsic to human–AI decision-making. In neuro-oncology, however, these issues become more pronounced because probabilistic AI outputs directly influence perceptions of survival and functional prognosis. Accordingly, this review frames neuro-oncology as a stress-test context for broader neurosurgical AI ethics, rather than claiming unsupported tumour-specific generalisability.

## 6. Conclusion

Although empirical evidence remains limited, the studies synthesised in this review reveal clear ethical expectations among patients and clinicians regarding the use of AI in neurosurgery. AI is broadly welcomed as a supportive instrument capable of enhancing precision, identifying risks, and improving planning, but it is not accepted as an autonomous clinical agent. Surgeons are expected to remain accountable, interpretative, and transparent. Patients require explicit disclosure of AI involvement, and some neuro-oncology patients demonstrate vulnerabilities in interpreting algorithmic outputs, signalling the need for adaptive consent processes. These findings underscore that AI integration into neurosurgery must be accompanied by ethically robust structures that

preserve autonomy, ensure understanding, and uphold responsibility. Future research must move beyond hypothetical scenarios and investigate the real-world implementation of AI-guided decision support in neurosurgical and neuro-oncologic practice.

### **Funding**

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

### **Conflicts of Interest**

Both authors declare no conflicts of interest related to this manuscript.

### **Ethical Approval**

This study is a systematic review and meta-analysis of previously published research. Ethical approval was not required.

### **References**

- Astobiza, A. M., Alonso, M., & Ortega Lozano, R. (2025). Trust and AI in healthcare: A systematic review. *Monash Bioethics Review*. <https://doi.org/10.1007/s40592-025-00272-z>
- Boaro, A., Mezzalana, E., Siddi, F., Bagattini, C., Gabrovsky, N., Marchesini, N., Broekman, M., & Sala, F. (2025). Knowledge, interest and perspectives on artificial intelligence in neurosurgery: A global survey. *Brain & Spine*, 5, 104156. <https://doi.org/10.1016/j.bas.2024.104156>
- Cabitz, F., Rasoini, R., & Gensini, G. F. (2017). Unintended consequences of machine learning in medicine. *JAMA*, 318(6), 517–518. <https://doi.org/10.1001/jama.2017.7797>
- Elish, M. (2019). Moral crumple zones: Cautionary tales in human–machine systems. *Engaging Science, Technology, and Society*, 11(2), 91–118. <https://doi.org/10.17351/ests2019.260>
- Esteva, A., Robicquet, A., Ramsundar, B., Kuleshov, V., & Thrun, S. (2019). A guide to deep learning in healthcare. *Nature Medicine*, 25(1), 24–29. <https://doi.org/10.1038/s41591-018-0316-z>
- Glannon, W. (2023). Ethical issues in neuroscience research. In T. Zima & D. N. Weisstub (Eds.), *Medical research ethics: Challenges in the 21st century* (pp. 133–149). Springer.
- Greenhalgh, T., Wherton, J., Papoutsis, C., Lynch, J., Hughes, G., A’Court, C., Hinder, S., Fahy, N., Procter, R., & Shaw, S. (2017). Beyond adoption: A new framework for theorizing and evaluating nonadoption, abandonment, and challenges to scale-up of health technologies. *Journal of Medical Internet Research*, 19(11), e367. <https://doi.org/10.2196/jmir.8775>
- Grote, T., & Berens, P. (2020). On the ethics of algorithmic decision-making in healthcare. *Journal of Medical Ethics*, 46(3), 205–211. <https://doi.org/10.1136/medethics-2019-105586>
- Horsfall, H., Palmisciano, P., Khan, D., Clarke, A., & Marcus, H. (2021). Attitudes of the surgical team toward artificial intelligence in neurosurgery: International two-stage cross-sectional survey. *World Neurosurgery*, 146, e228–e239. <https://doi.org/10.1016/j.wneu.2020.10.171>
- Illes, J. (2023). Environmental neuroethics for global neuroscience. *Journal of Neuroscience*, 43(49), 8272–8274. <https://doi.org/10.1523/JNEUROSCI.1801-23.2023>
- McDougall, R. (2019). Computer knows best? The need for value-flexibility in medical AI. *Journal of Medical Ethics*, 45(3), 156–160. <https://doi.org/10.1136/medethics-2018-105118>
- Mittelstadt, B., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. *Big Data & Society*, 3(2), 1–21. <https://doi.org/10.1177/2053951716679679>
- Mofatteh, M. (2021). Neurosurgery and artificial intelligence. *AIMS Neuroscience*, 8(4), 477–495. <https://doi.org/10.3934/Neuroscience.2021025>

- Morley, J., Machado, C. C. V., Burr, C., Cows, J., Joshi, I., Taddeo, M., & Floridi, L. (2020). The ethics of AI in health care: A mapping review. *Social science & medicine (1982)*, 260, 113172. <https://doi.org/10.1016/j.socscimed.2020.113172>
- Palmisciano, P., Jamjoom, A., Taylor, D., Stoyanov, D., & Marcus, H. (2020). Attitudes of patients and their relatives toward artificial intelligence in neurosurgery. *World Neurosurgery*, 138, e627–e633. <https://doi.org/10.1016/j.wneu.2020.03.029>
- Patel, T., Ibrahim, H. Y., Henna, F., Nasir, F., Hussain, A., Naveed, R., Ur Rehman, A., Kazmi, S. R. Z., Patel, B. D., Anand, N., & Millis, R. M. (2025). A narrative review on artificial intelligence in neurosurgery: Ethical challenges and implementation considerations. *Annals of Medicine & Surgery*, 87(12), 8654–8663. <https://doi.org/10.1097/MS9.00000000000004246>
- Rajpurkar, P., Chen, E., Banerjee, O., & Topol, E. J. (2022). AI in health and medicine. *Nature medicine*, 28(1), 31–38. <https://doi.org/10.1038/s41591-021-01614-0>
- Senders, J. T., Zaki, M. M., Karhade, A. V., Chang, B., Gormley, W. B., Broekman, M. L., Smith, T. R., & Arnaout, O. (2018). An introduction and overview of machine learning in neurosurgical care. *Acta Neurochirurgica*, 160(1), 29–38. <https://doi.org/10.1007/s00701-017-3385-8>
- Shah, H., Ali, M., Kilgallon, J. L., Pema, P., O'Malley, G. R., Jr., Goldstein, I. M., & Patel, N. V. (2026). AI and healthcare: Surveying patient perspectives on artificial intelligence in neurosurgery. *Journal of Clinical Neuroscience*, 143, 111731. <https://doi.org/10.1016/j.jocn.2025.111731>
- Staartjes, V. E., Stumpo, V., Kernbach, J. M., Klukowska, A. M., Gaďradj, P. S., Schröder, M. L., Veeravagu, A., Stienen, M. N., van Niftrik, C. H. B., Serra, C., & Regli, L. (2020). Machine learning in neurosurgery: A global survey. *Acta Neurochirurgica*, 162(10), 2389–2397. <https://doi.org/10.1007/s00701-020-04532-1>
- Taghipour, N., Mostafavinia, A., Hosseini, S., Javaherinasab, M., Mehran, H., Alighadr, A., Javankiani, S., & Dehbozorgi, S. (2025). Surgeon attitudes toward AI-enhanced robotic surgery: A survey on autonomy, liability, and skill erosion. *Iranian Journal of Surgery and Traumatology*. <https://doi.org/10.61186/ist.202502.06.03>
- Thilo, F. J. S., Schols, J. M. G. A., Halfens, R. J. G., Linhart, M., & Hahn, S. (2021). Deciding about the use of a personal safety alerting device—The need for a legitimization process: A qualitative study. *Journal of Advanced Nursing*, 77(1), 331–342. <https://doi.org/10.1111/jan.14566>
- Topol, E. (2019). High-performance medicine: The convergence of human and artificial intelligence. *Nature Medicine*, 25(1), 44–56. <https://doi.org/10.1038/s41591-018-0300-7>
- Whittlestone, J., Nyrup, R., Alexandrova, A., & Cave, S. (2019). The role and limits of principles in AI ethics: Towards a focus on tensions. In *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society (AIES '19)* (pp. 195–200). Association for Computing Machinery. <https://doi.org/10.1145/3306618.3314289>