The Declaration of Sydney: Ethical, Legal, and Professional Foundations of Computational Neurosurgery

Antonio Di Ieva[1,2,3]*, Amin Tavallaii[1,4]*, Margaret Somerville[5], Eric Suero Molina[6], Aimee DeGaetano[7], Allan McCay[8], Katrina Hutchison[9], Quinlan Buchlak[10,11], Rita Matulionyte[12], Wendy Rogers[9], Jeffrey V. Rosenfeld[13,14,15]

*These authors contributed equally

Affiliations:

1. Computational NeuroSurgery (CNS) Lab, Macquarie Medical School, Macquarie University, Sydney, NSW, Australia

2. Macquarie Neurosurgery & Spine, Macquarie University Hospital, Sydney, NSW, Australia

3. Department of Neurosurgery, Nepean Blue Mountains Local Health District, Kingswood, NSW, Australia

4. Department of Neurosurgery, Mashhad University of Medical Sciences, Mashhad, Iran

5. National School of Medicine, Professor of Bioethics, University of Notre Dame Australia, Sydney Campus, NSW, Australia

6. Department of Neurosurgery, University Hospital Münster, Münster, Germany

7. Department of Health Science and Medicine, Eastern Virginia Medical School, Old Dominion University, Norfolk, VA, United States

8. Sydney Institute of Criminology, Law School, University of Sydney, Sydney, NSW, Australia

9. Discipline of Philosophy, School of Humanities, Macquarie University, Sydney, NSW, Australia

10. National School of Medicine, University of Notre Dame Australia, Sydney Campus, NSW, Australia

11. Medical Sciences, Medtronic

12. Law School, Macquarie University, Sydney, NSW, Australia

13. Department of Neurosurgery, Alfred Hospital, Melbourne, Victoria, Australia

14. Department of Surgery, Monash University, Melbourne, Victoria, Australia

15. Adjunct Professor, F. Edward Hébert School of Medicine, Uniformed Services University of The Health Sciences, Bethesda, MD, USA

Corresponding Authors:

Antonio Di Ieva

Email: antonio.diieva@mq.edu.au, Address: Computational Neurosurgery Lab, 75 Talavera Road, Macquarie Park, NSW, Australia 2113

Amin Tavallaii

Email: amin.tavallaii@mq.edu.au | dr.amin.tavallaii@gmail.com, Address: Computational Neurosurgery Lab, 75 Talavera Road, Macquarie Park, NSW, Australia 2113

Abstract

Artificial intelligence (AI), data-driven methods, and emerging neurotechnologies are increasingly shaping how neurosurgery is practised, studied, and taught. Within this evolving landscape, *computational neurosurgery* has emerged as a transdisciplinary field that integrates computation, artificial and augmented intelligence, data science, and neurotechnology across the neurosurgical care pathway, encompassing applications in diagnosis, risk stratification and screening, clinical decision support, surgical planning, intraoperative support, postoperative management, research, education, and training, with the overarching aim of improving fundamental understanding, clinical decision-making, and patient outcomes. Computational methodologies have been shown to, or at least promise to, influence clinical judgment throughout neurosurgical care. While such developments hold considerable potential for enhancing precision and expanding access to care, they also introduce ethical, legal, and societal questions that are not readily addressed by generic approaches to AI governance. Neurosurgery occupies a particularly sensitive position in this context, given its direct engagement with the brain as the substrate of consciousness, identity, and agency.

In response to these challenges, an international and multidisciplinary task force was convened in association with the first World Conference of Computational Neurosurgery (WCCNS, Sydney, Australia, 13-15 February 2026). Its goal was to develop a shared ethical framework and embed it in neurosurgical practice, to guide the responsible use of AI and computational methods in neurosurgery and other complementary disciplines. Working through an iterative process of deliberation, collaborative drafting, and explicit documentation of consensus and disagreement, the task force produced two complementary outcomes: a concise public pledge, the Declaration of Sydney, and the present white paper, which elaborates the ethical and conceptual foundations underlying that pledge.

This white paper examines the emergence of computational neurosurgery as both a clinical and epistemic shift, and explores its implications for governance, data stewardship, clinical integration, education, equity, sustainability, and long-term oversight. With a patient-centred approach, the Declaration of Sydney affirms the primacy of patients, respect for human dignity, professional accountability, and transparency, while acknowledging the need for governance frameworks capable of adapting to evolving technologies and societal expectations. Taken together, the goal of the Declaration and the analysis presented in this Whitepaper is to benefit patients needing neurosurgical care, support responsible innovation, encourage transdisciplinary engagement, and contribute to the ethical maturation of computational neurosurgery as a global field. In such a perspective, the Declaration of Sydney emerges as a global framework for responsible computational neurosurgery, defining the articles and

ethical principles neurosurgeons should abide by when using AI and computational techniques.

Keywords: neurosurgery, computational neurosurgery, neuro-ethics, neuro-philosophy, neurotechnology, artificial intelligence

1. Introduction

Recent advances in artificial intelligence, data-driven analysis, and neurotechnology are reshaping how medicine is practised and understood. In neurosurgery, the growing capacity to integrate computational modelling with complex clinical and biological data has opened new possibilities for prediction, interpretation, and decision support that extend beyond the limits of unaided human cognition. These developments hold potential to improve diagnostic timing, refine surgical planning, support learning health systems, and extend specialised expertise to settings where it has historically been scarce(1). At the same time, they raise ethical, legal, and societal questions that are not easily resolved within existing medical norms or by governance frameworks designed for artificial intelligence in general(2).

Neurosurgery occupies a particularly sensitive position within this transformation. Unlike many other clinical specialties, it routinely intervenes in neural structures closely associated with thought, memory, emotion, identity, and personal agency. In this context, computational neurosurgery has emerged as a transdisciplinary field integrating computation, artificial and augmented intelligence, data science, and neurotechnology across the neurosurgical care pathway. Within this evolving domain, computational approaches are increasingly being explored and, in some settings, integrated across stages of care, from risk stratification and diagnosis to operative planning, intraoperative support, postoperative management, rehabilitation, and longer-term outcome assessment, while their adoption remains uneven across institutions, health systems, and global contexts and continues to evolve with ongoing technological development. Precisely because these technologies are expanding in capability while remaining variably embedded in practice, they bring into sharper focus ethical concerns that extend beyond familiar questions of safety and effectiveness. Issues of mental privacy, cognitive liberty, trust, human dignity, and professional accountability become increasingly salient, particularly in relation to technologies such as robotics-assisted neurosurgery, brain-machine and brain-computer interfaces, adaptive neurostimulation systems, and other invasive or non-invasive neurotechnologies that blur the boundaries between observation, prediction, and intervention(3).

Against this backdrop, an international multidisciplinary task force was convened under the auspices of the first World Conference of Computational Neurosurgery (WCCNS, Sydney, Australia, 13-15 February 2026) to develop a shared ethical and legal framework to guide the responsible use of artificial intelligence and other computational methods in neurosurgical care. The present paper provides the scholarly foundation for that effort. Drawing on perspectives from clinical practice, computational science, ethics, and law, it articulates the rationale, scope, and governing principles that underpin the Declaration of Sydney on the Ethical Use of AI in Neurosurgery.

2. Methods

2.1 Design and Governance Framework

This work was undertaken as an international, multidisciplinary consensus initiative focused on ethical and normative analysis rather than empirical investigation. The overall approach combined structured deliberation with iterative drafting and critical review, allowing ideas to be proposed, challenged, refined, and consolidated over time. Rather than seeking rapid convergence, the process emphasised careful reflection and transparency around both agreement and disagreement.

Because the project did not involve human participants, patient data, or experimental intervention, institutional research ethics approval was not required. Governance of the process instead relied on shared responsibility among task force members, with explicit commitments to openness, collegial oversight, and intellectual integrity guiding all stages of the work.

2.2 Task Force Composition

The Declaration of Sydney Task Force brought together contributors from a broad range of disciplinary and geographic backgrounds, including neurosurgery, neuroscience, philosophy, ethics, law, data science, global health, industry, academia, and medical education. Members were invited based on recognised expertise, scholarly contribution, and sustained engagement with computational, ethical, or regulatory dimensions of neurosurgical practice.

Throughout the process, the task force operated under shared principles of mutual respect and constructive dialogue. Deliberations were intentionally structured to encourage exchange across clinical, technical, ethical, and legal perspectives, with the goal of developing a framework that reflects the complexity of computational neurosurgery rather than privileging any single disciplinary viewpoint.

2.3 Consensus-Building Process

The consensus-building process unfolded over a series of structured yet iterative stages designed to balance inclusivity with progressive refinement. The initial phase focused on scoping and idea generation. Early meetings were devoted to defining the contours of computational neurosurgery and identifying the ethical, legal, clinical, and societal questions most relevant to the field. Contributions were made through live discussion as well as asynchronous written input, allowing time for reflection across different time zones and disciplinary traditions. At this stage, breadth of perspective was prioritised over resolution.

The second phase centred on collaborative drafting and focused deliberation. A shared online "live document" served as the principal workspace, enabling task force members to propose text, suggest article language, and comment directly on one another's contributions. Participation was intentionally inclusive, and divergent views were explicitly welcomed. Regular meetings were held online to discuss points of convergence and disagreement, refine terminology, and clarify the intended relationship between the public Declaration and its accompanying scholarly analysis. Formal meeting minutes were maintained throughout to document decisions, areas of consensus, and issues requiring further reflection.

The final phase involved synthesis and consolidation. In accordance with task force agreement, drafting shifted from expansive idea generation toward the development of two distinct but closely related outputs: a concise public Declaration intended for reading and endorsement (set on the 15th of February 2026 at the 1st WCCNS), and a comprehensive scholarly manuscript, presented here as a white paper, which preserves the depth of ethical analysis and contextual discussion underlying the pledge. During this stage, overlapping contributions were merged, redundant language reduced, and thematic coherence strengthened, while care was taken to preserve the substance and intent of contributors' positions, particularly where unresolved tensions were ethically significant.

2.4 Decision Rules and Editorial Principles

Editorial decisions were guided by a set of principles developed and reaffirmed during task force deliberations. For the public Declaration, priority was given to positions reflecting broad agreement across the group. The pledge was not intended to catalogue all views expressed, but to articulate a shared ethical foundation capable of collective endorsement. By contrast, the scholarly manuscript was deliberately more extensive. Divergent perspectives, unresolved debates, and emerging concepts were retained and contextualised rather than excluded, reflecting the view that ethical maturity requires transparency about disagreement as well as consensus.

A clear distinction was also maintained between the goals of the two outputs. The Declaration was shaped to be concise, accessible, and suitable for public reading and endorsement, while the preprint was developed to allow greater analytical depth, conceptual nuance, and engagement with scholarly and regulatory discourse. Across both documents, the task force avoided jurisdiction-specific legal prescriptions, favouring principle-based guidance capable of remaining relevant across diverse legal, cultural, and health-system contexts.

Taken together, these methodological and editorial commitments provided not only a framework for consensus-building, but also a lens through which to examine the evolving identity and scope of computational neurosurgery. The sections that follow therefore move beyond process description to consider how the field has emerged, why it raises distinct ethical and societal concerns, and why a sector-specific declaration is warranted.

3. The Emergence of Computational Neurosurgery

3.1 From Technique to Epistemic Transformation

Computational neurosurgery has emerged not simply as a collection of new tools, but as a broader transformation in how neurosurgical knowledge is produced, interpreted, and applied. Traditionally, neurosurgical decision-making has depended on the synthesis of clinical experience, anatomical understanding, imaging interpretation, and intraoperative judgment, often exercised under conditions of uncertainty and time pressure. While these "human ways of knowing" remain relevant and important, computational approaches now make it possible to translate complex neural phenomena into structured representations that can be modelled, simulated, and systematically explored.

This evolution has important epistemic consequences. Increasingly, algorithms mediate how evidence is generated and how patterns are recognised, with predictions drawn from population-scale datasets rather than individual experience alone(4). Decision-support systems may influence clinical reasoning in real time, shaping how risks are perceived and options weighed. In this sense, computational neurosurgery alters the relationship between data, knowledge, and judgment, introducing new dependencies alongside new forms of insight. These changes create opportunities for more consistent and data-informed care, while also raising questions about interpretability, uncertainty, and the locus of professional responsibility.

3.2 Scope of Application Across the Neurosurgical Pathway

Computational methods are increasingly embedded across the continuum of neurosurgical care. At present, some of the most mature and widely studied applications lie in diagnostics, particularly the use of deep learning, radiomics (and other multi-omics approaches), and data-driven analysis in neuroimaging and digital neuropathology to support detection, classification, and prognostic assessment of neurological disease. In the preoperative setting more broadly, machine learning models are being explored to assist diagnosis, assess risk, evaluate surgical candidacy, and inform individualised treatment planning. During surgery, navigation systems, robotics, computer vision, and real-time analytics contribute to precision and operative safety. After surgery, predictive and adaptive models support monitoring, complication detection, rehabilitation planning, and longer-term assessment of outcomes(5).

The scope of computational neurosurgery continues to expand through the growing use of implanted and semi-autonomous neurotechnologies. AI-augmented neural devices, including open- and closed-loop brain–computer interfaces, neurostimulation systems, and spinal or peripheral neural implants, may operate continuously within the body, adjusting neural activity in response to ongoing physiological signals. These technologies blur traditional temporal boundaries between planning, intervention, and follow-up, introducing new forms of persistent clinical influence and shared responsibility over time.

Beyond direct patient care, computational approaches are reshaping neurosurgical research, education, and health-system organisation. Simulation platforms, virtual and augmented reality environments, and adaptive learning tools are increasingly incorporated into training, while large-scale analytics inform publication practices, policy development, and decisions about resource allocation. In this way, computational neurosurgery extends beyond the operating theatre to influence how the field learns, evaluates itself, and plans for future capacity.

3.3 A Discipline or a Set of Practices?

Throughout task force deliberations, a recurring question was whether computational neurosurgery should be regarded as a distinct discipline or as a collection of practices embedded within neurosurgery. While perspectives varied, there was broad agreement that the field encompasses elements of both. It functions as a transdisciplinary domain of research and development, involving engineers, data scientists, neuroscientists, neuro-philosophers, lawyers, and ethicists, while also operating as a clinical practice in which neurosurgeons retain responsibility for patient care.

This dual character has significant implications for governance. Ethical and legal responsibilities must be distributed across those who design, develop, deploy, regulate, and finance computational systems, while preserving clear lines of clinical accountability. Therefore, computational neurosurgery cannot be governed as if AI were a neutral or purely technical addition to existing practice. Instead, it requires models of oversight that reflect its hybrid socio-technical nature, recognising that ethical considerations arise not only at the point of clinical use, but also earlier, in research priorities, system design choices, and investment decisions that shape which technologies are developed, scaled, and ultimately integrated into neurosurgical care.

4. The Case for a Sector-Specific Declaration

4.1 Limits of General AI Regulation

Recent efforts to address artificial intelligence through broad, cross-sector frameworks have made clear both their value and their limitations. While such approaches can establish high-level principles, they often struggle to account for the contextual specificity and risk profiles of clinical domains(6). The Declaration of Sydney is not intended to function as a regulatory instrument, nor does it seek to replace existing or future legal oversight. Rather, it is conceived as a principled ethical framework designed to guide the responsible development, deployment, and use of computational systems in neurosurgery, and to complement formal regulation by articulating shared normative commitments. At the same time, the Declaration recognises that computational neurosurgery exists within a rapidly evolving scientific and technological ecosystem; accordingly, it is intended not as a fixed or immutable statement, but as a living and adaptive instrument capable of evolving in response to advances in knowledge, practice, and societal expectations.

Neurosurgical applications of AI differ in important ways from consumer, commercial, or other low-stakes uses of algorithmic systems. They operate in settings characterised by clinical vulnerability, uncertainty, and pronounced asymmetries of knowledge and power. Although not all AI-enabled interventions in neurosurgery are strictly irreversible, many devices can be adjusted, deactivated, or removed, computational systems may nonetheless influence decisions and interventions that affect neural structure or function with potentially lasting consequences for cognition, identity, and quality of life. Even where recovery, neuroplasticity, or rehabilitation is possible, the ethical stakes remain high. These features underscore the need for guidance that is sensitive to the distinctive clinical, moral, and societal dimensions of neurosurgical care, rather than reliance on generic AI ethics frameworks alone(7).

Task force deliberations consistently highlighted that existing cross-sector approaches to AI ethics do not sufficiently capture these realities. What is required instead is guidance that is clinically grounded, proportionate to risk, and attentive to the particular forms of risk, harm, responsibility, and trust that arise when computational systems are applied to the human nervous system.

4.2 The Role of Declarations in Medical Governance

Declarations have long played a formative role in medical ethics, shaping professional norms and influencing policy in areas ranging from research governance to global health. Although non-binding, they articulate shared values, set expectations for responsible practice, and often serve as reference points for education, institutional standards, and regulatory development.

Importantly, declarations are not static instruments; they are capable of evolving in response to new evidence, technological change, and shifting societal expectations(8).

The Declaration of Sydney is intended to function within this tradition. It does not attempt to resolve every ethical or legal question raised by computational neurosurgery, nor does it prescribe detailed rules for compliance. Instead, it establishes a shared ethical foundation that affirms patient primacy, respect for human dignity, professional accountability, and global equity. By doing so, it aims to guide future research, clinical practice, and governance, while providing a common point of reference for ongoing dialogue as the field continues to develop.

5. Ethical, Legal, and Societal Challenges

5.1 Patient Vulnerability, Clinical Responsibility, and Consent

The integration of computational and AI-enabled systems into neurosurgical care does not lessen the ethical obligations of the profession and its practitioners; rather, it brings them into sharper focus. Neurosurgeons continue to have a fiduciary duty to act in the best interests of their patients, grounded in care, honesty, confidentiality, informed consent, and good professional judgment. Computational systems complicate this duty, particularly where they are designed or optimised for objectives, such as efficiency, throughput, or commercial performance, goals that may not fully align with giving priority to the welfare of individual patients.

Neurosurgical practice is especially vulnerable to ethical risk in this context. Decisions are often made under conditions of time pressure, uncertainty, and high clinical stakes, within environments shaped by complex human, institutional, and technological factors. Uncritical reliance on algorithmic outputs, automation bias, or excessive deference to computational recommendations can erode clinical judgment, weaken vigilance, and strain the clinician-patient relationship(9). The ethical challenge extends beyond questions of technical accuracy or performance to encompass the preservation of moral agency, professional accountability, and compassionate care.

Across task force deliberations, there was consistent agreement that AI systems in neurosurgery should function as “augmented intelligence” in relation to clinical decision-making. Computational tools may inform, support, or enhance judgment, but they cannot replace professional responsibility, empathy, or ethical reasoning. Neurosurgeons remain ultimately accountable for patient care. At the same time, inappropriate delegation of judgment to machines, particularly in diagnostic, prognostic, or treatment-planning contexts, risks deskilling, cognitive offloading, and diminished preparedness when systems fail, are unavailable, or produce misleading outputs.

The task force also recognised that computational neurosurgery is not limited to decision-support technologies. Implanted and semi-autonomous neural devices, including open- and closed-loop neurostimulation systems and brain-computer interfaces, may operate continuously within a patient's body, performing therapeutic or modulatory functions rather than assisting clinician judgment at a single point of care. Even in these settings, ethical responsibility and clinical accountability cannot be displaced. Decisions regarding deployment, configuration, monitoring, adjustment, and discontinuation remain clinical acts for which neurosurgeons and healthcare teams retain responsibility. Augmentation of function does not entail abdication of accountability.

Safe and ethical clinical integration therefore requires more than retrospective performance metrics alone. Robust evaluation must include preclinical testing, prospective clinical assessment, and ongoing real-world monitoring across diverse patient populations and care settings. Equally important is structured ethical reflection on the purposes, risks, and limits of deployment. Clear articulation of indications, contraindications, uncertainty, and known failure modes is essential, as are mechanisms that allow clinicians to question, override, or disengage from computational recommendations when appropriate.

Institutional responsibilities are central to this process. Healthcare systems must provide not only technical infrastructure, training, and oversight, but also governance arrangements that integrate ethical deliberation into clinical workflows and continuing professional development. Responsible deployment of computational systems in neurosurgery depends on aligning technical validation with sustained ethical appraisal, rather than treating ethical considerations as secondary or retrospective concerns. In short, ethics must be embedded at all levels and across all aspects of computational neurosurgery from the very inception of its development and clinical use. Ethics is not a hurdle to be jumped, a requirement to be considered at the initiation of a neurosurgical intervention and then forgotten, nor is it just an afterthought after the procedure is completed. It is integral to the entire neurosurgical process.

Informed consent presents an additional and evolving challenge. While many current AI applications are treated as clinical tools that do not require separate consent, this assumption becomes increasingly fragile as systems grow more adaptive, autonomous, or influential in clinical decision-making. Traditional consent models, often framed as one-time transactional events, may be inadequate for technologies that evolve over time, draw on external data streams, or modify their behaviour through learning(10). Meaningful consent in computational neurosurgery may therefore require ongoing disclosure, patient education, and opportunities for reassessment or withdrawal of consent. Importantly, this includes discussion of risks that extend beyond conventional neurosurgical harms, such as cybersecurity vulnerabilities, threats to data or mental privacy, and the potential for unintended cognitive or behavioural effects alongside physical risks. If research is involved, the fullest disclosure and clearest consent are required.

We note that what is required to obtain a valid informed consent can be complex, especially when new technologies are the subject. There is an extensive ethical and legal literature on what is required for informed consent in a wide range of situations, which will need to be expanded to take into account novel issues raised by computational neurosurgery.

5.2 Neural Data, Mental Privacy, and Data Stewardship

Computational neurosurgery depends on the use of highly sensitive forms of data, often collected at scale and over extended periods of time. These include neuroimaging, electrophysiological recordings, behavioural measures, intraoperative data, and, increasingly, neurogenomic and other multi-omics information. Unlike many other categories of medical data, neural data are closely entangled with features of mental life. They may reveal, or enable inferences about, cognition, intention, vulnerability, preference, or identity in ways that reach beyond conventional clinical descriptors(11). For this reason, the task force agreed that neural data warrant heightened ethical and legal protection.

There was broad consensus on the importance of robust data stewardship across the entire data lifecycle, encompassing privacy, security, consent, transparency, and responsible sharing. At the same time, deliberations revealed thoughtful disagreement about how best to frame these protections conceptually. Some participants favoured the explicit language of neurorights, drawing on emerging international discussions of possible novel rights such as mental privacy and cognitive liberty. Others cautioned that rights-based terminology may be legally contested or culturally variable, and argued for grounding protections within established frameworks of medical ethics. Despite these differences, there was strong agreement that the underlying values, respect for the mind, freedom of thought and speech, patient autonomy, dignity, and their freedom from exploitation, are non-negotiable, regardless of the language used to express them.

Mental privacy emerged as a particularly salient concern. Advances in neuroimaging, brain-computer interfaces, and predictive modelling increasingly allow inferences about mental states, intentions, or vulnerabilities that patients may neither anticipate nor explicitly disclose. Protecting against non-consensual access, misuse, surveillance, profiling, or discriminatory application of neural data is therefore essential. Consent processes must be meaningful and proportionate to risk, addressing not only immediate clinical use but also secondary applications such as algorithm development, data sharing, and future repurposing.

Responsible stewardship also requires careful attention to data provenance and representativeness. Transparent documentation of data sources, preprocessing methods, population characteristics, and known limitations is central to scientific integrity and to mitigating bias. Particular care is needed for neurogenomic and multi-omics data, where incidental findings, predictive implications, and risks of stigma or discrimination may arise. Governance mechanisms such as data access committees, ethics review processes, and patient or community representation play an important role in aligning data practices with patient values and the public interest.

Finally, the task force emphasised that data stewardship in computational neurosurgery cannot be separated from questions of global justice. Data practices that extract value from low- and middle-income settings while concentrating benefits in high-resource environments risk undermining trust and ethical legitimacy. Equitable collaboration, fair benefit-sharing, capacity building, and inclusive participation in governance are therefore essential components of responsible data use in this field. Ethics requires that those who bear the burdens of research, such as its risks and harms, must also share its benefits.

5.3 Trust, Accountability, and System-Level Governance

Trust lies at the heart of neurosurgical practice and takes on heightened importance as computational and AI-enabled systems become more deeply integrated into care. This trust operates at several interconnected levels: patients' trust in their clinicians and care teams; clinicians' trust in the reliability, limitations, and appropriate use of computational tools; and wider public trust in the institutions responsible for governing their development and deployment(12). When responsibly designed and used, such systems may strengthen trust by supporting better outcomes and greater transparency. Conversely, trust may be undermined when technologies are perceived as opaque, externally imposed, or when the persons using them are insufficiently accountable or as having a conflict of interest(13). Task force discussions repeatedly underscored the distinction between earned trust, built through evidence, transparency, oversight, and ethical responsibility ("trust me because I will show that you can trust me"), and blind trust ("trust me because I know what is best for you"). Blind trust is especially suspect when the fiduciary has a conflict of interest, for instance, could financially gain or receive enhanced professional recognition from the use of the technology.

Accountability is inseparable from trust and remains one of the most complex ethical and legal challenges in computational neurosurgery. When AI-supported decisions contribute to harm, responsibility must not be obscured or diffused across socio-technical systems. Clear and traceable allocation of accountability among clinicians, healthcare institutions, developers, and regulators is essential, both to prevent moral hazard and to ensure that patients retain meaningful avenues for explanation, remedy, and redress. Computational systems that influence neurosurgical decision-making should not function as unaccountable “black boxes,” whether at the level of clinical use or institutional governance.

Governance responsibilities in computational neurosurgery operate across multiple, interdependent levels. At the level of individual practice, neurosurgeons retain fiduciary obligations toward patients, including duties of care, honesty, transparency, and to obtain informed consent. At the institutional level, hospitals, academic centres, and professional societies are responsible for establishing governance structures, ethics oversight, credentialing processes, training requirements, and quality assurance mechanisms. At the level of public oversight, regulators and legal systems must define standards for validation, approval, liability, and post-market surveillance that are proportionate to clinical risk. Beyond national boundaries, international coordination is required to address cross-border data flows, multinational technology development, and persistent inequities in access and benefit-sharing. Across all of these levels, the task force emphasised the ethical importance of incorporating the perspectives of patients and individuals with lived experience of neural devices or AI-enabled interventions, both to inform design and to ensure governance remains responsive to real-world impacts.

Unlike many traditional medical technologies, computational systems may evolve over time. Models can be updated, performance can drift, and new biases may emerge as systems interact with changing data and clinical contexts. Governance must therefore be conceived as a dynamic process rather than a static framework. Continuous monitoring, incident reporting, audit mechanisms, and periodic review are essential, alongside structures such as registries, custodial bodies, and sustained engagement with affected users. These mechanisms help ensure that oversight evolves in step with scientific evidence, technological capability, and societal expectations.

Taken together, earned trust, clear accountability, and good governance constitute an integrated ethical infrastructure for computational neurosurgery. Without accountability, trust, whether on the part of patients, clinicians, or society, cannot be sustained. Without trust, governance lacks legitimacy. And without governance, innovation risks advancing more quickly than ethical and legal safeguards can adapt. Addressing these dimensions collectively is therefore essential to ensuring that computational neurosurgery advances patient welfare, maintains professional integrity, and retains public confidence.

6. Education, Training, and Assessment in the Era of Computational Neurosurgery

The rise of computational neurosurgery calls for a rethinking of how neurosurgeons are educated, trained, assessed, and credentialed. Unlike earlier technological advances, AI-enabled systems increasingly shape diagnostic reasoning, procedural planning, and intraoperative decision support. As a result, competence in neurosurgery can no longer be understood purely in technical terms. It now also encompasses epistemic understanding, ethical judgment, and an ability to remain professionally accountable in environments where computational systems actively influence clinical thinking.

6.1 Beyond Tool Familiarity: Cognitive and Ethical Competence

Task force discussions repeatedly emphasised that education in computational neurosurgery must go well beyond familiarity with specific tools or interfaces. Neurosurgeons and trainees need a working understanding of how algorithmic systems reason, where their limitations lie, and how bias, uncertainty, and failure can arise. Without this conceptual grounding, there is a risk that AI-generated outputs will be accepted uncritically, leading to automation bias, erosion of independent clinical reasoning, and inappropriate delegation of judgment.

This educational responsibility extends to AI-enabled and implanted neural devices. As such technologies become more prevalent, practitioners must be prepared to recognise risks related not only to physical safety, but also to mental privacy, freedom of thought, cybersecurity vulnerabilities, and the possibility of unintended cognitive or behavioural effects. Devices that operate continuously within a patient's body demand a level of ethical attentiveness that cannot be reduced to technical proficiency alone.

Ethics, in this context, cannot be confined to a single course or introductory module. Participants stressed that ethical reflection should be embedded longitudinally throughout training and aligned with the full lifecycle of AI and neurotechnology, from funding of research, data generation and system design to deployment, post-market surveillance, and long-term use. As neural systems increase in sophistication and autonomy, education must also prepare practitioners to recognise tensions that may arise between patient welfare and external influences, such as commercial incentives, data extraction practices, or optimisation goals embedded in system design. Treating these considerations as integral to professional competence is essential if ethical responsibility, including protection of patients' mental privacy and autonomy, is to be understood as a core component of clinical excellence rather than an external constraint.

6.2 Deskilling, Skill Preservation, and Augmentation

Concerns about deskilling emerged frequently during task force discussions, particularly in relation to foundational neurosurgical competencies such as anatomical knowledge, clinical reasoning, operative planning, and the management of complications. While computational tools may enhance performance, excessive reliance on them, especially in the absence of deliberate skill-preservation strategies, could weaken clinical resilience in situations where AI systems fail, are unavailable, or generate misleading outputs.

Educational frameworks should therefore be designed to ensure that computational tools augment rather than replace human expertise. Simulation platforms, AI-assisted rehearsal, and decision-support systems may serve as valuable training aids, but they must be paired with clear expectations for independent competence. Assessment and credentialing processes should not assume continuous access to AI systems, and core clinical abilities must remain demonstrably intact even when technological support is absent.

At the same time, augmentation in computational neurosurgery is not limited to individual clinicians. Emerging models of AI-augmented multidisciplinary teams suggest that computational systems may enhance collective expert decision-making by supporting data synthesis, pattern recognition, and shared situational awareness across specialties, potentially strengthening rather than diminishing expert judgment when appropriately governed and critically applied(14).

6.3 Assessment Integrity in the GenAI Era

The widespread availability of generative AI introduces new challenges for assessment and credentialing. When AI systems can produce polished reports, plans, or analyses with minimal effort, traditional assessment methods that focus solely on final outputs become less reliable indicators of individual competence(15).

Task force members highlighted the need for assessment approaches that capture learning and judgment over time, rather than relying on isolated artefacts. Programmatic assessment models, which draw on longitudinal evidence across multiple contexts, were seen as particularly well suited to this environment. Such approaches can emphasise reflective practice, transparency about AI use, and the ability to critically appraise algorithmic outputs.

Several contributors noted the importance of assessing not only what was produced, but how it was produced. This includes disclosure of AI assistance, justification for tool selection, and demonstration of clinical reasoning independent of computational outputs. Addressing what has been described as an emerging “attribution gap”, where authorship, reasoning, and accountability become blurred, is essential for maintaining trust in professional credentialing and sustaining public confidence in AI-augmented neurosurgical practice.

7. Global Equity, Access, and the Risk of Technological Stratification

Computational neurosurgery is often presented as a means of widening access to expertise and narrowing longstanding gaps in care. In principle, data-driven tools could help extend advanced neurosurgical knowledge to regions where specialist resources are scarce. At the same time, task force discussions repeatedly underscored a central concern: without deliberate attention to equity, these same technologies may deepen, rather than reduce, existing global disparities. It merits noting, regarding research carried out in developing countries, that ethics demands that those who bear the burdens of research, for instance by serving as subjects, must also share in its benefits.

7.1 Opportunity for Capacity Building

When responsibly designed and deployed, computational tools may support neurosurgical capacity in low- and middle-income settings in meaningful ways. AI-enabled diagnostic support, tele-mentoring platforms, simulation-based training, and decision-support systems have the potential to reduce dependence on local subspecialist availability and to strengthen clinical decision-making where human resources are limited. Such approaches may enable earlier diagnosis, more effective triage, and better-informed referral pathways, contributing to improved outcomes even in resource-constrained environments.

Importantly, these benefits are not automatic. Their realisation depends on contextual adaptation, local relevance, and alignment with existing health-system needs. Tools designed without attention to infrastructure, workflow, or cultural context may fail to deliver promised gains, regardless of technical sophistication.

7.2 Risks of Digital Extraction and Unequal Benefit

Alongside these opportunities, task force members expressed concern about patterns of “ethics dumping” and digital extraction. In such scenarios, data from resource-limited settings are used to train academic or commercial models without meaningful return in the form of capacity building, shared governance, or local benefit. There is also a risk that algorithms developed primarily from high-income population data may perform poorly when deployed elsewhere, reinforcing inequities rather than alleviating them.

Several contributors emphasised that global equity cannot be reduced to access alone. Genuine equity requires participation in governance, involvement in validation and evaluation, and fair distribution of benefits arising from innovation. Without intentional design and accountability, computational neurosurgery risks entrenching a two-tiered system, one in which high-resource settings capture the advantages of technological progress while others absorb its risks without corresponding gains.

8. Environmental and Infrastructural Sustainability

During task force discussions, increasing attention was given to the environmental and infrastructural implications of computational neurosurgery. Large-scale computational systems rely on substantial energy consumption, specialised hardware, and complex global supply chains(16). Although these considerations are often peripheral in discussions of medical AI, participants recognised them as ethically relevant, particularly within healthcare systems that are themselves committed to minimising harm and responsibly stewarding finite resources.

Sustainability was therefore framed not as an external constraint, but as part of the broader ethical context in which computational neurosurgery operates. Contributors encouraged awareness of environmental impact alongside clinical benefit, including attention to model efficiency, hardware lifecycles, and procurement practices. While no single technical solution was endorsed, the task force emphasised that responsible innovation should include reflection on resource use and long-term infrastructural consequences, rather than treating computational capacity as an unlimited or ethically neutral input.

An important ethical issue relating to minimising harm was the potential of the advancement of computational simulations and refinement of bio-models and bio-computation to reduce and, where scientifically feasible, replace the use of animals in neurosurgical research and training.

9. Living Governance and Custodial Responsibility

The task force consistently emphasised that computational neurosurgery should not be understood as a discrete technological shift, but as an ongoing sociotechnical transformation. Given the pace at which computational methods, neuro-technologies, and regulatory landscapes continue to evolve, static ethical guidance risks becoming outdated or misaligned with practice(2). For this reason, participants supported the idea of a living Declaration, capable of evolving over time in response to new evidence, emerging risks, and changing societal expectations.

To support this approach, the task force endorsed the establishment of a multidisciplinary custodian group charged with the long-term stewardship of the Declaration. This group would be responsible for periodic review, public engagement, and ongoing alignment with developments in science, ethics, and law. Importantly, such custodianship was not conceived as an enforcement mechanism. Rather, it was understood as a normative role, aimed at sustaining transparency, facilitating dialogue, and fostering coherence across related global initiatives, while preserving the Declaration's ethical intent as the field continues to mature. The structure, responsibilities, and operating principles of this custodianship model are described in detail in Supplementary Material 1 (Custodianship Charter).

10. The Declaration

The Declaration of Sydney, presented below, represents the Task Force's effort to distil these complex considerations into a concise set of shared principles suitable for public commitment, while acknowledging that ongoing scholarly work, debate, and revision will remain necessary as the field continues to evolve.

Declaration of Sydney on the Ethical Use of Artificial Intelligence in Neurosurgery

Definition of Computational Neurosurgery

Computational neurosurgery is a transdisciplinary field that integrates computation, artificial and augmented intelligence, data science, and neurotechnology across the neurosurgical care pathway. Its applications include diagnosis, risk stratification and screening, clinical decision support, surgical planning, intraoperative support, postoperative management, research, education, and training. Computational neurosurgery aims to improve basic understanding, clinical decision-making, and patient outcomes in neurosurgery.

Goal of the Declaration

The goal of this Declaration is to establish a shared ethical foundation for computational neurosurgery. Without limiting the breadth of this goal, the Declaration seeks to embed ethical, responsible, and human-centred development and use of computational and AI-enabled technologies across all stages and levels of computational neurosurgery. Achieving this goal will help ensure that innovations advance beneficial patient care while minimising risks and harms, preserving professional accountability, and sustaining earned public trust.

Preamble

We recognise that:

a. understanding the complexity of the human nervous system and its disorders remains among the most significant challenges in medicine and science;
b. the convergence of computational science, artificial intelligence, neurotechnology, neuroimaging, data science, robotics, and clinical neurosurgery has created unprecedented opportunities to advance brain health and improve diagnostic and therapeutic precision;
c. computational neurosurgery has emerged as a multidisciplinary field integrating scientific discovery with clinical practice to enhance patient outcomes and transform neurosurgical knowledge;
d. computational technologies raise profound ethical issues and legal ramifications, including risks to privacy, autonomy, equity, trust, and the integrity of clinical decision-making;
e. there is a need to ensure that computational advances are developed and deployed in ways that support clinical judgement, minimise risks and harms to patients, uphold patient safety, human dignity, and earn public trust;
f. there is a moral imperative to ensure that AI-enabled innovations reduce global disparities in access to neurosurgical care;
g. computational neurosurgery should be used to augment human expertise, and not to replace it.

We therefore present the Declaration of Sydney to identify and articulate shared values, guide responsible innovation, and establish an ethical foundation for the evolving field of computational neurosurgery.

Articles of the Declaration

Article 1 – Primacy of Patient-Centred Care and Respect for Human Dignity

The use of computational and AI-enabled tools in neurosurgery must always prioritise the avoidance of harm and the provision of beneficent and compassionate care to the individual patient, including, but not limited to, their safety, welfare, autonomy, and respect for their dignity.

Article 2 – Ethical Use of Artificial Intelligence in Neurosurgery

Computational methods, including AI, and their applications must be used responsibly, in conformity with established ethical principles, sound clinical judgment, and in accordance with the principle “First, do no harm”.

Article 3 – Human Oversight and Clinical Accountability

Computational tools should support, not replace, professional judgment and accountability.

Article 4 – Data Rights, Privacy, and Mental Integrity

The collection, use, and analysis of neurosurgical and neural data must be undertaken and governed in accordance with legal and ethical frameworks that include requirements of respect for individuals' privacy, autonomy, and mental integrity, and mandate robust protections against misuse, discrimination, or exploitation.

Article 5 – Data Stewardship, and Informed Consent to Data Collection, Storage, and Use

All data collected and used in computational neurosurgery must be collected, stored, shared, and processed within rigorous frameworks that ensure accuracy, responsible governance, transparency of use, and compliance with ethical and legal requirements.

Article 6 – Accountability, Traceability, and Explainability

Computational systems should provide clear documentation of their capabilities and limitations, and, where possible, explainable decision pathways sufficient to support clinical understanding, accountability, and earned trust.

Article 7 – Responsible Development, Validation, and Deployment

AI systems used in neurosurgery must undergo rigorous scientific development, appropriate clinical validation, risk-aware deployment (including minimising ethical risks), and ongoing monitoring throughout their lifecycle.

Article 8 – Scientific Integrity and Research Responsibility

Research in computational neurosurgery must uphold high standards of scientific rigour, reproducibility, transparency, and ethical conduct, with responsible publication and independent evaluation.

Article 9 – Equity, Access, and Global Neurosurgery

Advances in computational neurosurgery should promote equitable access to care, support global capacity building, and contribute to reducing disparities in access to neurosurgical services, worldwide.

Article 10 – Education, Training, and Skill Preservation

AI-enabled educational tools should strengthen neurosurgical training while preserving foundational competencies, critical reasoning, professional autonomy, and earned trust in clinical expertise.

Article 11 – Transdisciplinary and Consumer Collaboration and Shared Governance

Ethical development and deployment of computational neurosurgery should involve collaboration among clinicians, scientists, engineers, ethicists, policymakers, patients, and society.

Article 12 – Sustainability

The development, validation, and deployment of computational neurotechnologies should be environmentally responsible and evaluated to minimise harmful impact on the planet's ecosystems.

Article 13 – Support of Simulation

The advancement of computational simulations and refinement of biomodels and biocomputation should be actively supported and prioritised to reduce and, where scientifically feasible, replace the use of animals in neurosurgical research and training.

Article 14 – Brain-Computer Interface Governance

The development, validation, and application of brain-computer interfaces (BCIs) and other forms of neurotechnology to advance patient care require informed consent for surgical intervention, rigorous regulatory oversight, and strong safeguards to prevent non-consensual data extraction, processing, and misuse, including manipulative, coercive, malevolent, or criminal interference.

Article 15 – Computational Neurosurgery and Human Enhancement

The neurosurgical implantation or application of computational and neurotechnological systems not required as therapy, but intended to enhance normal brain function, requires careful ethical, legal, and societal consideration prior to implementation.

Final Clause and Pledge

We, the undersigned, gathered at the 1st World Conference of Computational Neurosurgery in Sydney, Australia, on the 15th February 2026, affirm our shared commitment to the ethical, responsible, and human-centred advancement of computational and AI-enabled technologies in neurosurgical medical practice and research.

By endorsing this Declaration, we each pledge to uphold its principles within our institutions, practices, and communities; to promote patient welfare, professional integrity, and earned public trust; and to steward the continued ethical evolution of computational neurosurgery for the benefit of present and future generations.

In witness of this commitment, we hereby endorse and sign the "Declaration of Sydney on the Ethical Use of Artificial Intelligence in Neurosurgery".

11. Discussion and Limitations

This paper represents an early, coordinated attempt to articulate ethical, legal, and professional responsibilities for the use of computational and AI-enabled technologies in neurosurgery at an international level. By pairing a concise public pledge with a more detailed scholarly analysis, the Declaration of Sydney is intended to serve both as a shared ethical statement for the field and as a foundation for deeper reflection on implementation, governance, and future inquiry.

A defining strength of this initiative is its deliberate separation between these two functions. The Declaration itself is intentionally brief, principle-driven, and accessible, designed for public endorsement and broad professional uptake. The accompanying preprint, by contrast, preserves the complexity of the task force's discussions, including areas of uncertainty, disagreement, and ongoing debate. This structure allows the field to articulate common commitments without flattening important differences or prematurely closing questions that remain unsettled.

Throughout the drafting process, several areas of productive tension became apparent and are acknowledged rather than resolved in this work. These include differing views on whether emerging concepts such as neurorights should be foregrounded explicitly or situated within existing human rights and medical ethics frameworks; varying perspectives on how precautionary governance should be balanced against innovation and clinical flexibility; and uncertainty about how informed consent should evolve as AI systems move from passive tools toward more adaptive and influential roles in care. Rather than advancing a single doctrinal position, the Declaration seeks to identify shared ethical responsibilities while allowing space for contextual interpretation and future refinement.

This work also has clear limitations. Although the task force brought together a wide range of disciplinary and geographic perspectives, it does not fully capture the diversity of global neurosurgical practice or patient experience. Voices from low-resource settings, Indigenous communities, and patient advocacy groups remain insufficiently represented and should be more actively engaged in future revisions and custodial processes. In addition, the principle-based nature of the Declaration means that it does not address jurisdiction-specific legal requirements, which may limit its immediate applicability within certain regulatory environments. Finally, the pace of technological change in computational neuroscience and neurotechnology means that specific risks, examples, or governance mechanisms discussed here may evolve more quickly than formal publication timelines.

It is also important to emphasise what this work does not claim to be. The Declaration of Sydney is not a regulatory instrument, a clinical guideline, or a legal standard of care. Its influence rests on ethical legitimacy, professional endorsement, and continued engagement rather than formal enforcement. Nonetheless it might prove useful as a source for consideration in relation to future regulatory and other instruments. Whether it proves meaningful will depend on sustained stewardship, transparent mechanisms for revision, and genuine uptake by clinicians, institutions, educators, industry actors, and policymakers. In this sense, the Declaration should be understood not as a conclusion, but as an invitation to ongoing ethical work as computational neurosurgery continues to develop.

12. Conclusions

Computational neurosurgery is entering a decisive phase. The growing integration of artificial intelligence, advanced data analytics, neurotechnology, and human-machine collaboration is changing how neurosurgeons interpret disease, plan interventions, and care for patients. These developments bring genuine promise: the potential for more precise treatment, broader access to expertise, and new scientific insight. At the same time, they raise ethical, legal, and societal questions that cannot be addressed adequately by general-purpose AI frameworks or by technical performance alone.

The Declaration of Sydney reflects a collective commitment to guiding this transformation responsibly. It affirms that computational systems in neurosurgery must remain anchored in patient welfare, professional accountability, and respect for human dignity. It emphasises that AI should support, rather than displace, clinical judgment; that neural data and mental life warrant particular care and protection; and that innovation must proceed alongside transparency, ethical reflection, and attention to global equity.

By coupling a concise public pledge with a detailed scholarly analysis, this initiative provides both a clear statement of shared values and a foundation for continued ethical engagement. The Declaration is intentionally not definitive or closed. Instead, it is designed to function as a living framework, one that can evolve as technologies mature, evidence accumulates, and societal expectations change.

Ultimately, the trajectory of computational neurosurgery will not be shaped by technical capability alone. It will be shaped by the choices made about how these tools are designed, governed, and used in practice. The Declaration of Sydney invites the global neurosurgical community to engage in that stewardship together, ensuring that computational advances serve patients, respect vulnerability, and earn trust over time.

References:

1. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. Nature medicine. 2019;25(1):44–56.
2. Guidance W. Ethics and governance of artificial intelligence for health. World Health Organization. 2021:1–165.
3. Clausen J. Ethical brain stimulation–neuroethics of deep brain stimulation in research and clinical practice. European Journal of Neuroscience. 2010;32(7):1152–62.
4. Leonelli S. Data-centric biology: A philosophical study: University of Chicago Press; 2019.
5. Senders JT, Arnaout O, Karhade AV, Dasenbrock HH, Gormley WB, Broekman ML, et al. Natural and artificial intelligence in neurosurgery: a systematic review. Neurosurgery. 2018;83(2):181–92.
6. OECD. AI in Health: Huge potential, huge risks. Paris: OECD Publishing; 2024.
7. Di Ieva A, Suero Molina E, Somerville MA, Beheshti A, Staartjes VE, Serra C, et al. Declaration of Computational Neurosurgery. In: Di Ieva A, Suero Molina E, Liu S, Russo C, editors. Computational Neurosurgery. Cham: Springer Nature Switzerland; 2024. p. 11–20.
8. Association WM. Declaration of Helsinki: medical research involving human participants. World Medical Association. 2024.
9. Lyell D, Coiera E. Automation bias and verification complexity: a systematic review. Journal of the American Medical Informatics Association. 2017;24(2):423–31.
10. Food US, Administration D. Cybersecurity in Medical Devices: Quality System Considerations and Content of Premarket Submissions: Guidance for Industry and Food and Drug Administration Staff. Silver Spring, MD: U.S. Food and Drug Administration; 2025.
11. Ienca M, Andorno R. Towards new human rights in the age of neuroscience and neurotechnology. Life sciences, society and policy. 2017;13(1):5.
12. O'neill O. Autonomy and trust in bioethics: Cambridge University Press; 2002.
13. Matulionyte R, Suero Molina E, Di Ieva A. Neurosurgery, Explainable AI, and Legal Liability. In: Di Ieva A, Suero Molina E, Liu S, Russo C, editors. Computational Neurosurgery. Cham: Springer Nature Switzerland; 2024. p. 543–53.
14. Di Ieva A. AI-augmented multidisciplinary teams: hype or hope? The Lancet. 2019;394(10211):1801.
15. Kasneci E, Seßler K, Küchemann S, Bannert M, Dementieva D, Fischer F, et al. ChatGPT for good? On opportunities and challenges of large language models for education. Learning and individual differences. 2023;103:102274.

16. Strubell E, Ganesh A, McCallum A, editors. Energy and policy considerations for deep learning in NLP. Proceedings of the 57th annual meeting of the association for computational linguistics; 2019.