



## OPEN ACCESS

This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

<sup>1</sup>Department of Neurosurgery, Stanford University School of Medicine, Stanford, CA, USA [ROR::](#)

<sup>2</sup>Nuffield Department of Clinical Neurosciences, University of Oxford, Oxford, UK

<sup>3</sup>Department of Orthopaedic Surgery, University of California, Los Angeles, CA, USA

<sup>4</sup>Department of Medicine, University of California, Los Angeles, CA, USA

<sup>5</sup>Department of Neurosurgery, Northwell, Zucker School of Medicine at Hofstra/Northwell, New Hyde Park, NY, USA

<sup>6</sup>Department of Neurosurgery, University of Buffalo Jacobs School of Medicine and Biomedical Sciences, Buffalo, NY, USA

<sup>7</sup>Department of Neurological Surgery, University of Miami Miller School of Medicine, FL, USA

<sup>8</sup>Department of Radiation Oncology, Stanford University School of Medicine, Stanford, CA, USA

<sup>9</sup>Department of Veteran Affairs, University of California, Los Angeles, CA, USA

<sup>10</sup>Viterbi School of Engineering, University of Southern California, CA, USA

<sup>11</sup>Department of Psychology, University of Amsterdam, Amsterdam, The Netherlands

<sup>12</sup>Department of Philosophy, California State Polytechnic University, Pomona, CA, USA

<sup>13</sup>Department of Public Health, Harvard TH Chan School of Public Health, Harvard University, MA, USA

<sup>14</sup>Premier Science, London, UK  
\*Indicates Co-first authorship

# Artificial Intelligence Across the Surgical Oncology Continuum: Decision Support, Operative Intelligence, and a Translation-First Roadmap

Matthew Abikenari<sup>1,2,\*</sup> , Iman Enayati<sup>3\*</sup>, Kimia Mohseni<sup>4</sup>, Vivek Sanker<sup>1</sup>, Faina Abyazova<sup>5</sup>, Luis Vargas<sup>6</sup>, Vratko Himic<sup>7</sup>, Yijiang Chen<sup>8</sup>, Alisher Baibussinov<sup>9</sup>, Sri Polkampally<sup>1</sup>, Shirley Liu<sup>1</sup>, James Poe<sup>1</sup>, Claire Kim<sup>1</sup>, Danial Davani<sup>10</sup>, Aryan Jain<sup>1</sup>, Rene Freichel<sup>11</sup>, Derek Abikenari<sup>12</sup>, Ahmed Kerwan<sup>13</sup> and Riaz Agha<sup>14</sup>

## ABSTRACT

Artificial intelligence (AI) and machine learning (ML) methods are increasingly being developed for surgical oncology, a domain in which perioperative decision-making must account for uncertainty and substantial variability in patient anatomy, tumor extent, imaging protocols, devices, and operative workflows. Although proof-of-concept performance has been reported across multiple tasks, there remain limited data regarding the extent to which these systems generalize across institutions and whether they confer measurable clinical utility under routine conditions. In parallel with broader advances in medical AI and surgical robotics, there has been a growing effort to operationalize AI/ML across the perioperative continuum. The aim of this narrative review is to define the role of AI/ML in surgical oncology and to delineate the evidentiary and methodological requirements for translation into reliable clinical systems. For clarity, we structure the field into three domains that align with the perioperative workflow: (1) clinical translation and decision support, (2) preoperative planning, and (3) intraoperative navigation with robotic assistance/control. Clinical translation centers on decision support, including perioperative risk stratification, outcome prediction, and workflow-integrated recommendations. Preoperative planning is dominated by imaging-based methods for lesion detection and classification, anatomic segmentation, and image registration. Intraoperative navigation and robotic assistance/control emphasize perception and guidance, such as localization, tissue and instrument tracking, geometric reconstruction, and augmented reality visualization, alongside emerging autonomy-enabling approaches, such as learning from demonstration and reinforcement learning, implemented within supervised human–robot interaction paradigms. Across these domains, we evaluate how data provenance, endpoint selection, external and prospective validation, robustness to dataset shift, uncertainty calibration, interpretability, and safety assurance constrain performance and determine translational readiness.

**Keywords:** Artificial intelligence, Machine learning, Surgical oncology, Perioperative decision-making, Preoperative planning, Intraoperative navigation, Robotic surgery, Clinical translation.

## Introduction

Surgical oncology integrates technically demanding resection and reconstruction with perioperative decision-making constrained by biologic uncertainty and substantial heterogeneity in patient anatomy, tumor extent, prior treatment response, and operative environment. Even when intent is curative, the operative plan is frequently refined intraoperatively based on unanticipated patterns of invasion, occult metastatic disease, tissue planes distorted by neoadjuvant therapy, and evolving physiologic tolerance for blood loss or prolonged anesthesia. In this setting, the margin for error is narrow: oncologic adequacy must be balanced against preservation of function and minimization of morbidity, and these trade-offs must be negotiated using information that is incomplete, distributed across modalities, and variably reliable.

Artificial intelligence (AI) refers broadly to computational methods that perform tasks typically requiring human intelligence; machine learning (ML) denotes a subset of AI in which models learn statistical structure from data to generate predictions or recommendations without explicit rule-based programming. Contemporary AI/ML in perioperative medicine is dominated by data-driven approaches, particularly deep learning, applied to high-dimensional signals such as cross-sectional imaging, digital pathology, endoscopic video, and structured and unstructured clinical data. In surgical oncology, preoperative applications have been shaped by imaging-centered tasks, including lesion detection and characterization, risk stratification, treatment response assessment, and selection of candidates for neoadjuvant or adjuvant strategies. Radiomics and related quantitative imaging methods exemplify this trajectory by converting medical images into features that can be integrated with clinical and molecular variables for prognostic and predictive modeling.<sup>1</sup>

Preoperative planning has also increasingly incorporated anatomic reconstruction and simulation, motivated by the need to anticipate anatomic variants, determine feasible resection planes, and select minimally invasive trajectories that preserve critical structures. Three-dimensional (3D) reconstruction workflows, often combined with AI-assisted segmentation and registration, have been particularly prominent in thoracic surgery, where complex bronchovascular branching

Correspondence to:  
Matthew Abikenari, MS  
mattabi@stanford.edu

Additional material is published  
online only. To view please visit the  
journal online.

Cite this as: Abikenari M<sup>1,2</sup>,  
Enayati P<sup>3</sup>, Mohseni K<sup>4</sup>, Sanker  
V<sup>1</sup>, Ablyazova F<sup>5</sup>, Vargas L<sup>6</sup>, Hemic  
V<sup>7</sup>, Chen Y<sup>8</sup>, Baibussinov A<sup>9</sup>,  
Polkampally S<sup>1</sup>, Liu S<sup>1</sup>, Poe J<sup>1</sup>, Kim  
C<sup>1</sup>, Davani D<sup>10</sup>, Jain A<sup>1</sup>, Freichel  
R<sup>11</sup>, Abikenari D<sup>12</sup>, Kerwan A<sup>13</sup> and  
Agha R<sup>14</sup>

Artificial Intelligence Across the  
Surgical Oncology Continuum:  
Decision Support, Operative  
Intelligence, and a Translation-First  
Roadmap.

Premier Journal of Science  
2026;20:100269

DOI: [https://doi.org/10.70389/  
PJS.100269](https://doi.org/10.70389/PJS.100269)

Peer Review

Received: 16 January 2026

Last revised: 1 March 2026

Accepted: 3 March 2026

Version accepted: 3

Published: 9 March 2026

Ethical approval: N/a

Consent: N/a

Funding: No industry funding

Conflict of Interest Disclosure:  
No pertinent conflicts of interest  
relevant to this manuscript.

Author contribution:

Matthew Abikenari - Writing –  
review & editing, Writing – original  
draft, Visualization, Validation,  
Supervision, Data curation,  
Conceptualization

Guarantor: Matthew Abikenari

Provenance peer review:  
unsolicited

Data availability statement: N/a

Financial Disclosure: None.

and limited operative exposure create dependence on preoperative imaging for safe dissection.<sup>2</sup> Across surgical oncology more broadly, these imaging-based approaches have progressed from offline analysis toward workflow-integrated tools that aim to provide interpretable anatomic abstractions and patient-specific risk estimates at the point of care.

The intraoperative transition from open surgery to endoscopic and robot-assisted platforms has further expanded the volume of machine-readable data while simultaneously reducing direct tactile feedback. Within this environment, computer vision-enabled perception (eg, tissue and instrument localization), workflow modeling (eg, phase recognition), and context-aware guidance have been proposed as mechanisms to reduce operator dependence and improve the consistency of oncologic technique. A systematic review of AI integration in robotic cancer surgery underscores both the breadth of proof-of-concept tasks, ranging from instrument tracking and skill assessment to intraoperative decision support, and the methodological heterogeneity that still characterizes the literature, including variable dataset provenance, inconsistent endpoint definitions, and limited external validation.<sup>3</sup> Parallel advances in multimodal AI seek to fuse complementary signals (eg, imaging, text, and video) to support perioperative decision-making, documentation, and surveillance; these approaches are conceptually attractive in oncologic surgery because key decisions often hinge on concordance across modalities rather than any single data stream.<sup>4</sup>

Despite technical progress, translation of AI/ML systems into reliable clinical support requires evidence that extends beyond retrospective performance within a single institution. Models intended to influence surgical oncology decisions must demonstrate calibration and robustness to dataset shift (systematic differences between training and deployment distributions), integrate uncertainty in ways that are actionable for clinicians, and evaluated against clinically meaningful endpoints under realistic workflow constraints. Importantly, many published systems are derived from retrospective, single-institution datasets and are evaluated using internal validation strategies, limiting inference about transportability and clinical impact.<sup>3</sup> Clinical implementation has been reported but remains comparatively uncommon to the volume of proof-of-concept studies. A large-scale implementation study in colorectal cancer surgery deployed an AI-based prediction model as part of perioperative decision support and evaluated outcomes using routine clinical data capture, illustrating both the potential clinical impact of workflow-integrated prediction and the interpretive constraints of non-randomized implementation designs.<sup>5</sup> Complementary evidence from international external validation of an ML-based risk prediction model for 90-day mortality after gastrectomy for cancer highlights the methodological necessity of testing transportability across institutions and case-mix distributions before widespread adoption.<sup>6</sup>

Clinical deployment also raises domain-specific ethical and epistemic considerations. In surgical oncology, AI-informed recommendations may affect treatment selection, extent of resection, and tolerance for operative risk, thereby shaping patient outcomes- and values-sensitive trade-offs. These realities sharpen questions of equipoise, liability, transparency, and informed consent, particularly when models are embedded in proprietary devices or closed robotic ecosystems and when failure modes are difficult to characterize a priori.<sup>7</sup> Consequently, translation must couple algorithmic performance with safety assurance practices that explicitly define intended use, identify hazards, and evaluate human–system interaction under conditions that approximate routine clinical care.

Intraoperative navigation and robotic assistance/control constitute a distinct translational domain in which performance is constrained by dynamic anatomy, occlusion, deformation, and real-time computational requirements. Representative work includes real-time vascular anatomical image navigation for laparoscopic surgery, in which deep learning is used to recognize key vascular structures in endoscopic video as a substrate for context-aware guidance.<sup>8</sup> Optical stereotactic navigation has also been explored in rectosigmoid cancer surgery using deep learning, supported by 3D modeling to improve localization and support more standardized dissection strategies.<sup>9</sup> In robot-assisted settings, safety-oriented perception has been approached through AI-based hazard detection methods that attempt to identify high-risk configurations of instruments and tissue interaction, reflecting a shift toward proactive risk mitigation rather than post hoc error attribution.<sup>10</sup> There are still many unresolved issues. One of the major surgical complications, with higher occurrence in cancer patients, is intraoperative hemorrhages, which if detected early, can be more efficiently controlled. Aim: This paper proposes a hazard detection system which incorporates the advantages of both Artificial Intelligence (AI For oncologic tasks in which decision thresholds are tightly coupled to margins, uncertainty-aware modeling is increasingly recognized as critical; uncertainty estimation has been evaluated, for example, in intraoperative margin detection using mass spectrometry to better characterize confidence in tissue classification.<sup>11</sup> In parallel, prospective evaluations of non-contact, AI-assisted intraoperative 3D navigation technologies in lung cancer surgery illustrate continued interest in reducing friction in intraoperative data acquisition while maintaining surgeon autonomy.<sup>12</sup>

Collectively, the contemporary literature suggests that AI/ML methods can enable consistent extraction of clinically relevant signals from heterogeneous perioperative data, but also that translational readiness is principally determined by methodological rigor and deployment context rather than by headline accuracy. Surgical oncology provides a stringent test bed for clinical AI: models must generalize across variable anatomy and imaging protocols, support de-

cisions that are safety-critical and time-constrained, and remain reliable when confronted with unexpected intraoperative conditions. In this narrative review, we define the role of AI/ML in surgical oncology and delineate the evidentiary and methodological requirements for translation into reliable clinical systems. For clarity, we structure the field into three domains aligned with the perioperative workflow: (1) clinical translation and decision support, (2) preoperative planning, and (3) intraoperative navigation with robotic assistance/control. Across these domains, we evaluate how data provenance, endpoint selection, external and prospective validation, robustness to dataset shift, uncertainty calibration, interpretability, and safety assurance constrain performance and determine translational readiness.

Lastly, to maintain the surgical oncology focus of this review, the remainder of this review schematizes the capabilities of ML systems against perioperative cancer-related decisions and actions such as oncologic surgery candidacy and risk tailoring, imaging-based characterization/planning for tumor and critical structure relationships, and intraoperative perception/guidance for margin assessment and complication avoidance. Where clinical AI frameworks are cited in general terms (eg, reporting, risk of bias, device control), they have been included insofar as they provide translation-relevant constraints for perioperative oncologic systems.

#### **Methodological Transparency**

This review was designed with the goal of translational readiness in mind rather than maximal topical relevance. We conducted a structured literature review according to the following methods. We searched PubMed/MEDLINE, Embase, IEEE Xplore, and the Cochrane Library for English-language publications from January 2010 to the most recent date before submission. Search concepts were designed as follows: (“surgical oncology” or procedure- or organ-specific cancer surgery) AND (“artificial intelligence,” “machine learning,” “deep learning,” “computer vision,” “decision support,” “robotic surgery,” “augmented reality,” “navigation,” “risk prediction”). We also searched reference lists of relevant systematic reviews and landmark studies to ensure that any pertinent literature would not be missed. Studies were eligible for inclusion if they developed, validated, or clinically used AI/ML systems that were intended to affect or were used for perioperative decision-making in cancer surgery, preoperative planning, or intraoperative navigation in cancer surgery. We did not include non-oncologic applications unless they provided a generalizable translational framework that is pertinent to cancer surgery. Two reviewers screened the abstracts for relevance. The abstracts were then used to screen the full text for inclusion based on the study’s clear intention to use the AI/ML system for the aforementioned applications. Data were abstracted for dataset source and representativeness, labeling strategy and reference standards,

internal or external or prospective validation, calibration and uncertainty estimation, decision-analytic evaluation, workflow integration, and safety features. To minimize bias in the review process and to improve the interpretability of the study results for the reader in terms of the applicability to the development of AI/ML systems for cancer surgery, the review is designed to assess the applicability and the risk of bias in the study results according to the domains in the PROBAST/PROBAST+AI criteria (participants, predictors, outcomes, analysis) while focusing on the reporting criteria in the TRIPOD+AI, CONSORT-AI, SPIRIT-AI, DECIDE-AI criteria as applicable.

Lastly, to make the evidence assembly process explicit and reproducible, we provide a summary of identification, screening, eligibility, and inclusion criteria, including details of sources searched, de-duplication strategy, inclusion and exclusion criteria, and final included items by domain. For ease of reading and conciseness of the main text, we have placed expanded framework tables, such as comprehensive standards tables and expanded failure mode tables, throughout the manuscript.

#### **Clinical Translation and Decision Support**

To help evidence maturity become clearer and avoid over-reliance on narrative emphasis, we provide domain-specific evidence maps that provide a summary of representative studies in surgical oncology (Tables 1–4). Each table includes information on dataset provenance (single-site vs multi-site, registry vs electronic health record [EHR], imaging), validation approach (internal, external, internal–external, prospective), calibration and handling of uncertainty, decision-analytic evaluation, and, most importantly, what specific clinical action is intended to be triggered by model output (eg, pathway enrollment, increased surveillance, changes in surgery approach, increased margins, sampling).

#### **Scope and Clinical Tasks**

In surgical oncology, *clinical translation* denotes the pathway by which AI and ML methods move from retrospective proof-of-concept toward deployment as workflow-integrated tools that measurably improve patient care under routine conditions.<sup>5,13</sup> This translational interface is dominated by *clinical decision support* (CDS), defined here as computer-based systems that provide patient-specific assessments or recommendations intended to aid clinical decision-making.<sup>14,15</sup> Decisions such as operative candidacy, extent and timing of resection, perioperative optimization, postoperative triage, and surveillance planning directly shape both short-term morbidity and long-term oncologic benefit.<sup>5,16,17</sup> At this stage, model outputs must be not only accurate but also actionable, interpretable, and deliverable at the moment a decision is made, properties repeatedly associated with CDS effectiveness in broader health systems research and explicitly emphasized in surgical risk calculator design.<sup>14,18</sup>

**Table 1 | Surgical oncology AI operational taxonomy with translation maturity rubric**

Operational Tier	Module or Task	Primary Clinical Product	Typical Model Family	Translation Maturity Rubric for This Module	Minimum Translation Bar	Dominant Harm If It Fails
<b>Decision Support</b>	Risk prediction, complication prediction, therapy guidance	Decision-relevant outputs for tumor board and pre-op planning	Supervised ML, deep nets, survival models	Retrospective discrimination → internal validation → external validation → prospective workflow study → impact study with patient or process endpoints → monitoring	External validation and prospective assessment should become routine; interface and workflow integration are scientific outputs	Silent distribution-shift failure, false reassurance, inappropriate plan changes
<b>Pre-op Planning</b>	Classification	Actionable probabilities for diagnosis or risk phenotype	CNN classifiers, multimodal models	Benchmark performance → multi-site generalizability → calibration and uncertainty → clinical impact on planning decisions	Outputs must align with decisions rather than outcomes; handle missing and imbalanced data	Overconfident wrong probabilities bias operative aggressiveness
<b>Pre-op Planning</b>	Detection	Lesion and landmark localization that reduces search burden	Detection networks, regression heads	Phantom or single-site tests → robustness across scanners and protocols → error propagation testing into segmentation and registration	Must be treated as upstream dependency that can cascade into later modules	Cascading localization error creates systematic downstream overlay error
<b>Pre-op Planning</b>	Segmentation	Operative geometry: tumor margins, organs at risk, vessels, corridors	U-Net variants, encoder-decoder nets	Accuracy on curated sets → robustness to label heterogeneity → clinically relevant geometry metrics	Annotation scalability is the binding constraint; weak and self-supervised are necessities	Wrong geometry yields unsafe distances, corridors, or resection maps
<b>Pre-op Planning</b>	Registration	Alignment across modalities, timepoints, and pre-op model to intra-op view	Deformable registration, learned deformation fields	Retrospective registration error → stress tests under deformation → downstream overlay safety tests	Confident subtle error is worse than no system; overlays require trust	Confident wrong overlay guides the surgeon toward danger
<b>Intra-op Navigation</b>	Shape instantiation	Recover usable 3D geometry from partial or 2D views	3D reconstruction, surface models	Lab or phantom → robustness under deformation and occlusion → failure-behavior characterization	Should be evaluated by modes and speed of failure, not frame accuracy	Wrong 3D model leads to wrong collision avoidance or distance-to-critical-structure
<b>Intra-op Navigation</b>	Endoscopic navigation	Depth, pose, mapping in non-rigid spaces	Depth estimation, odometry, SLAM, self-supervised	Phantom or limited datasets → domain shift robustness → OOD detection and safe degradation	Systems must be tested for failure behavior across environments	Catastrophic navigation drift with false confidence
<b>Intra-op Navigation</b>	Tissue tracking	Continuous lock on regions of interest under motion and deformation	Tracking-by-detection, online updating	Controlled settings → robustness in real OR conditions → governance for non-stationary updating	Tight constraints during low confidence; governance challenges for online learning	Tracker drift causes unsafe AR and automation bias
<b>Intra-op Navigation</b>	Augmented reality overlays	Intuitive overlay of pre-op models onto operative field	AR overlay pipelines	Lab alignment → non-rigid organ overlay validation → human factors evaluation	Periodically incorrect overlay is dangerous; uncertainty must be represented	Surgeon overtrusts an intermittently wrong overlay

**Table 1 | Surgical oncology AI operational taxonomy with translation maturity rubric**

Operational Tier	Module or Task	Primary Clinical Product	Typical Model Family	Translation Maturity Rubric for This Module	Minimum Translation Bar	Dominant Harm If It Fails
<b>Robotics and Autonomy</b>	Instrument segmentation and tracking	Tool localization as prerequisite for safe automation	U-Net tools, tracking smoothing	Benchmark → robustness under blood, smoke, occlusion → integrated safety tests	Perception is prerequisite for bounded autonomy	Mislocalized tool leads to unsafe action
<b>Robotics and Autonomy</b>	Learning from demonstration	Surgical primitives and reusable motion policies	Imitation learning, primitive decomposition	Benchmark datasets → OOD detection → safe handoff	Must detect edge conditions and transfer control swiftly to human	Unrecognized edge case leads to unsafe continuation
<b>Robotics and Autonomy</b>	Reinforcement learning	Narrow task-level autonomy via policies	RL with sim-to-real	Simulation → transfer learning → robust safety constraints	Sim-to-real transfer is a major practical concern; autonomy is bounded	Policy fails at rare events and creates instantaneous harm
<b>Robotics and Autonomy</b>	Human–robot interface	Command-intention detection and safe interaction	Multimodal intent models	Lab tests → OR noise and occlusion validation → extremely low tolerance error audits	Low tolerance for errors; single wrong command can be dangerous	Erroneous command triggers dangerous actuation

**Table 2 | Integrity-chain audit table: how upstream failures propagate to downstream harms**

Upstream Module	Downstream Module	Dependency You Claim	How Failure Propagates	Concrete Harm Example	Translation Mitigation You Advocate
<b>Detection</b>	Segmentation	Landmark and ROI constraints	Missed or shifted detection biases segmentation seed and geometry	Incorrect tumor margin near critical structure	Evaluate cascade explicitly, not modules in isolation
<b>Segmentation</b>	Registration	Geometry informs alignment and overlays	Wrong anatomy geometry biases deformation alignment	AR overlay is systematically shifted despite looking plausible	Treat overlay trustworthiness as safety-critical
<b>Registration</b>	AR overlays	Overlay depends on correct alignment	Small registration error becomes high-confidence visual guidance	Confident wrong overlay increases surgeon confidence wrongly	Report uncertainty; fail safely when confidence is low
<b>Tracking</b>	AR overlays	Overlay stability depends on tracking	Tracker drift makes overlay intermittently correct and intermittently wrong	Surgeon trusts overlay during a wrong interval	Human factors evaluation and uncertainty-first outputs
<b>Shape Instantiation</b>	Robotics constraints	Collision avoidance requires usable 3D geometry	Incorrect surface model breaks distance-to-structure constraints	Instrument trajectory violates safe boundary	Require multimodal sensing and defined low-confidence behavior
<b>Decision Support</b>	Pre-op planning	Planning decisions depend on risk and advisability	Miscalibrated risk pushes corridor choice or aggressiveness	Overly aggressive plan in fragile patient	External and prospective validation; workflow-native interfaces
<b>Pre-op Planning Pipeline</b>	Intra-op navigation	Pre-op model is used as intra-op reference	Normalized pre-op outputs become wrong under deformation	Navigation uses outdated geometry	Closed-loop consistency across operating suite
<b>Intra-op Navigation</b>	Bounded autonomy	Perception is prerequisite for action	Perception errors produce unsafe actuation	Autonomous subtask continues under OOD conditions	OOD detection and swift transfer of control
<b>Any Module</b>	Downstream clinical workflow	Optional systems get dropped	Poor integration reduces adoption and increases misuse	Tool ignored or used incorrectly under time pressure	Interface and workflow are primary scientific outputs

**Table 3 | Translation-grade evaluation crosswalk to reporting and staged-evidence frameworks**

Evaluation Element	What to Report in Surgical Oncology AI Papers	Applies Most to	Guideline and Staged-Evidence Crosswalk
<b>Cohort Definition and Target Decision</b>	Define the decision the model supports, not just the outcome	Decision support, classification	TRIPOD+AI for prediction model reporting; DECIDE-AI for decision-support evaluation framing
<b>Missing Data and Imbalance Handling</b>	Explicit handling strategies and sensitivity analyses	Decision support	TRIPOD+AI and DECIDE-AI emphasize transparent data handling and evaluation design
<b>Internal vs External Validation Separation</b>	Distinguish internal validation from true external testing	Decision support, all modules	TRIPOD+AI; IDEAL encourages staged evidence as systems mature
<b>Prospective Evaluation in Real Workflow</b>	Prospective studies with workflow endpoints, time to intervention, concordance	Decision support, intra-op systems	DECIDE-AI specifically targets early-stage prospective evaluation; CONSORT-AI applies when randomized trials exist; IDEAL later stages
<b>Interface and Human Factors</b>	Describe interface, tumor board integration, AR human factors testing	Decision support, AR	DECIDE-AI highlights integration into practice; CONSORT-AI requires reporting AI-specific trial elements when trialed
<b>Uncertainty as First-Class Output</b>	Calibration, confidence, abstention rules, and defined fail behavior	All modules, especially AR and robotics	DECIDE-AI and TRIPOD+AI support transparent uncertainty and safety behavior reporting
<b>Failure-Mode Characterization</b>	Evaluate modes and speed of failure, not only average accuracy	Intra-op navigation, tracking	IDEAL style staged evaluation aligns with progressively harsher real-world testing
<b>Cascade and System-Level Evaluation</b>	Test integrity chain and upstream–downstream error propagation	Pre-op to intra-op to AR	IDEAL supports system maturation; DECIDE-AI supports system evaluation rather than isolated metrics
<b>Dataset Strategy for Domain Shift</b>	Build datasets with domain shift, label quality, temporal changes	All modules	TRIPOD+AI encourages transparent dataset and modeling description; staged evidence expects multi-site robustness
<b>Post-Deployment Monitoring</b>	Define monitoring signals, drift detection, auditability	Deployed systems	IDEAL long-term evaluation concepts; AI trial reporting literature emphasizes this as necessary for trustworthy adoption

**Table 4 | Risk-of-harm and failure-mode matrix mapped to mitigations and what is actually tested**

Failure Mode	Where It Manifests	Why It Happens	Why It Is Dangerous in Surgery	Mitigation Strategy You Propose	What the Literature Commonly Tests vs Rarely Tests
<b>Distribution Shift</b>	Decision support, classification	Multi-site variation in patients, imaging, practice	Quiet failure with plausible outputs	Routine external validation and prospective evaluation	Common: retrospective discrimination; Rare: prospective impact and multi-site robustness
<b>Confident Wrong Overlays</b>	Registration and AR	Subtle misalignment, non-rigid deformation, tracking drift	Elevates surgeon confidence toward wrong action	Uncertainty-first overlays; defined low-confidence behavior; multimodal sensing	Common: offline overlay accuracy; Rare: human factors harm testing under realistic conditions
<b>Cascading Upstream Error</b>	Detection to segmentation to registration	Pipeline dependency	Downstream harm scales from small upstream bias	Explicit cascade testing and integrity-chain evaluation	Common: module benchmarks; Rare: end-to-end safety tests
<b>Tracker Drift</b>	Intra-op tracking and AR	Occlusion, lighting, deformation, online updating	Intermittent wrong guidance at critical moments	Constraints during low confidence; governance for online updates	Common: controlled tracking benchmarks; Rare: governance and non-stationary behavior audits

**Table 4 | Risk-of-harm and failure-mode matrix mapped to mitigations and what is actually tested**

Failure Mode	Where It Manifests	Why It Happens	Why It Is Dangerous in Surgery	Mitigation Strategy You Propose	What the Literature Commonly Tests vs Rarely Tests
<b>OOD Edge Cases</b>	Robotics autonomy	Rare bleeds, unusual anatomy, unexpected complications	Policy performs worst at the edge where patients actually are	OOD detection and swift control transfer to human	Common: simulation success; Rare: validated OOD detection and safe handoff metrics
<b>Sim-to-Real Brittleness</b>	Reinforcement learning	Training in simulation with domain gap	Unsafe actuation when real-world dynamics differ	Conservative bounded autonomy; transfer learning validation	Common: simulation results; Rare: robust sim-to-real safety demonstration
<b>Command-Intention Misfire</b>	Human–robot interface	OR noise, occlusion, cognitive load	Single wrong command can be dangerous	Extremely low tolerance error audits; multimodal redundancy	Common: lab HRI demos; Rare: OR-grade validation under noise and stress
<b>Dataset Label Heterogeneity</b>	Segmentation, detection	Annotation cost, heterogeneity, weak labels	Geometry errors appear precise and persuasive	Weak or self-supervised scaling plus robust evaluation	Common: single-site curated labels; Rare: heterogeneity stress tests
<b>Security and System Breakdown</b>	Deployed systems	Cybersecurity gaps, device failures	Instantaneous harm in high-stakes environment	Secure, auditable systems and governance	Common: privacy discussion; Rare: engineering-grade failure accountability

Contemporary perioperative decision-making is anchored by traditional risk stratification tools (eg, comorbidity indices and general surgical risk calculators), yet these instruments can be imperfectly calibrated to specific oncologic procedures, patient populations, and evolving perioperative pathways.<sup>18–21</sup> ML-based CDS has therefore concentrated on three clinically actionable tasks: (1) preoperative prediction of postoperative morbidity, mortality, and resource use (eg, prolonged length of stay); (2) perioperative process-of-care decisions (eg, escalation of monitoring intensity or thromboprophylaxis); and (3) oncologic outcome prediction (eg, early recurrence, metastasis, survival) intended to tailor adjuvant therapy counseling and surveillance strategies.<sup>5,17,22–24</sup>

**Data Substrates and Modeling Paradigms**

Decision-support systems in surgical oncology are typically trained on multimodal clinical data rather than imaging alone.<sup>22,25,26</sup> Common substrates include structured EHR variables (demographics, comorbidities, laboratory values, medications), perioperative process data (procedure type, operative duration, transfusion, enhanced recovery pathway adherence), and longitudinal physiologic measurements.<sup>22,27,28</sup> Unstructured text has particular value because key prognostic and procedural descriptors are often documented in operative notes, pathology reports, and radiology narratives.<sup>25,26</sup> In ovarian cancer surgery, for example, natural language processing (NLP) applied to preoperative computed tomography (CT) reports improved prediction of morbidity and mortality beyond structured variables, illustrating how extraction of clinically salient detail from free text can enrich

perioperative risk models without requiring new data collection.<sup>26</sup>

Patient-generated health data are an emerging input modality for post-discharge decision support.<sup>29</sup> Remote telemonitoring platforms capturing symptoms and physiologic parameters have been evaluated for the prediction of postoperative complications after cancer surgery, addressing a recognized vulnerability in surgical pathways: many complications manifest after discharge, when traditional in-hospital monitoring has ended, and clinical contact is intermittent.<sup>29</sup> These studies suggest that integration of patient-reported and sensor-derived data may extend the temporal window of actionable risk detection, but they also raise implementation requirements (device adherence, data governance, and clinical response capacity) that must be incorporated into translational evaluation rather than treated as downstream operational details.

Lastly, the notion of regulatory readiness for perioperative AI now extends to include a lifecycle plan for controlled updates and performance management, not merely pre-market performance and accuracy. In the case of the United States, the Food and Drug Administration (FDA) has issued guidance for AI-enabled medical devices that supports iterative modification via the Predetermined Change Control Plan (PCCP), which addresses what changes can occur and how they can be assessed to provide reasonable assurance of safety and effectiveness. At the same time, FDA cybersecurity guidance also addresses security-by-design within the quality system and the integration of cybersecurity documents within the device submission package for those devices with cybersecurity risk. In the EU, the Medical Device Regulation (MDR 2017/745) has

also placed greater responsibility on manufacturers for post-market surveillance (Articles 83–86) and for software IT security within general safety and performance requirements. In practice, we would advocate the alignment of OR-related quality and cybersecurity processes (secure health software lifecycle processes IEC 81001-5-1) and the establishment of triggers for escalation and rollback based upon statistically significant performance degradation on previously determined clinical endpoints (including calibration), sustained increases in out-of-distribution (OOD) rates and tracking and registration failures beyond control limits, safety events plausibly linked to the AI model performance, or significant cybersecurity events; mitigation should include rollback to the last known good model performance, silent mode revalidation, and auditable change history consistent with the PCCP.<sup>15,29</sup>

Methodologically, most decision-support applications remain supervised prediction models trained on retrospective cohorts, benchmarking logistic regression or regularized regression against tree-based ensembles and, less commonly, neural networks.<sup>22,30–32</sup> Several comparative evaluations suggest that increased model complexity does not guarantee improved performance.<sup>22,30</sup> For example, in a population-based analysis of mortality prediction using administrative diagnosis codes, boosted trees did not outperform logistic regression, emphasizing that the added computational and interpretability costs of complex models should be justified by demonstrable gains in calibration, discrimination, or clinical utility relevant to the intended decision.<sup>14,22,33</sup>

#### Perioperative Risk Stratification and Workflow-Integrated Recommendations

A central translational premise of perioperative CDS is to identify individuals at elevated risk for adverse outcomes and to trigger targeted pathways that reduce preventable morbidity without disproportionate resource consumption.<sup>5,14</sup> In colorectal cancer surgery, Rosen et al. reported a registry-based ML model predicting 1-year mortality, deployed as a clinical decision support tool coupled to personalized perioperative care pathways; the implementation was associated with reductions in postoperative complications, improved “textbook” outcomes, shorter length of stay, and lower costs, illustrating an end-to-end translational pattern from prediction to workflow-embedded intervention and measurable system-level impact.<sup>5</sup> This kind of coupling between risk prediction and standardized, protocolized responses is also consistent with broader CDS literature showing that systems are more likely to change practice when they deliver recommendations (or pathways) within routine workflow rather than standalone risk estimates.<sup>14,15</sup>

Across oncologic procedures, ML-based models have been developed to predict composite postoperative morbidity as well as specific complications that motivate discrete prophylactic or monitoring interventions.<sup>21,28,31,33,34</sup> Examples include models for postoperative complications after liver surgery, early

postoperative complications after radical gastrectomy, and complication risk stratification after lung cancer surgery.<sup>21,31,33</sup> In a thoracic oncology cohort, an elastic net model predicting postoperative complications after lung cancer surgery achieved good discrimination and improved risk prediction compared with the Charlson Comorbidity Index, demonstrating that procedure-specific perioperative models can outperform general comorbidity summaries when appropriately developed and temporally validated.<sup>19,21</sup> Similarly, in head and neck cancer surgery, ML models have been used to predict prolonged length of stay and were compared against the ACS-NSQIP risk calculator and conventional statistical models, highlighting a recurring translational theme: models must be evaluated not only in isolation but also against the clinical baselines that actually shape care decisions.<sup>18,20</sup>

Risk prediction for procedure-specific events illustrates both the opportunity and the translational constraints of CDS. Anastomotic complications after esophagectomy (leak, stricture) are high-impact outcomes with potential for algorithmically guided monitoring and early intervention.<sup>23,35,36</sup> Recent studies have developed models combining clinical variables with imaging-derived features to predict anastomotic leakage, reporting discrimination and calibration metrics, as well as threshold-based evaluation intended to support clinical decision-making.<sup>23,35</sup> In parallel, automated ML approaches have been applied to predict post-esophagectomy strictures, suggesting that algorithmic pipelines can be designed for scalability when feature extraction and model selection are automated, although automation does not obviate the need for external validation, drift monitoring, and transparent reporting.<sup>36,37</sup>

Dynamic decision support aligned to perioperative trajectories is increasingly emphasized because physiologic and laboratory changes often precede clinically recognized deterioration.<sup>27,28</sup> In a multicenter gastrectomy cohort, time-sequential ML models using serial laboratory and vital-sign measurements improved early prediction of postoperative complications compared with baseline models, supporting the concept that CDS should be designed as a longitudinal process rather than a single preoperative score.<sup>28</sup> Complementary evidence comes from dynamic modeling of postoperative venous thromboembolism risk after colorectal cancer surgery, where repeated updates to risk estimates reflected evolving clinical states and supported individualized prophylaxis and monitoring decisions in a multicenter context.<sup>27</sup>

Prospective validation is particularly important for perioperative CDS because apparent performance in retrospective data can be inflated by outcome misclassification, missing data mechanisms, and site-specific practice patterns.<sup>13,37,38</sup> Notably, prospective evaluation has begun to appear in some surgical oncology prediction studies.<sup>28,34</sup> In lung cancer surgery, Chen et al. developed and prospectively validated an explainable ML model for postoperative pulmonary complications, reporting discrimination, calibration, decision

curve analysis, and feature attribution via SHapley Additive exPlanations (SHAP), an unusually comprehensive suite of translational evaluation components for a single procedure-focused model.<sup>34</sup> Together, these studies illustrate a maturation of perioperative CDS from static, internally validated scores toward longitudinal, interpretable, and prospectively assessed systems aligned with intervention thresholds and workflow integration.<sup>27,34,37</sup>

### Oncologic Outcomes and Treatment-Tailoring Decision Support

Decision support in surgical oncology must also address endpoints that determine cancer control and long-term survivorship, particularly early recurrence and metastasis after ostensibly curative resection.<sup>17,39</sup> In esophageal cancer, Rahman et al. developed ML models to predict early recurrence after surgery using multinational data and applied internal–external validation across centers, providing an instructive template for transportability assessment when multicenter external validation is not yet feasible.<sup>39</sup> In pancreatic cancer surgery, an interpretable ML model was developed to predict early liver metastasis after resection and was externally validated with calibration and decision-curve analyses; notably, the model was also implemented as an accessible application, reflecting an explicit translation intent. These examples underscore that oncologic CDS often targets events (recurrence, metastasis) that are clinically meaningful precisely because they can change downstream actions such as adjuvant therapy selection, surveillance intensity, and enrollment into clinical trials.<sup>15,39</sup>

In glioblastoma, for instance, a quintessential conundrum in oncology, the same reasoning applies, in that despite ‘maximal safe’ resection, early failure on imaging and failure patterns at the edge of resection are frequent and directly inform intensification of treatment, trial participation, and follow-up intensity.<sup>40,41</sup> A CDS approach using preoperative magnetic resonance imaging (MRI) radiomics, intraoperative factors such as extent of resection, and molecular factors such as MGMT promoter methylation or IDH mutation could risk-stratify for early failure and inform postoperative treatment and follow-up.<sup>42</sup> Notably, in glioblastoma, this CDS system would need to directly address issues related to uncertainties in the differentiation of true failure from treatment-associated radiation changes and pseudoprogression.

Certain CDS applications bridge the boundary between decision support and intraoperative perception by producing staging or resectability assessments that directly influence operative strategy.<sup>24,32</sup> The artificial intelligence laparoscopic exploration system (AiLES) is an AI system designed to recognize intra-abdominal metastasis during laparoscopic gastric cancer surgery using real-time video, with evaluation against surgeon assessments; this form of intraoperative decision support can alter the surgical plan and downstream therapy selection when occult metastatic disease is detected.<sup>24</sup> In parallel, perioperative personalized decision

support reviews have emphasized the need to integrate educational and feedback components with prediction outputs so that CDS improves not only prediction but also clinical action and decision quality.<sup>15,32</sup>

### Evidence Maturity, External Validation, and Clinical Utility

A consistent finding across systematic reviews is that much of the surgical oncology AI literature remains concentrated in retrospective development studies, with heterogeneous endpoints and limited prospective evaluation demonstrating improved patient outcomes or workflow efficiency.<sup>13,37</sup> Bektaş et al., in a systematic review of ML applications in upper gastrointestinal cancer surgery, noted substantial heterogeneity in outcomes and methodological approaches, limiting comparative synthesis and emphasizing the need for standardized reporting and validation.<sup>37</sup> Similarly, in rectal cancer surgery, a systematic review of AI models predicting surgical difficulty highlighted variability in definitions, imaging features, and validation strategies, illustrating that even seemingly bounded decision support targets (eg, “difficulty”) can suffer from endpoint ambiguity that impairs translation.<sup>37</sup>

Generalizability across institutions and time is a core translational constraint in perioperative CDS because case mix, perioperative protocols, coding practices, and follow-up capture differ across sites and evolve over time.<sup>13</sup> Several surgical oncology studies have begun to address this through explicit external validation designs.<sup>6,27,29,39</sup> Dal Cero et al. performed an international external validation of a machine learning model for 90-day mortality after gastrectomy, providing direct evidence that transportability must be tested under clinical and geographic shifts rather than assumed from internal performance alone.<sup>6</sup> Complementary internal–external validation strategies, as used in multinational recurrence prediction after esophageal cancer surgery, similarly represent pragmatic approaches for early transportability assessment when full external validation cohorts are not yet available.<sup>37,39</sup>

From a clinical translation perspective, discrimination is necessary but insufficient.<sup>43,44</sup> Poorly calibrated models can misestimate absolute risk and therefore misdirect threshold-based decisions (eg, intensive monitoring, admission to higher-acuity units, or initiation of prophylaxis).<sup>18,43</sup> Consequently, calibration assessment and decision-analytic evaluation should be considered minimum requirements for CDS systems intended to change clinical actions.<sup>43,44</sup> This emphasis is reflected both in methodological guidance and in several recent surgical oncology studies reporting calibration curves, Brier scores, and decision curve analysis (DCA) alongside discrimination.<sup>23,34,35,40</sup> The reporting of DCA is particularly relevant to decision support because it quantifies net benefit across thresholds, connecting model outputs to the clinical consequences of false positives and false negatives rather than relying solely on rank-based discrimination.<sup>34,44</sup>

### Human Factors, Equity, and Reporting Standards

Even when predictive performance is favorable, CDS adoption depends on usability, clinician trust, and integration at the time and location of decision-making.<sup>14,45</sup> A systematic review of clinical decision support systems identified features associated with improved clinical practice, including automatic provision within workflow and delivery of recommendations rather than assessments alone, design principles that are directly relevant to AI/ML tools intended for perioperative oncology care.<sup>14</sup> Qualitative research on decision support interventions targeting shared decision-making similarly underscores clinician concerns about appropriateness of recommendations, time costs, and alignment with clinical judgment, anticipating barriers to AI/ML CDS if outputs are poorly contextualized or not coupled to feasible actions.<sup>45</sup> For patient-centered surgical decisions, evidence from randomized evaluation of conversation aids in breast cancer surgery further illustrates that the quality of decisions depends on aligning clinical options with patient values and understanding; prediction outputs are therefore only one component of the translational goal of improving decision quality.<sup>15,46</sup>

Equity considerations are central to translational readiness because prediction models trained on historical healthcare data may encode structural inequities in access, treatment, and documentation.<sup>13,47</sup> In a widely cited example outside oncology, Obermeyer et al. demonstrated substantial racial bias in a commercial risk prediction algorithm arising from the use of healthcare costs as a proxy for health needs; analogous proxy-label and documentation biases are plausible in surgical oncology when outcomes, comorbidities, and follow-up are differentially captured across groups.<sup>25,37,47</sup> Systematic review evidence in cancer pathways also indicates inconsistent reporting of subgroup performance and limited attention to equity in prospective evaluations, reinforcing the need to treat fairness assessment as a core translational requirement rather than an optional post hoc analysis.<sup>13,38,42</sup>

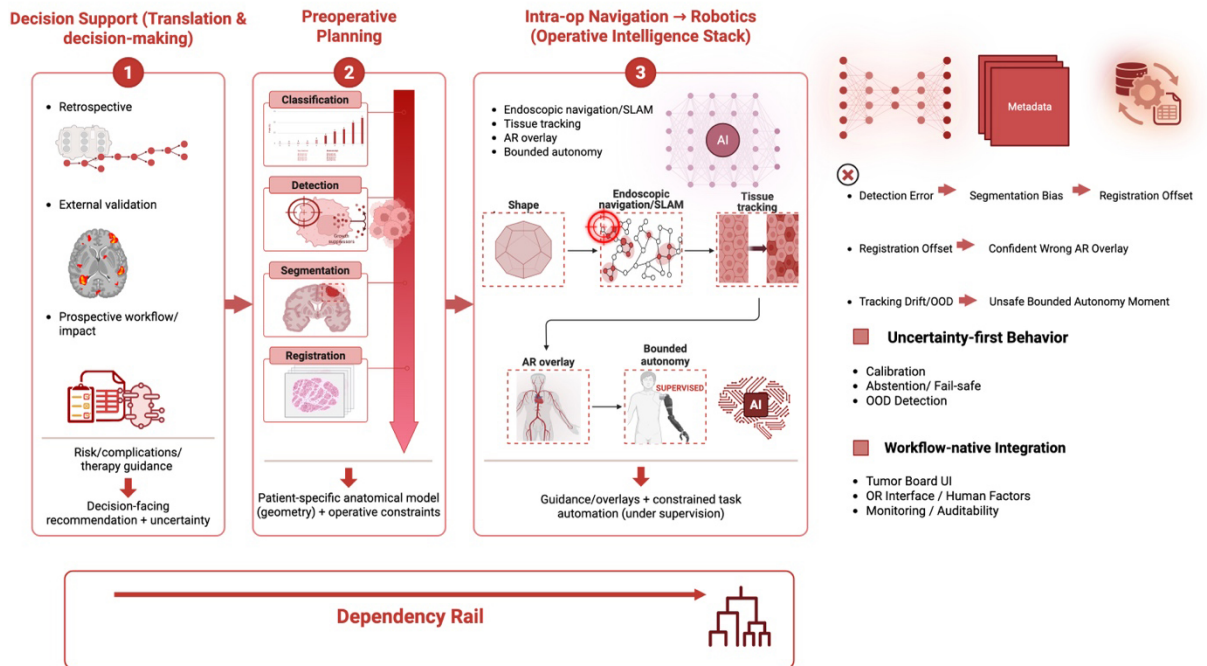
Finally, reporting and evaluation standards are prerequisites for reliable translation.<sup>42,48–50</sup> TRIPOD and the updated TRIPOD+AI statement provide minimum reporting items for prediction model studies using regression or ML methods, including a transparent description of data sources, outcome definitions, handling of missingness, and full reporting of performance metrics.<sup>48,49</sup> PROBAST and the updated PROBAST+AI tool support structured assessment of risk of bias and applicability, offering a shared framework for judging whether a model is likely to generalize and whether reported performance is credible.<sup>42,50</sup> For prospective and interventional evaluation of AI-enabled decision support, CONSORT-AI and SPIRIT-AI extend trial reporting and protocol guidance, while DECIDE-AI provides a reporting guideline for early-stage clinical evaluation of decision support systems driven by AI, an appropriate framework for perioperative CDS that may be introduced first in limited workflows before broader rollout.<sup>13,15,51</sup>

Human factors readiness of perioperative AI systems for use will demand interface design elements that support safety in situations of uncertainty, drift, and incorrectness. For example, at the UI level, uncertainty-first design principles must be implemented through quantification of probabilities and explicit threshold bands with prespecified actions, insufficient evidence states when probabilities are too low, drifting detection and notification, and hard constraints against providing confident guidance through overlays and trajectories when upstream uncertainty bands are breached. At the workflow level, each AI system will need a compact failure mode playbook that maps common types of failures (eg, detection miss, segmentation bias, registration offset, tracking drift/OOD, cybersecurity issues) to standard responses such as confirmation steps, escalation protocols, fail-safe transitions to manual or alternative modes of imaging, documentation requirements, and review triggers after failure events. In parallel, equity assessment will need to be addressed as a translational deliverable through a simple audit template that reports stratified performance and calibration across clinically meaningful subgroups and settings (eg, by sex, age, race/ethnicity if available, comorbidity burden, language and insurance status if applicable), subgroup-specific net benefit analysis through DCA, and prespecified mitigation approaches such as targeted data enrichment and reweighting and domain adaptation, subgroup-dependent thresholds with clinically justified trade-offs and monitoring with rollback criteria if disparate errors are seen (Figure 1).

### Regulatory, Safety, and Lifecycle Governance for Perioperative Oncology AI

When ML systems are used to provide patient-specific information for the guidance of perioperative cancer care, they are likely to be categorized as Software as a Medical Device (SaMD), where the intention is to diagnose, treat, mitigate, or in any way affect or influence the care provided to the patient. For such ML systems to be translated for practical use in the field of medicine, they must align with the lifecycle approach that covers the major regulatory regions or the regulatory harmonization process (FDA, EMA, IMDRF), as opposed to the retrospective approach that is based purely on the system's historical performance. For the ML system to be translated for practical use in the field of medicine, the system must be designed to meet the regulatory requirements through the intended use, the definition of the system's role in the clinical setting (whether assistive or advisory), the hazards associated with the system's use, as well as the associated mitigations in line with the traditional approach to risk management. This means that the system must meet the requirements for the software development lifecycle (IEC 62304), usability engineering for the system's interfaces (IEC 62366), as well as the systematic approach to risk management (ISO 14971) that applies to the field of surgical oncology, where time pressure, irreversible actions, and the consequences associated with errors are the norm. The system's governance must not

An Uncertainty-First Operative Intelligence Stack for AI-Enabled Surgical Care



**Fig 1 | An uncertainty-first operative intelligence stack for AI-enabled surgical care** This diagram represents the end-to-end, dependency-conscious approach to the systematic integration of AI across the surgical continuum of care, from decision support to planning and navigation with bounded autonomy. (1) Decision support: Intra-operatively and retrospectively validated models are used to make inferences regarding risk assessment, complication, and therapy, which are quantitatively paired with model-fit or solution-space uncertainties to facilitate decision-making for clinicians. (2) Preoperatively, patient-specific models of anatomy and geometry are created via the use of pipelines of classification, detection, segmentation, and registration, which incorporate operative conditions to provide the patient with a surgical model. (3) Intraoperatively, navigation and robotics provide the operative intelligence stack that incorporates techniques for endoscopic navigation and SLAM, tissue tracking, and augmented reality overlays, which utilize supervised, bounded autonomy. Note the failure modes depicted in the figure: detection failure, segmentation bias, registration offset, and tracking drift with out-of-distribution (OOD) inputs, which propagate to confident but erroneous guidance. As the dependency rail makes manifest, the reliability of the former models of intelligence conditionally bounds the latter model's autonomy to emphasize the imperative for safe AI applications. For each domain, translational readiness is constrained by: data provenance and representativeness; endpoint definition and clinical action linkage; internal, external, and prospective validation; robustness to dataset shift; calibration and uncertainty communication; human factors and workflow integration; and lifecycle governance (safety, monitoring, and cybersecurity). The framework emphasizes that headline accuracy is insufficient when downstream actions are irreversible, time-constrained, and safety-critical. An accompanying domain-level evidence map should make the action linkage explicit (what decision changes, by whom, and at what threshold) and align evaluation to that decision rather than headline accuracy. Minimum reporting should include external and prospective validation (including silent trials) plus calibration and clinical-utility analyses (reliability curves and decision-curve analysis) to justify operating points and fail-safe triggers.

only be limited to the pre-deployment validation but must extend to the post-deployment monitoring to ensure that the system is performing as required in the real-world setting.<sup>52-56</sup>

Collectively, the evidence to date indicates that clinical translation of ML decision support in surgical oncology will depend less on incremental improvements in discrimination and more on rigorous alignment of model endpoints with actionable decisions, transparent reporting, robust external and prospective validation, user-centered workflow integration, and ongoing monitoring for calibration drift and inequitable performance.<sup>13,15,49,50</sup>

**Preoperative Planning**

Preoperative planning in surgical oncology refers to computational and procedural methods that transform pre-treatment data, most commonly cross-sectional imaging, into patient-specific representations intended to guide operative strategy, define resection targets, anticipate technical constraints, and coordinate adjunct technologies (eg, navigation, fluorescence, or robotic assistance).<sup>1,2,51,52</sup> As deployed in current clinical workflows, planning systems typically emphasize geometric and spatial reasoning, what anatomy is present, where the tumor is related to critical structures,

and how operative goals map onto feasible dissection planes, rather than solely providing probabilistic risk estimates.<sup>2,57</sup> Across organ systems, the dominant technical primitives underlying these workflows remain lesion characterization (including staging-relevant inference), anatomic segmentation and reconstruction, and multimodal fusion/registration.<sup>1,2,53-56,58</sup>

### Imaging-Derived Characterization and Staging

Radiologic staging and anatomic definition remain central determinants of operability and approach selection; accordingly, AI/ML methods have been applied to extract structured staging signals from imaging, radiology narratives, and complementary tissue sources.<sup>1,59-62</sup> Radiomics-based strategies, whether relying on engineered feature families or learned representations, have been positioned as a means to quantify tumor phenotype and context in ways that may complement conventional human interpretation, while also amplifying sensitivity to imaging protocol variability and preprocessing choices.<sup>1</sup> In breast oncology, convolutional neural network analysis of multiparametric MRI has been evaluated for preoperative prediction of axillary lymph node metastasis, directly targeting a staging component that may influence the extent of axillary surgery and systemic therapy sequencing.<sup>59</sup> In gastric cancer, multimodal approaches that combine CT-derived information with digital pathology features (whole-slide imaging [WSI]) have been reported for staging-related prediction, illustrating how “preoperative planning” increasingly extends beyond imaging alone when pre-treatment tissue is available (eg, endoscopic biopsy).<sup>60</sup>

Natural language processing (NLP) has also been explored to convert narrative radiology into structured, decision-relevant staging information. Automated esophageal cancer staging from free-text radiology reports has been evaluated using large language models (LLMs), supporting the feasibility of extracting standardized staging descriptors from routine documentation and potentially reducing manual abstraction burden in multidisciplinary workflows.<sup>61</sup> Because many planning decisions hinge on a synthesis of imaging, histology, and clinical narrative, LLM-centered approaches may be attractive for workflow integration; however, their translational value depends on rigorous evaluation under realistic variability in report style, missingness, and institutional terminology, as well as safeguards against unsupported inference, and values that integrate restorative ventures for at risk social groups in healthcare.<sup>61,63</sup>

### Planning-Oriented Targets Beyond Categorical Stage

Planning outputs extend beyond the categorical stage to intermediate, decision-relevant targets such as anticipated technical difficulty, required exposure, and margin risk.<sup>41,62</sup> In rectal cancer surgery, AI models applied to preoperative MRI have been studied for predicting surgical difficulty, an endpoint that can influence the choice of minimally invasive vs open approach, need for specialized instrumentation, and

staffing/experience requirements.<sup>41</sup> In oral squamous cell carcinoma, pathomics-based models have been studied for predicting positive surgical margins, illustrating an emerging paradigm in which pre-treatment or perioperative tissue-derived features are translated into margin-focused planning inputs.<sup>62</sup> These examples underscore a broader trend: planning systems often seek to forecast “actionable constraints” (eg, limited working space, high-risk dissection planes) rather than only long-term outcomes, but such targets require careful operationalization and validation because they can be sensitive to surgeon factors, institutional practice patterns, and differences in operative technique.<sup>1,41</sup>

### Segmentation, Quantitative Anatomy, and Patient-Specific Modeling

Most preoperative planning workflows depend on delineation of tumors, organs-at-risk, and critical structures, making anatomic segmentation an enabling task for 3D reconstruction, simulation, and downstream guidance.<sup>2,54-56,58,64-67</sup> In thoracic oncology, semantic segmentation of chest CT has been used to recognize patient-specific pulmonary vessel variants relevant to resection planning, illustrating how segmentation can be directed toward clinically consequential anatomic variability rather than organ boundary extraction alone.<sup>58</sup> In rectal cancer, 3D reconstruction has been described as a means to improve diagnosis and surgical planning, reflecting sustained interest in translating pelvic MRI into spatially coherent representations that can support operative strategy and multidisciplinary discussion.<sup>55</sup>

Segmentation also supports simulation-oriented planning that seeks to anticipate intraoperative views and dissection trajectories.<sup>64-66</sup> An AI-driven 3D simulation system for gastric cancer surgery has been reported as a retrospective validation study, consistent with an early translational pathway in which automated model construction and anatomy recognition are first evaluated for feasibility and concordance with expert interpretation before the impact on procedural decisions is tested prospectively.<sup>64</sup> Similarly, AI-based technology to generate a 3D model for rectal cancer surgery planning from MRI has been reported as a step toward preoperative simulation, highlighting an approach in which imaging-derived models are used for rehearsal and spatial understanding rather than direct automation of operative steps.<sup>65</sup> At a more granular anatomic scale, automated segmentation of male pelvic floor soft tissues have been proposed for preoperative simulation and morphologic assessment in lower rectal cancer surgery, reflecting the planning value of extracting patient-specific pelvic anatomy that may influence exposure and dissection strategy.<sup>66</sup>

Planning applications have also been described in hepatobiliary contexts using intelligent image segmentation approaches, aligning with long-standing clinical needs for patient-specific visualization of vascular and biliary anatomy and for volumetric assessment in resection planning.<sup>67</sup> Although organ-specific reviews emphasize potential benefits of AI/ML for liver can-

cer surgery, the evidentiary standard for preoperative planning systems remains high because errors in anatomic delineation or spatial relationships can directly affect operative decisions.<sup>52</sup>

#### Quantitative Biomarkers Derived from Segmentation

Quantitative biomarkers derived from segmentation can inform preoperative optimization and approach selection when measurement performance is adequate, and the clinical mapping from measurement to action is explicit.<sup>68-70</sup> Automated body composition assessment from routine CT has been evaluated as a scalable method to measure muscle and adipose compartments, including validation-focused studies of deep learning derived measurements from standard-of-care CT examinations.<sup>68,69</sup> Related work has linked imaging-derived sarcopenic obesity with prognostic outcomes, supporting the premise that quantitative imaging may identify phenotypes relevant to prehabilitation, nutritional optimization, and candidacy assessments.<sup>70</sup> These approaches may be particularly attractive because they leverage imaging already obtained for staging, but translational use requires transparent reporting of measurement error, robustness across scanners and protocols, and clear articulation of how biomarker thresholds would change management.<sup>1,68,70</sup>

#### Image Fusion, Registration, and Deformation-Aware Planning

Preoperative planning frequently requires integration across modalities (eg, multiphase CT and MRI), across timepoints, and across coordinate frames that differ from the intraoperative configuration; therefore, image registration and fusion are recurrent bottlenecks for constructing coherent 3D models that can be inspected by clinicians and, in some settings, transferred to guidance systems.<sup>2,53,54,56</sup> Reviews of 3D reconstruction and organ modeling emphasize that static preoperative models can be undermined by soft-tissue deformation and physiologic motion (eg, respiration), motivating deformation-aware planning and, where feasible, updating or reconciling models with intraoperative sensing.<sup>2,54</sup> The urologic robotics literature similarly situates preoperative modeling within a broader imaging-robotics ecosystem, highlighting both the promise of patient-specific reconstructions and the technical challenges of aligning preoperative images with intraoperative anatomy.<sup>53</sup>

#### Simulation, Rehearsal, and Operator-Facing Visualization

Because a preoperative plan must be interpretable and actionable for surgeons and operative teams, many planning systems emphasize operator-facing 3D visualization and rehearsal rather than prediction outputs alone.<sup>2,54-56,58,64,65</sup> Work on 3D lung model development for minimally invasive lung cancer surgery has explicitly framed a progression from static reconstructions

toward real-time or dynamic modeling, consistent with the clinical observation that thoracoscopic and robotic approaches increase dependence on accurate spatial representations when tactile feedback is limited.<sup>54</sup> This emphasis on visualization is also apparent in gastrointestinal oncology planning efforts that use AI-enabled reconstruction to support preoperative simulation in gastric and rectal cancer operations.<sup>64,65</sup>

Virtual reality and 3D-printed models have been systematically reviewed as planning tools in image-guided, robot-assisted nephron-sparing surgery, providing evidence synthesis on how patient-specific 3D representations may support surgeon understanding, rehearsal, and intra-team communication.<sup>57</sup> Complementary discussions of robotic partial nephrectomy in the era of 3D virtual reconstructions similarly reflect continued interest in operationalizing preoperative 3D models as part of routine planning.<sup>71</sup> While these technologies are conceptually aligned with surgical oncology goals, improving spatial understanding and reducing unanticipated anatomy, the translational question is whether such tools measurably improve decision quality or outcomes when implemented at scale, accounting for the time and expertise required for model generation and review.<sup>57,71</sup>

#### Selecting and Planning Adjunct Imaging and Guidance Technologies

Preoperative planning also includes selection and orchestration of adjunct intraoperative sensing modalities and tracers, with plans often specifying timing of administration, expected contrast behavior, and decision thresholds.<sup>56,72,73</sup> Reviews of precision cancer surgery describe fluorescence-guided approaches that seek to improve tumor margin identification and lymphatic mapping, while emphasizing constraints imposed by probe pharmacokinetics, tissue optical properties, and imaging hardware.<sup>72</sup> A randomized translational study of protein- and peptide-based probes illustrates that probe selection and deployment strategy are empirically tractable components of the planning process, but also highlights that clinical utility depends on matching probe behavior to a defined decision task (eg, margin assessment or lesion localization).<sup>73</sup> Robotic- and image-guided workflows have likewise been described for fluorescence detection, reflecting ongoing efforts to integrate targeted imaging into operative systems.<sup>74</sup>

Lymphatic mapping provides a complementary example in which tracer workflow and imaging protocol determine intraoperative localization targets. Indocyanine green (ICG)-based lymph node mapping has been described in cancer surgery contexts, and preoperative sentinel node mapping work in gynecologic oncology demonstrates how imaging protocol design and tracer deployment can be codified into planning pathways that influence operative navigation targets.<sup>75,76</sup> Optical coherence tomography (OCT) adds a further dimension by providing microstructural imaging that can, in principle, support margin-focused

decision-making; systematic and scoping reviews describe OCT in intraoperative tissue assessment and AI-enabled OCT for disease detection and diagnosis, respectively, emphasizing both potential and the need for clinically grounded validation.<sup>77-79</sup>

#### Methodological and Evidentiary Requirements for Translation

For AI/ML-enabled preoperative planning, translational readiness depends on demonstrating that outputs are reliable under routine imaging heterogeneity, that errors are bounded and detectable, and that any claimed benefit is linked to decision-relevant endpoints rather than surrogate technical metrics alone.<sup>1,2,54,56,68</sup> Validation studies in segmentation and quantitative imaging (eg, body composition measurement) illustrate the importance of reporting agreement with reference standards and clarifying conditions under which manual correction remains necessary.<sup>68,69</sup> Because planning models often function as upstream dependencies for downstream navigation, simulation, or robotic workflows, failure modes should be characterized not only at the level of segmentation overlap but also in terms of clinically meaningful spatial errors (eg, mislocalization relative to vessels, planes, or margins).<sup>2,54,56,58</sup>

Surgeon-facing generative systems and automated staging tools introduce additional requirements beyond conventional imaging models, including provenance-aware synthesis, safeguards against hallucinated or non-evidence-based assertions, and evaluation under realistic documentation variability.<sup>61</sup> Current single-center retrospective planning studies, whether focused on 3D reconstruction, simulation, or prompt-driven assistance, should therefore be interpreted as feasibility signals; subsequent work must incorporate external validation, prospectively defined endpoints, and, where possible, decision-impact study designs that quantify how planning outputs change operative choices and whether those changes improve patient-centered outcomes.<sup>1,54,57,61-65</sup>

#### Intraoperative Navigation with Robotic Assistance/Control

Intraoperative navigation with robotic assistance/control comprises computational systems that interpret operative data streams in real time and translate those interpretations into actionable guidance, visualization, or robot-mediated behaviors during surgery.<sup>3,53,80,81</sup> In surgical oncology, this domain has expanded alongside the adoption of minimally invasive and robotic approaches across urologic, colorectal, foregut, and pancreatic procedures, which standardize endoscopic viewpoints and yield high-fidelity digital signals (eg, video and robotic platform telemetry).<sup>3,53,80,81</sup> Contemporary frameworks described in systematic and specialty-focused reviews generally position intraoperative AI/ML as progressing from perception and measurement to navigation and decision support, and only later (and more cautiously) toward shared autonomy or control-constraining behaviors.<sup>3,53,80,81</sup>

#### Data Modalities and the Primacy of Intraoperative Perception

Across robotic oncologic surgery, the dominant substrate for intraoperative ML is surgical video, reflecting the ubiquity and standardization of endoscopic and robotic camera feeds and the relative ease of rendering outputs as overlays within existing display pipelines.<sup>3,53,80</sup> Proposed enabling architectures for AI-enhanced processing of da Vinci robot-assisted video emphasize that translational feasibility is conditioned by system-level constraints, deterministic latency, robust video ingestion, and reliable handling of large data streams, which can be as decisive as model accuracy for operating room deployment.<sup>82</sup>

Video-based perception has been used to recognize anatomic structures directly tied to oncologic safety and postoperative function. Deep learning approaches for vessel and tissue recognition during third-space endoscopy illustrate a generalizable pattern: models are trained to detect structures that may be visually subtle yet clinically consequential when injured, with outputs intended to support safer dissection in minimally invasive contexts.<sup>83</sup> In urologic oncology, an AI-based intraoperative nerve recognition system has been described for nerve sparing during robotic prostatectomy, exemplifying a navigation-adjacent use case in which “recognize-and-overlay” can provide moment-to-moment spatial cues during technically demanding dissection.<sup>84</sup> These approaches align with broader reviews of robotic oncologic surgery that highlight computer vision-based recognition of anatomy and “danger zones” as a near-term, clinically plausible tier of intraoperative AI support.<sup>3,53,80,81</sup>

Intraoperative perception has also been extended to targets that are explicitly oncologic rather than purely anatomic. In laparoscopic gastric cancer surgery, the AiLES was developed for real-time recognition of intra-abdominal metastasis during exploration, directly addressing a clinically consequential failure mode in which small or occult lesions are overlooked, and staging is thereby misclassified.<sup>24</sup> Reviews of intraoperative imaging tools in gastrointestinal oncology likewise emphasize that the clinical value of intraoperative perception increases when outputs are tied to specific intraoperative actions (eg, escalation to biopsy/frozen section, targeted inspection, or altered resection strategy) rather than solely to descriptive visualization.<sup>56</sup>

Functional state estimation provides a complementary intraoperative objective because tissue perfusion and viability influence anastomotic strategy and can interact with oncologic extent through resection planning.<sup>56,85,86</sup> AI-based real-time microcirculation imaging has been evaluated for assessing colonic perfusion status, representing a class of intraoperative systems that infer physiologic states from subtle imaging cues that are difficult to quantify reliably by human vision alone.<sup>85</sup> In addition, intraoperative near-infrared (NIR) functional imaging has been studied in colorectal cancer settings, reflecting efforts to quantify perfusion-related and function-related signals that

could be integrated into navigation workflows and assessed against endpoints such as leak risk, complication rates, or re-intervention.<sup>86</sup>

#### **Geometry-Aware Navigation: Localization, Registration, and Coordinate Consistency**

Navigation systems must establish spatial relationships between recognized structures, instruments, and (when applicable) preoperative models, often under conditions of soft-tissue deformation, camera motion, and variable insufflation.<sup>3,53,80</sup> Real-time vascular anatomical image navigation has been demonstrated for laparoscopic surgery as an approach to aligning vascular anatomy with intraoperative views to support orientation and safer dissection around critical vessels.<sup>8</sup> In rectosigmoid oncology, the feasibility of optical stereotactic navigation supported by deep learning based 3D modeling has been reported, highlighting the clinical interest in “coordinate-consistent” guidance for tumor localization and margin-conscious dissection within anatomically constrained pelvic spaces.<sup>9</sup>

Evidence from thoracic oncology similarly indicates active development of geometry-aware navigation strategies. A prospective application study of non-contact, AI-assisted intraoperative 3D navigation in lung cancer surgery demonstrates a pragmatic approach in which spatial information is acquired and reconstructed intraoperatively to mitigate the mismatch between preoperative imaging and intraoperative anatomy.<sup>12</sup> Across these demonstrations, the translational challenge is not merely producing a geometric solution but maintaining accuracy under clinically routine perturbations (eg, deformation, retractors, motion, variable insufflation), and predefining failure detection and escalation workflows when navigation assumptions no longer hold.<sup>3,9,12,53</sup>

Visualization is a central component of intraoperative navigation because guidance must be interpretable at the point of action. Reviews of extended reality and “metaverse” concepts in surgery discuss how augmented visualization environments could provide spatial context and training ecosystems, while emphasizing limited clinical evidence for outcome benefit and the nontrivial challenges of integrating head-mounted or mixed reality interfaces into sterile workflows.<sup>87</sup> In robotic radical prostatectomy, specialty-focused reviews similarly frame advanced visualization (including potential augmented guidance) as an emerging direction that will require rigorous clinical validation and human factors evaluation before routine adoption.<sup>81</sup>

#### **Imaging-Augmented Guidance in Robotic Oncology**

Optical contrast agents and intraoperative imaging can function as “signal amplifiers” for tumor, lymphatics, or function, and ML methods are increasingly used to standardize interpretation and generate overlays that can be integrated into robotic visualization pipelines.<sup>53,56</sup> Robotic assistance can stabilize imaging geometry and enable instrument-mounted sensing, as illustrated by robotic “click-on” fluorescence detection

using surgical instruments to characterize molecular tissue aspects, a design pattern that can support localized assessment within an otherwise video-dominated workflow.<sup>74</sup>

Deep learning enabled fluorescence imaging has been developed for surgical guidance tasks aligned with oncologic decision-making. For oral cancer, in silico trained deep learning methods have been described for fluorescence-based depth quantification and for fluorescence-guided margin classification in preclinical models, representing an approach that seeks to convert fluorescence signal into explicitly interpretable, decision-relevant outputs rather than qualitative visualization alone.<sup>88,89</sup> More broadly, reviews of precision cancer surgery emphasize that the translational value of imaging-augmented guidance depends on linking the imaging signal (and any ML-derived interpretation) to prespecified intraoperative decisions (eg, additional resection, targeted sampling, or altered dissection plane) and demonstrating benefit against clinically meaningful endpoints.<sup>56</sup>

Near-infrared fluorescence and functional imaging also support navigation decisions beyond margin assessment, including lymphatic mapping and perfusion-oriented guidance.<sup>56,86,90</sup> In upper gastrointestinal oncology, indocyanine green (ICG) lymph node mapping has been described for cancer surgery, and prospective mapping of lymphatic drainage patterns with NIR fluorescence has been studied during robotic-assisted minimally invasive Ivor Lewis esophagectomy.<sup>75,91</sup> These studies illustrate that intraoperative guidance is often probabilistic and pathway-dependent (eg, drainage patterns and node basins) and therefore benefits from clear protocols describing how intraoperative findings are intended to modify the operative plan.<sup>56</sup>

#### **Label-Free Tissue Assessment and Uncertainty-Aware Inference**

Label-free tissue characterization provides an alternative approach to intraoperative guidance by aiming to classify tissue state directly from biophysical signatures, potentially enabling margin assessment without exogenous tracers.<sup>11,92,93</sup> Perioperative tissue assessment using combined mass spectrometry and histopathology imaging has been described as a framework for integrating molecular and morphologic information to inform intraoperative decisions in cancer surgery.<sup>93</sup> Within this domain, uncertainty estimation for surgical margin detection using mass spectrometry has been reported as a strategy to calibrate confidence and support decision thresholds suited to high-stakes intraoperative use.<sup>11</sup>

Representation-learning strategies have also been applied to reduce dependence on exhaustive labeling while preserving clinically relevant performance, exemplified by image-driven self-supervised learning for mass spectrometry-based tissue assessment.<sup>92</sup> Although these methodological advances may improve scalability, systematic syntheses in robotic cancer surgery converge on the need for external validation and

explicit specification of action policies when model outputs are uncertain, particularly when outputs are used to trigger irreversible actions such as widening margins, converting operative plans, or escalating to additional resection.<sup>3,53</sup> Accordingly, translational protocols should predefine how uncertainty will be communicated to the surgical team, how clinicians should respond to uncertain outputs (eg, confirmatory pathology), and how performance will be monitored over time once deployed.<sup>3,11,53,92</sup>

#### **Microscopic Intraoperative Imaging and AI-Enabled Interpretation**

Microscopic intraoperative imaging modalities such as OCT have been evaluated as tools for margin-focused guidance and tissue characterization when macroscopic imaging is insufficient to resolve microstructural features at the resection interface.<sup>77,79</sup> A systematic review of OCT in oncologic surgery summarizes evidence for intraoperative tissue assessment and highlights translational barriers, including limited standardization, variable acquisition conditions, and uncertainty about how best to integrate OCT-derived information into operative workflows.<sup>79</sup> A scoping review focused on AI in advancing OCT further reinforces that algorithmic interpretation is a key enabler for converting high-dimensional OCT data into actionable intraoperative outputs, while also underscoring the need for robust validation and clinically grounded study designs.<sup>77</sup>

Hybrid systems that fuse OCT with ML inference have been reported for breast cancer surgical margin visualization, illustrating a design pattern in which high-resolution imaging is paired with automated classification to generate decision-supportive outputs intraoperatively.<sup>78</sup> Translation of such systems into routine oncologic practice requires demonstration of robustness to blood, motion, and specular artifacts; prespecified uncertainty-aware operating points; and empirical evidence that OCT-based guidance improves margin-related outcomes without disproportionate resection or operative burden.<sup>77-79</sup>

#### **Robotic Assistance, Safety Monitoring, and the Boundary Between Guidance and Control**

Robotic platforms create opportunities for continuous performance monitoring and algorithmic safety layers that detect and mitigate hazards, but the evidentiary threshold increases substantially when systems move from passive guidance to action-constraining control.<sup>3,53,80</sup> AI-based hazard detection in robotic-assisted single-incision oncologic surgery exemplifies a safety-oriented direction by targeting the identification of hazardous states in a modality characterized by constrained instrument motion, limited triangulation, and heightened collision risk.<sup>10</sup>

Robotic data streams can also be used to evaluate and standardize operative performance, with potential downstream implications for oncologic outcomes and complication risk. A study associating skill and

errors with outcomes in robotic rectal cancer surgery illustrates how intraoperative performance metrics can be linked to clinical endpoints, supporting the concept that intraoperative analytics may provide actionable feedback for quality improvement.<sup>80,94</sup> At the same time, outcome-linked analytics emphasize the need to avoid overinterpreting surrogate performance measures unless they are validated against patient-centered endpoints and contextualized by case complexity, anatomy, and oncologic extent.<sup>3,53,80,81,94,95</sup>

Engineering frameworks for AI-enhanced processing of da Vinci robot-assisted video reinforce that translation requires not only accurate models but also reliable system-level integration, including secure data handling, deterministic latency, and compatibility with the robotic platform's data interfaces.<sup>82</sup> Systematic synthesis of AI integration into robotic cancer surgery indicates that most published systems remain at the level of feasibility or retrospective validation, with limited evidence from prospective studies and limited clarity regarding how ML outputs are incorporated into standardized intraoperative actions.<sup>3,53</sup> This evidence profile supports a conservative near-term posture: prioritize decision support, safety monitoring, and uncertainty-aware guidance before pursuing control-constraining behaviors that would require substantially higher safety assurance.<sup>3,10,53</sup>

#### **Translational Requirements: Real-Time Operation, Robustness, and Clinically Testable Utility**

Across intraoperative navigation and robotic assistance/control, translational readiness depends on demonstrating reliable real-time operation under routine variability, explicit handling of uncertainty, and measurable benefit in decision-relevant endpoints.<sup>3,53,56</sup> Feasibility studies in vascular and stereotactic navigation demonstrate that geometry-aware guidance can be operationalized in oncologic procedures, while prospective application of AI-assisted 3D navigation in lung surgery shows that intraoperative navigation can be evaluated in forward-looking designs rather than solely retrospective settings.<sup>8,9,12</sup> Prospective demonstration of real-time metastasis recognition during exploration further illustrates how intraoperative perception can be tied to immediate oncologic decisions, providing a model for future studies that specify workflow integration and decision thresholds a priori.<sup>24</sup>

For imaging-enhanced guidance (eg, fluorescence and NIR functional imaging), translation requires separating device- and protocol-specific variability from biologically meaningful signal and validating ML-derived overlays across acquisition settings, patient factors, and institutions.<sup>56,86,88,89</sup> For label-free modalities (eg, mass spectrometry and OCT), uncertainty-aware inference and robustness to intraoperative artifacts are central because both false reassurance and false alarms can produce clinically meaningful harm (eg, inadequate margins vs unnecessary tissue loss).<sup>11,77-79</sup> These requirements are emphasized across

systematic and specialty-focused reviews, which report that the clinical evidence for many AI-enabled intraoperative systems remains dominated by retrospective analyses and early feasibility work, with relatively few prospective evaluations tied to patient-centered outcomes.<sup>3,53,80,81,95</sup> Prospective application studies in navigation (eg, non-contact 3D guidance) and real-time disease detection (eg, AiLES) provide useful templates for study designs that connect real-time inference to standardized intraoperative actions and measurable outcomes.<sup>12,24</sup> Finally, because robotic surgery continues to expand across complex oncologic procedures (including pancreatic surgery), the bar for safe, clinically effective integration of intraoperative AI should remain anchored in workflow-aware validation, transparent uncertainty handling, and rigorous prospective evaluation rather than technical performance alone.<sup>3,10,53</sup>

In addition, outside the perioperative realm, there have been equal advancements in AI for psychiatry and mental health applications, including risk assessment (suicide attempts, relapse, and acute decompensation), symptom profiling, and treatment recommendation with multimodal data sources like longitudinal electronic health records, wearable and smartphone-based digital phenotyping, speech/language processing, and neuroimaging.<sup>96-98</sup> In precision oncology, the combination of radiology, digital pathology, and genomic profiling enabled by AI has facilitated tumor profiling, prediction of treatment response (including immunotherapy), and computational triage for molecular analyses, thereby propelling the transition of high-dimensional biomarkers into decision-making.<sup>99-101</sup> These parallel applications collectively point to a shift towards data-driven, personalized decision support systems, while also pointing out common issues in fairness, generalization, explainability, and validation that are just as applicable in surgical oncology.

### Conclusion

Across surgical oncology, AI/ML research increasingly spans the full perioperative continuum, with maturation that varies substantially by domain and by proximity to real-world clinical integration.<sup>53,95</sup> In clinical translation and decision support, recent work demonstrates that deployment and monitoring of prediction systems in routine colorectal cancer surgical workflows are feasible, but such implementation-level evidence remains uncommon to the volume of retrospective development studies.<sup>5,53,95</sup> Consistent with broader concerns in clinical prediction and medical AI, the primary constraint on translation is not proof-of-concept discrimination but the combination of generalizability, calibration, and clinically meaningful utility under routine conditions.<sup>42-44,48-50</sup>

Preoperative planning remains dominated by imaging-based methods, particularly segmentation and localization, where performance may appear strong in curated datasets yet degrade under

cross-site variation in acquisition protocols and label practices. The translational requirement is therefore not solely architectural sophistication but demonstrable robustness to dataset shift, transparent reporting of data provenance, and external validation that reflects target deployment heterogeneity.<sup>42-44,48-50</sup> In parallel, perioperative risk prediction models for thoracic and esophageal oncologic surgery increasingly integrate imaging with clinical variables, and emerging reports emphasize explainability; however, interpretability claims must be matched to calibrated risk estimates and decision-relevant evaluations to support counseling, triage, and perioperative planning.<sup>23,34,43,44</sup>

Intraoperative navigation and robotic assistance/control are advancing in tandem with surgical robotics and training-focused AI, yet the evidentiary base remains weighted toward feasibility demonstrations and educational applications rather than prospective, outcome-linked clinical utility.<sup>53,81</sup> The methodological gap is well captured by contemporary guidance: early-stage evaluations should explicitly report human-AI interaction, workflow integration, safety mitigations, and error-case behavior (DECIDE-AI), while later-stage interventional studies should adhere to AI-specific trial and protocol standards (CONSORT-AI, SPIRIT-AI).<sup>15,38,51</sup>

Future progress in surgical oncology AI will depend on aligning model development with deployment realities: multicenter and temporally robust evaluation, explicit calibration assessment and updating strategies, decision-analytic utility assessment at clinically relevant thresholds, and systematic appraisal of bias and subgroup performance when outcomes and labels reflect structural differences in access and care pathways. These requirements are increasingly codified in updated reporting and risk-of-bias instruments (TRIPOD+AI, PROBAST+AI) and in good ML practice guidance emphasizing lifecycle governance, monitoring, and safety assurance for medical-device ML systems.<sup>49,50</sup> When these evidentiary expectations are met, AI/ML systems in surgical oncology are more likely to transition from high in silico performance to reproducible, safe, and clinically useful tools that improve perioperative decision-making rather than merely predicting outcomes.

### List of Abbreviations

AI: Artificial intelligence  
 AiLES: Artificial intelligence Laparoscopic Exploration System  
 CDS: Clinical Decision Support  
 CT: Computed Tomography  
 DCA: Decision Curve Analysis  
 FDA: Food and Drug Administration  
 HER: Electronic Health Record  
 LLM: Large Language Model  
 MDR: Medical Device Regulation  
 ML: Machine Learning  
 MRI: Magnetic Resonance Imaging  
 NIR: Near-Infrared

NLP: Natural language processing  
 OCT: Optical coherence tomography  
 PCCP: Predetermined Change Control Plan  
 WSI: Whole-Slide Imaging

## References

- Williams TL, Saadat LV, Gonen M, Wei A, Do RK, Simpson AL. Radiomics in surgical oncology: applications and challenges. *Comput Assist Surg (Abingdon)*. 2021;26(1):85–96. <https://doi.org/10.1080/24699322.2021.1994014>
- Lin G, Li X, Wang D, Chen X. Advances in the application of three-dimensional reconstruction in thoracic surgery: a comprehensive review. *Thorac Cancer*. 2025;16(17):e70159. <https://doi.org/10.1111/1759-7714.70159>
- Leszczyńska A, Obuchowicz R, Strzelecki M, Seweryn M. The integration of artificial intelligence into robotic cancer surgery: a systematic review. *J Clin Med*. 2025;14(17):6181. <https://doi.org/10.3390/jcm14176181>
- Buck L, Kohler J, Risch J, Incesu RB, Hügelmann K, Weiss ML, et al. The emerging role of multimodal artificial intelligence in urological surgery. *Curr Oncol*. 2025;32(12):665. <https://doi.org/10.3390/curroncol32120665>
- Rosen AW, Ose I, Gögenur M, Andersen LP, Bojesen RD, Vogelsang RP, et al.; AID-SURG study group. Clinical implementation of an AI-based prediction model for decision support for patients undergoing colorectal cancer surgery. *Nat Med*. 2025;31(11):3737–48. <https://doi.org/10.1038/s41591-025-03942-x>
- Dal Cero M, Gibert J, Grande L, Gimeno M, Osorio J, Bencivenga M, et al.; On behalf of the Spanish Eurecca Esophagogastric Cancer Group and the European Gastrodata Study Group. International external validation of risk prediction model of 90-day mortality after gastrectomy for cancer using machine learning. *Cancers (Basel)*. 2024;16(13):2463. <https://doi.org/10.3390/cancers16132463>
- Vakili-Ojaroood M, Naseri A, Shirinzadeh-Dastgiri A, Saberi A, HaghighiKian SM, Rahmani A, et al. Ethical considerations and equipoise in cancer surgery. *Indian J Surg Oncol*. 2024;15(S3 Suppl 3):363–73. <https://doi.org/10.1007/s13193-024-02023-8>
- Kitaguchi D, Takeshita N, Matsuzaki H, Igaki T, Hasegawa H, Kojima S, et al. Real-time vascular anatomical image navigation for laparoscopic surgery: experimental study. *Surg Endosc*. 2022;36(8):6105–12. <https://doi.org/10.1007/s00464-022-09384-7>
- Ten Brink R, Schram R, Kats-Ugurlu G, Kwee T, Hemmer P, Havenga K, et al. Feasibility of optical stereotactic navigation for rectosigmoid cancer with deep learning-supported 3D modelling. *Eur J Surg Oncol*. 2025;51(11):110397. <https://doi.org/10.1016/j.ejso.2025.110397>
- Rus G, Andras I, Vaida C, Crisan N, Gherman B, Radu C, et al. Artificial intelligence-based hazard detection in robotic-assisted single-incision oncologic surgery. *Cancers (Basel)*. 2023;15(13):3387. <https://doi.org/10.3390/cancers15133387>
- Fooladgar F, Jamzad A, Connolly L, Santilli A, Kaufmann M, Ren K, et al. Uncertainty estimation for margin detection in cancer surgery using mass spectrometry. *Int J Comput Assist Radiol Surg*. 2022;17(12):2305–13. <https://doi.org/10.1007/s11548-022-02764-3>
- Li C, Song J, Yan R, Zhang B, Wang B, Fan J, et al. Non-contact artificial intelligence-assisted intraoperative 3D navigation technology prospective application study in lung cancer surgery. *J Thorac Dis*. 2025;17(11):9610–21. <https://doi.org/10.21037/jtd-2025-1136>
- Macheka S, Ng PY, Ginsburg O, Hope A, Sullivan R, Aggarwal A. Prospective evaluation of artificial intelligence (AI) applications for use in cancer pathways following diagnosis: a systematic review. *BMJ Oncol*. 2024;3(1):e000255. <https://doi.org/10.1136/bmjonc-2023-000255>
- Kawamoto K, Houlihan CA, Balas EA, Lobach DF. Improving clinical practice using clinical decision support systems: a systematic review of trials to identify features critical to success. *BMJ*. 2005;330(7494):765. <https://doi.org/10.1136/bmj.38398.500764.8F>
- Vasey B, Nagendran M, Campbell B, Clifton DA, Collins GS, Denaxas S, et al.; DECIDE-AI expert group. Reporting guideline for the early-stage clinical evaluation of decision support systems driven by artificial intelligence: DECIDE-AI. *BMJ*. 2022;28(5):924–33. <https://doi.org/10.1136/bmj-2022-070904>
- Alselaïm NA, Alsemari MA, Alyabsi M, Al-Mutairi AM. Factors associated with 30-day mortality and morbidity in patients undergoing emergency colorectal surgery. *Ann Saudi Med*. 2023;43(6):364–72. <https://doi.org/10.5144/0256-4947.2023.364>
- Xu J, Zhou J, Hu J, Ren Q, Wang X, Shu Y. Development and validation of a machine learning model for survival risk stratification after esophageal cancer surgery. *Front Oncol*. 2022;12:1068198. <https://doi.org/10.3389/fonc.2022.1068198>
- Bilimoria KY, Liu Y, Paruch JL, Zhou L, Kmiecik TE, Ko CY, et al. Development and evaluation of the universal ACS NSQIP surgical risk calculator: a decision aid and informed consent tool for patients and surgeons. *J Am Coll Surg*. 2013;217(5):833–42.e1. <https://doi.org/10.1016/j.jamcollsurg.2013.07.385>
- Abikenari M, Jain B, Xu R, Jackson C, Huang J, Bettogowda C, et al. Bridging imaging and molecular biomarkers in trigeminal neuralgia: toward precision diagnosis and prognostication in neuropathic pain. *Med Res Arch, [S.l.]*, v. 13, n. 5, May 2025. ISSN 2375-1924. Available at: <https://esmed.org/MRA/mra/article/view/6605>. Date accessed: 15 January 2026. <https://doi.org/10.18103/mra.v13i5.6605>
- Namavarian A, Gabinet-Equihua A, Deng Y, Khalid S, Ziai H, Deutsch K, et al. Length of stay prediction models for oral cancer surgery: machine learning, statistical and ACS-NSQIP. *Laryngoscope*. 2024;134(8):3664–72. <https://doi.org/10.1002/lary.31443>
- Kadomatsu Y, Emoto R, Kubo Y, Nakanishi K, Ueno H, Kato T, et al. Development of a machine learning-based risk model for postoperative complications of lung cancer surgery. *Surg Today*. 2024;54(12):1482–9. <https://doi.org/10.1007/s00595-024-02878-y>
- Cowling TE, Cromwell DA, Bellot A, Sharples LD, van der Meulen J. Logistic regression and machine learning predicted patient mortality from large sets of diagnosis codes comparably. *J Clin Epidemiol*. 2021;133:43–52. <https://doi.org/10.1016/j.jclinepi.2020.12.018>
- Klontzas ME, Ri M, Koltsakis E, Stenqvist E, Kalarakis G, Boström E, et al. Prediction of anastomotic leakage in esophageal cancer surgery: a multimodal machine learning model integrating imaging and clinical data. *Acad Radiol*. 2024;31(12):4878–85. <https://doi.org/10.1016/j.acra.2024.06.026>
- Chen H, Gou L, Fang Z, Dou Q, Chen H, Chen C, et al. Artificial intelligence-assisted real-time recognition of intra-abdominal metastasis during laparoscopic gastric cancer surgery. *NPJ Digit Med*. 2025;8(1):9. <https://doi.org/10.1038/s41746-024-01372-6>
- Barber EL, Garg R, Persenaire C, Simon M. Natural language processing with machine learning to predict outcomes after ovarian cancer surgery. *Gynecol Oncol*. 2021;160(1):182–6. <https://doi.org/10.1016/j.ygyno.2020.10.004>
- Rossi LA, Melstrom LG, Fong Y, Sun V. Predicting post-discharge cancer surgery complications via telemonitoring of patient-reported outcomes and patient-generated health data. *J Surg Oncol*. 2021;123(5):1345–52. <https://doi.org/10.1002/jso.26413>
- Yan YD, Wu XW, Li Y, Lin HW, Zhang ZT, Jia D, et al.; CRC-VTE investigators. Machine learning to predict venous thromboembolism after colorectal cancer surgery: a Chinese dynamic modelling study. *Int J Surg*. 2025; <https://doi.org/10.1097/S109759000000000004036>
- Ri M, Nunobe S, Narita T, Seto Y, Kawazoe Y, Ohe K, et al. Time-sequential prediction of postoperative complications after gastric cancer surgery using machine learning: a multicenter cohort study. *Gastric Cancer*. 2025;28(6):1273–81. <https://doi.org/10.1007/s10120-025-01658-y>
- Zeng S, Li L, Hu Y, Luo L, Fang Y. Machine learning approaches for the prediction of postoperative complication risk in liver resection patients. *BMC Med Inform Decis Mak*. 2021;21(1):371. <https://doi.org/10.1186/s12911-021-01731-3>
- Lin V, Tsouchnika A, Allakhverdiev E, Rosen AW, Gögenur M, Clausen JS, et al. Training prediction models for individual risk assessment of postoperative complications after surgery for colorectal cancer. *Tech Coloproctol*. 2022;26(8):665–75. <https://doi.org/10.1007/s10151-022-02624-x>

- 31 Lee S, Oh HJ, Yoo H, Kim CY. Machine learning insight: unveiling overlooked risk factors for postoperative complications in gastric cancer. *Cancers (Basel)*. 2025;17(7):1225. <https://doi.org/10.3390/cancers17071225>
- 32 Khan MF, Cahill RA. Peroperative personalised decision support and analytics for colon cancer surgery– Short report. *Eur J Surg Oncol*. 2021;47(2):477–9. <https://doi.org/10.1016/j.ejso.2020.04.010>
- 33 Li R, Zhao Z, Yu J. Development and validation of a machine learning model for predicting early postoperative complications after radical gastrectomy. *Front Oncol*. 2025;15:1631260. <https://doi.org/10.3389/fonc.2025.1631260>
- 34 Chen S, Deng T, Yang Q, Li J, Shen J, Luo X, et al. Development and validation of an explainable machine learning model for predicting postoperative pulmonary complications after lung cancer surgery: a machine learning study. *EClinicalMedicine*. 2025;86:103386. <https://doi.org/10.1016/j.eclinm.2025.103386>
- 35 Yang X, Dou F, Tang G, Xiu R, Zhao X. Interpretable machine learning model for predicting anastomotic leak after esophageal cancer surgery via LightGBM. *BMC Cancer*. 2025;25(1):976. <https://doi.org/10.1186/s12885-025-14387-3>
- 36 Hu J, Liu Q, He W, Wu J, Zhang D, Sun C, et al. Automated machine learning model for predicting anastomotic strictures after esophageal cancer surgery: a retrospective cohort study. *Surg Endosc*. 2025;39(6):3737–48. <https://doi.org/10.1007/s00464-025-11759-5>
- 37 Bektaş M, Burchell GL, Bonjer HJ, van der Peet DL. Machine learning applications in upper gastrointestinal cancer surgery: a systematic review. *Surg Endosc*. 2023;37(1):75–89. <https://doi.org/10.1007/s00464-022-09516-z>
- 38 Liu X, Cruz Rivera S, Moher D, Calvert MJ, Denniston AK, Chan AW, et al.; SPIRIT-AI and CONSORT-AI Working Group. Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: the CONSORT-AI extension. *Nat Med*. 2020;26(9):1364–74. <https://doi.org/10.1038/s41591-020-1034-x>
- 39 Rahman SA, Walker RC, Lloyd MA, Grace BL, van Boxel GI, Kingma BF, et al.; OCCAMS Consortium. Machine learning to predict early recurrence after oesophageal cancer surgery. *Br J Surg*. 2020;107(8):1042–52. <https://doi.org/10.1002/bjs.11461>
- 40 Abikenari M, Schonfeld E, Choi J, Kim LH, Lim M. Revisiting glioblastoma classification through an immunological lens: a narrative review. *Glioma*. 2024;7(2)3–9. [https://doi.org/10.4103/glioma.glioma\\_4\\_24](https://doi.org/10.4103/glioma.glioma_4_24)
- 41 Vargas LO, Himic V, Otaner F, Abikenari M, Chandar J, Govindarajan V, et al. Modulating the glioma microenvironment with laser interstitial thermal therapy: mechanisms and therapeutic implications. *J Neurooncol*. 2025;176(1):99. <https://doi.org/10.1007/s11060-025-05305-5>
- 42 Abikenari M, Enayati I, Fountain DM, Leite MI. Navigating glioblastoma therapy: a narrative review of emerging immunotherapeutics and small-molecule inhibitors. *Microbes Immun*. 2025;2(4):132–43. <https://doi.org/10.36922/mi.5075>
- 43 Van Calster B, McLernon DJ, van Smeden M, Wynants L, Steyerberg EW. Topic Group Evaluating diagnostic tests and prediction models of the STRATOS initiative. Calibration: the Achilles heel of predictive analytics. *BMC Med*. 2019;17(1):230. <https://doi.org/10.1186/s12916-019-1466-7>
- 44 Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. *Med Decis Making*. 2006;26(6):565–74. <https://doi.org/10.1177/0272989X06295361>
- 45 Caldon LJ, Collins KA, Reed MW, Sivell S, Austoker J, Clements AM, et al.; BresDex Group. Clinicians' concerns about decision support interventions for patients facing breast cancer surgery options: understanding the challenge of implementing shared decision-making. *Health Expect*. 2011;14(2):133–46. <https://doi.org/10.1111/j.1369-7625.2010.00633.x>
- 46 Yen RW, Durand MA, Harris C, Cohen S, Ward A, O'Malley AJ, et al. Text-only and picture conversation aids both supported shared decision making for breast cancer surgery: analysis from a cluster randomized trial. *Patient Educ Couns*. 2020;103(11):2235–43. <https://doi.org/10.1016/j.pec.2020.07.015>
- 47 Obermeyer Z, Powers B, Vogeli C, Mullainathan S. Dissecting racial bias in an algorithm used to manage the health of populations. *Science*. 2019;366(6464):447–53. <https://doi.org/10.1126/science.aax2342>
- 48 Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *BMJ*. 2015;350(4):g7594. <https://doi.org/10.1136/bmj.g7594>
- 49 Collins GS, Moons KG, Dhiman P, Riley RD, Beam AL, Van Calster B, et al. TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. *BMJ*. 2024;385:e078378. <https://doi.org/10.1136/bmj-2023-078378>
- 50 Moons KG, Damen JA, Kaul T, Hooft L, Andaur Navarro C, Dhiman P, et al. PROBAST+AI: an updated quality, risk of bias, and applicability assessment tool for prediction models using regression or artificial intelligence methods. *BMJ*. 2025;388:e082505. <https://doi.org/10.1136/bmj-2024-082505>
- 51 Cruz Rivera S, Liu X, Chan AW, Denniston AK, Calvert MJ, Darzi A, et al.; SPIRIT-AI and CONSORT-AI Working Group; SPIRIT-AI and CONSORT-AI Steering Group; SPIRIT-AI and CONSORT-AI Consensus Group. Guidelines for clinical trial protocols for interventions involving artificial intelligence: the SPIRIT-AI extension. *Nat Med*. 2020;26(9):1351–63. <https://doi.org/10.1038/s41591-020-1037-7>
- 52 Chan SY, Twhog P. Artificial intelligence in liver cancer surgery: predicting success before the first incision. *World J Gastroenterol*. 2025;31(16):107221. <https://doi.org/10.3748/wjg.v31.i16.107221>
- 53 International Medical Device Regulators Forum (IMDRF). Software as a Medical Device (SaMD): key definitions. *IMDRF/SaMD WG/ N10FINAL:2013*. Published December 18, 2013. Accessed February 25, 2026. <https://www.imdrf.org/sites/default/files/docs/imdrf/final/technical/imdrf-tech-131209-samd-key-definitions-140901.pdf>
- 54 International Medical Device Regulators Forum (IMDRF). Software as a Medical Device (SaMD): Clinical Evaluation. *IMDRF/SaMD WG/N41FINAL:2017*. Published September 21, 2017. Accessed February 25, 2026. [https://www.imdrf.org/sites/default/files/docs/imdrf/final/technical/imdrf-tech-170921-samd-n41-clinical-evaluation\\_1.pdf](https://www.imdrf.org/sites/default/files/docs/imdrf/final/technical/imdrf-tech-170921-samd-n41-clinical-evaluation_1.pdf)
- 55 García-Granero Á, Jerí-McFarlane S, Ochogavía A, Gamundí-Cuesta M, García-Granero E, González-Argenté FX. 3D reconstructions in rectal cancer. New tools for better diagnosis and surgical planning. *Cir Esp (Engl Ed)*. 2025;103(9):800198. <https://doi.org/10.1016/j.cireng.2025.800198>
- 56 US Food and Drug Administration. Software as a Medical Device (SaMD). Updated December 4, 2018. Accessed February 25, 2026. <https://www.fda.gov/medical-devices/digital-health-center-excellence/software-medical-device-samd>
- 57 Checcucci E, De Cillis S, Porpiglia F. 3D-printed models and virtual reality as new tools for image-guided robot-assisted nephron-sparing surgery: a systematic review of the newest evidences. *Curr Opin Urol*. 2020;30(1):55–64. <https://doi.org/10.1097/MOU.0000000000000686>
- 58 Chen X, Xu H, Qi Q, Sun C, Jin J, Zhao H, et al. AI-based chest CT semantic segmentation algorithm enables semi-automated lung cancer surgery planning by recognizing anatomical variants of pulmonary vessels. *Front Oncol*. 2022;12:1021084. <https://doi.org/10.3389/fonc.2022.1021084>
- 59 Wang Z, Sun H, Li J, Chen J, Meng F, Li H, et al. Preoperative Prediction of Axillary Lymph Node Metastasis in Breast Cancer Using CNN Based on Multiparametric MRI. *J Magn Reson Imaging*. 2022 Sep;56(3):700–9. <https://doi.org/10.1002/jmri.28082>
- 60 Tan Y, Feng LJ, Huang YH, Xue JW, Long LL, Feng ZB. A comprehensive radiopathological nomogram for the prediction of pathological staging in gastric cancer using CT-derived and WSI-based features. *Transl Oncol*. 2024;40:101864. <https://doi.org/10.1016/j.tranon.2023.101864>
- 61 Yao Y, Cen X, Gan L, Jiang J, Wang M, Xu Y, et al. Automated esophageal cancer staging from free-text radiology reports: large language model evaluation study. *JMIR Med Inform*. 2025;13:e75556. <https://doi.org/10.2196/75556>

- 62 Tang Z, Feng S, Liu Q, Ban Y, Zhang Y. Using pathomics-based model for predicting positive surgical margins in patients with esophageal squamous cell carcinoma: a comparative study of decision tree and nomogram. *Int J Gen Med*. 2024;17:5869–82. <https://doi.org/10.2147/IJGM.S495296>
- 63 Awad MH, Sanchez M, Abikenari MA. The values work of restorative ventures: the role of founders' embodied embeddedness with at-risk social groups. *J Bus Ventur Insights*. 2022;18:e00337. <https://doi.org/10.1016/j.jbvi.2022.e00337>
- 64 Kaida S, Murakami Y, Masaki Y, et al. Artificial intelligence-driven 3-dimensional simulation system for enhanced preoperative planning in gastric cancer surgery: a retrospective validation study. *J Gastrointest Surg*. Published online December 2025:102295. <https://doi.org/10.1016/j.gassur.2025.102295>
- 65 Hamabe A, Ishii M, Kamoda R, Sasuga S, Okuya K, Okita K, et al. Artificial intelligence-based technology to make a three-dimensional pelvic model for preoperative simulation of rectal cancer surgery using MRI. *Ann Gastroenterol Surg*. 2022;6(6):788–94. <https://doi.org/10.1002/ags3.12574>
- 66 Aisu Y, Okada T, Itatani Y, Masuo A, Tani R, Fujimoto K, et al. Automatic segmentation of male pelvic floor soft tissue structures for anatomical simulation and morphological assessment in lower rectal cancer surgery. *Tech Coloproctol*. 2025;29(1):176. <https://doi.org/10.1007/s10151-025-03218-z>
- 67 Wang F, Xiao C, Jia T, Pan L, Du F, Wang Z. Hepatobiliary surgery based on intelligent image segmentation technology. *Open Life Sci*. 2023 Aug;18(1):20220674. <https://doi.org/10.1515/biol-2022-0674>
- 68 Hofmann FO, Heiliger C, Tschaidse T, Jarmusch S, Auhage LA, Aghamaliyev U, et al. Validation of body composition parameters extracted via deep learning-based segmentation from routine computed tomographies. *Sci Rep*. 2025;15(1):11909. <https://doi.org/10.1038/s41598-025-96238-6>
- 69 Suthakaran R, Cao K, Arafat Y, Yeung J, Chan S, Master M, et al. Body composition assessment by artificial intelligence can be a predictive tool for short-term postoperative complications in Hartmann's reversals. *BMC Surg*. 2024 Apr;24(1):111. <https://doi.org/10.1186/s12893-024-02408-0>
- 70 Kim J, Han SH, Kim HI. Detection of sarcopenic obesity and prediction of long-term survival in patients with gastric cancer using preoperative computed tomography and machine learning. *J Surg Oncol*. 2021;124(8):1347–55. <https://doi.org/10.1002/jso.26668>
- 71 Amparore D, Piramide F, De Cillis S, Verri P, Piana A, Pecoraro A, et al.; Renal Cancer Working Group of the Young Academic Urologists (YAU) and European Association of Urology (EAU). Robotic partial nephrectomy in 3D virtual reconstructions era: is the paradigm changed? *World J Urol*. 2022;40(3):659–70. <https://doi.org/10.1007/s00345-022-03964-x>
- 72 Cheng H, Xu H, Peng B, Huang X, Hu Y, Zheng C, et al. Illuminating the future of precision cancer surgery with fluorescence imaging and artificial intelligence convergence. *NPJ Precis Oncol*. 2024;8(1):196. <https://doi.org/10.1038/s41698-024-00699-3>
- 73 Iresjö BM, Smedh U, Engström C, Persson J, Mårtensson C, Lundholm K. A randomized translational study on protein and glucose metabolism in skeletal muscles evaluated by gene-ontology, following preoperative oral carbohydrate loading compared to overnight peripheral parenteral nutrition (PPN) before major cancer surgery. *J Transl Med*. 2024;22(1):675. <https://doi.org/10.1186/s12967-024-05484-1>
- 74 van Oosterom MN, van Leeuwen SI, Mazzone E, Dell'Oglio P, Buckle T, van Beurden F, et al. Click-on fluorescence detectors: using robotic surgical instruments to characterize molecular tissue aspects. *J Robot Surg*. 2023;17(1):131–40. <https://doi.org/10.1007/s11701-022-01382-0>
- 75 Müller D, Stier R, Straatman J, Babic B, Schiffmann L, Eckhoff J, et al. ICG-lymphknoten-mapping in der Tumorchirurgie des oberen Gastrointestinaltrakts. *Chirurgie (Heidelb)*. 2022;93(10):925–33. <https://doi.org/10.1007/s00104-022-01659-y>
- 76 Hoogendam JP, Hobbelenk MG, Veldhuis WB, Verheijen RH, van Diest PJ, Zweemer RP. Preoperative sentinel node mapping with (99m) Tc-nanocolloid SPECT-CT significantly reduces the intraoperative sentinel node retrieval time in robot assisted laparoscopic cervical cancer surgery. *Gynecol Oncol*. 2013;129(2):389–94. <https://doi.org/10.1016/j.ygyno.2013.02.020>
- 77 Alikarami M, Faraj TA, Hama NH, Hosseini AS, Habibi P, Samiei Mosleh I, et al. Artificial intelligence in advancing optical coherence tomography for disease detection and cancer diagnosis: A scoping review. *Eur J Surg Oncol*. 2025;51(9):110188. <https://doi.org/10.1016/j.ejso.2025.110188>
- 78 Levy Y, Rempel D, Nguyen M, Yassine A, Sanati-Burns M, Salgia P, et al. The fusion of wide field optical coherence tomography and AI: advancing breast cancer surgical margin visualization. *Life (Basel)*. 2023;13(12):2340. <https://doi.org/10.3390/life13122340>
- 79 Pavone M, Innocenzi C, Carles E, Bizzarri N, Moro F, Ferrari FA, et al. Cutting edge microscopic intraoperative tissue assessment for guidance in oncologic surgery: a systematic review of the role of optical coherence tomography. *Ann Surg Oncol*. 2025;32(3):2191–205. <https://doi.org/10.1245/s10434-024-16632-8>
- 80 Abikenari M, Awad MH, Korouri S, Mohseni K, Abikenari D, Freichel R, et al. Reframing clinical AI evaluation in the era of generative models: toward multidimensional, stakeholder-informed, and safety-centric frameworks for real-world health care deployment. *PJS*. 2025;11:100089. <https://doi.org/10.70389/PJS.100089>
- 81 Tanzim U, Khan IA, Abikenari M, Al-Deen RH, Miah L, Blaaza M, et al. Transforming patient-provider communication: the role of artificial intelligence in advancing health literacy – a comprehensive review. *PJS*. 2025;13:100095. <https://doi.org/10.70389/PJS.100095>
- 82 Cheng Q, Dong Y (Hussein AF, editor). Da Vinci robot-assisted video image processing under artificial intelligence vision processing technology. *Comput Math Methods Med*. 2022. pp. 1–10. <https://doi.org/10.1155/2022/2752444>
- 83 Ebigbo A, Mendel R, Scheppach MW, Probst A, Shahidi N, Prinz F, et al. Vessel and tissue recognition during third-space endoscopy using a deep learning algorithm. *Gut*. 2022;71(12):2388–90. <https://doi.org/10.1136/gutjnl-2021-326470>
- 84 Checcucci E, Piana A, Volpi G, Piazzolla P, Amparore D, De Cillis S, et al. Three-dimensional automatic artificial intelligence driven augmented-reality selective biopsy during nerve-sparing robot-assisted radical prostatectomy: A feasibility and accuracy study. *Asian J Urol*. 2023;10(4):407–15. <https://doi.org/10.1016/j.ajur.2023.08.001>
- 85 Park SH, Park HM, Baek KR, Ahn HM, Lee IY, Son GM. Artificial intelligence based real-time microcirculation analysis system for laparoscopic colorectal surgery. *World J Gastroenterol*. 2020;26(44):6945–62. <https://doi.org/10.3748/wjg.v26.i44.6945>
- 86 Boland PA, Hardy NP, Moynihan A, McEntee PD, Loo C, Fenlon H, et al. Intraoperative near infrared functional imaging of rectal cancer using artificial intelligence methods – now and near future state of the art. *Eur J Nucl Med Mol Imaging*. 2024;51(10):3135–48. <https://doi.org/10.1007/s00259-024-06731-9>
- 87 Checcucci E, Vecchia A, Puliatti S, De Backer P, Piazza P, Kowalewski KF, et al. Metaverse in surgery – origins and future potential. *Nat Rev Urol*. 2026;23(1):50–63. <https://doi.org/10.1038/s41585-024-00941-4>
- 88 Won NJ, Bartling M, La Macchia J, Markevich S, Holtshousen S, Jagota A, et al. Deep learning-enabled fluorescence imaging for surgical guidance: in silico training for oral cancer depth quantification. *J Biomed Opt*. 2025;30(S1 Suppl 1):S13706. <https://doi.org/10.1117/1.JBO.30.S1.S13706>
- 89 Kurosawa H, Won NJ, Wunder JB, Patil S, Bartling M, Najjar E, et al. Deep learning-enabled fluorescence imaging for oral cancer margin classification in preclinical models. *J Biomed Opt*. 2025;30(S3 Suppl 3):S34109. <https://doi.org/10.1117/1.JBO.30.S3.S34109>
- 90 Abosheisha M, Nasr E, Ali M, Tamanna R, Bin Halim S, Rakib Hasan M, et al. Near-infrared fluorescence imaging in general surgery: applications in vascularization, tumor margin detection, and biliary anatomy. *Cureus*. 2025;17(9):e92194. <https://doi.org/10.7759/cureus.92194>
- 91 Müller DT, Schiffmann LM, Reisewitz A, Chon SH, Eckhoff JA, Babic B, et al. Mapping the lymphatic drainage pattern of esophageal cancer with near-infrared fluorescent imaging during robotic assisted minimally invasive Ivor Lewis esophagectomy (RAMIE) – First results of the prospective ESOMAP feasibility trial. *Cancers (Basel)*. 2023;15(8):2247. <https://doi.org/10.3390/cancers15082247>
- 92 Connolly L, Fooladgar F, Jamzad A, Kaufmann M, Syeda A, Ren K, et al. ImSpect: image-driven self-supervised learning for surgical margin evaluation with mass spectrometry. *Int J Comput Assist Radiol Surg*. 2024;19(6):1129–36. <https://doi.org/10.1007/s11548-024-03106-1>

- 93 Connolly L, Jamzad A, Kaufmann M, Farquharson CE, Ren K, Rudan JF, et al. Combined mass spectrometry and histopathology imaging for perioperative tissue assessment in cancer surgery. *J Imaging*. 2021;7(10):203. <https://doi.org/10.3390/jimaging7100203>
- 94 Boal M, Reali C, Duhoky R, Gill T, Khan J, Miskovic D, et al. Association of skill and errors with outcomes in robotic rectal cancer surgery. *Surg Endosc*. 2025. <https://doi.org/10.1007/s00464-025-12393-x>.
- 95 Quero G, Mascagni P, Kolbinger FR, Fiorillo C, De Sio D, Longo F, et al. Artificial intelligence in colorectal cancer surgery: present and future perspectives. *Cancers (Basel)*. 2022;14(15):3803. <https://doi.org/10.3390/cancers14153803>.
- 96 Siddarth P, Abikenari M, Grzenda A, Cappelletti M, Oughli H, Liu C, et al. Inflammatory markers of geriatric depression response to Tai Chi or Health Education Adjunct Interventions. *Am J Geriatr Psychiatry*. 2023;31(1):22–32. <https://doi.org/10.1016/j.jagp.2022.08.004>
- 97 Ajam Oughli H, Siddarth P, Lum M, Tang L, Ito B, Abikenari M, et al. Peripheral Alzheimer's disease biomarkers are related to change in subjective memory in older women with cardiovascular risk factors in a trial of yoga vs. memory training: lien établi entre les biomarqueurs périphériques de la maladie d'Alzheimer et l'amélioration de la mémoire subjective chez les femmes âgées présentant des facteurs de risque cardiovasculaire dans le cadre d'un essai comparant le yoga à l'entraînement de la mémoire. *Can J Psychiatry*. 2025;0(0):7067437251343291. <https://doi.org/10.1177/07067437251343291>
- 98 Basereh M, Abikenari MA, Sadeghzadeh S, et al. ConvNeXt-driven detection of Alzheimer's disease: a benchmark study on expert-annotated alzaset mri dataset across anatomical planes. *Diagnostics (Basel)*. 2025;15(23):2997. Published November 25, 2025. <https://doi.org/10.3390/diagnostics15232997>.
- 99 Luo X, Wang X, Eweje F, Zhang X, Yang S, Quinton R, et al. Ensemble learning of foundation models for precision oncology. *arXiv preprint*. 2025; arXiv:2508.16085.
- 100 Abikenari M, Liu J, Ha JH, Annagiri S, Himic V, Medikonda R, et al. Emerging trends in cell-based therapies: contemporary advances and ethical considerations in translational neurosurgical oncology. *J Neurooncol*. 2025;175(1):1–20. <https://doi.org/10.1007/s11060-025-05170-2>
- 101 Li Y, Wu X, Fang D, Luo Y. Informing immunotherapy with multi-omics driven machine learning. *NPJ Digit Med*. 2024;7(1):67. <https://doi.org/10.1038/s41746-024-01043-6>.