MDFORLIVES IN-DEPTH INSIGHT REPORT

AI in the Operating Room: How Surgeons Balance Alerts, Judgment and Liability

A global peer report on where intraoperative AI adds value, how surgeons resolve disagreement, which real-world conditions expose model limits, and why workflow, training, and accountability shape adoption.

Surgeons · 6 countries · Global peer perspectives

64	88	6	72.7%
Complete responses	Survey records	Countries represented	Completion rate

THE CLINICAL QUESTION

When intraoperative AI disagrees with surgical judgment, how are surgeons deciding what to trust and what to verify?

The dominant pattern is neither automatic deference to the software nor automatic rejection. 40.3% pause to seek a secondary maneuver or alternative imaging, 19.4% consult a colleague or assistant, and 29.9% continue on manual intuition. Only 10.4% say they alter the surgical trajectory purely for defensive medicine and liability mitigation.

That response sits inside a broader transition. 73.5% report pilot-only use or no AI integration in their primary operating suite. At the same time, respondents already identify clear use cases, failure conditions, automation-bias risks, liability concerns, and safeguards that shape whether an algorithm can become a trusted intraoperative signal.

> **Coverage:** United States, United Kingdom, Canada, Italy, France, Germany. The survey contains 88 records, including 64 complete and 24 partial responses, for a 72.7% completion rate. Findings describe the opinions and reported practice patterns of participating surgeons. Question-level response bases vary because not every survey record answered every item.

EXECUTIVE SUMMARY

The pattern in one view

The survey captures a profession negotiating AI before deep integration becomes routine. Surgeons see value, but they want defined tasks, verifiable disagreement pathways, realistic performance in complex anatomy, and governance that protects both clinical judgment and professional accountability.

01 / ADOPTION

Deep OR integration remains uncommon

73.5% are in pilot-only or no-integration environments, showing that routine intraoperative AI is still uneven.

02 / UTILITY

Phase recognition leads current value

43.3% choose automated surgical phase recognition as the application with the greatest current clinical utility or safety return.

03 / FAILURE

Complex anatomy is the stress test

54.5% identify distorted anatomy or multi-pathology as the most frequent cause of AI visual failure.

04 / SAFEGUARDS

Governance and training lead

74.2% choose either shared-liability frameworks or mandatory unplugged training as the most important safeguard.

> **Reading across the findings:** The strongest pattern is bounded automation. Surgeons appear more willing to use AI when its role is defined, disagreement can be escalated, uncertainty is visible, manual skill remains intact, and responsibility is clear before a complication tests the policy.

SECTION 01

AI integration remains transitional rather than routine

CONTEXT

Intraoperative AI is often discussed as though the operating room has already crossed into real-time algorithmic assistance. The survey shows a more uneven deployment landscape.

DATA

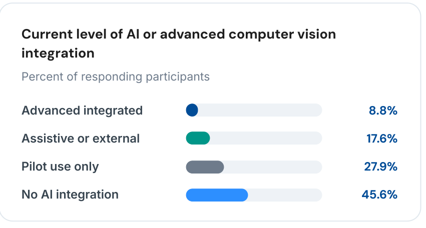
8.8% report advanced integrated AI, 17.6% report assistive or external AI, 27.9% report pilot use only, and 45.6% report no AI integration in their primary operating suite.

INSIGHT

The profession is forming opinions about autonomy, liability, and training while many surgeons still work in environments where AI is absent or experimental.

WHY IT MATTERS

Adoption policy can mature ahead of direct user experience. That raises the value of phased implementation, competency development, and clear limits on intended use.



The deployment distribution also reframes the rest of the report. The survey is not measuring a mature AI-native operating environment. It is capturing how surgeons imagine and experience the transition from conventional imaging to increasingly algorithm-assisted workflows.

This is clinically important because experience with a post-operative analytics platform is not the same as experience with a real-time tissue overlay. The closer an AI function moves to the moment of irreversible action, the more consequential trust, latency, false positives, override rules, and liability become.

The 2026 AORN guideline is relevant to this transition because it frames AI integration as a lifecycle and governance problem, not merely a technology purchase. Its emphasis on clinical oversight, education, privacy, safety, and governance mirrors the implementation pressures visible in this survey.

SECTION 02

Current utility is strongest where AI organizes the operation rather than directs the dissection

CONTEXT

Not every intraoperative AI use case carries the same clinical consequence. Phase recognition, anatomical mapping, early-warning alerts, and technique benchmarking sit at different distances from a real-time operative decision.

DATA

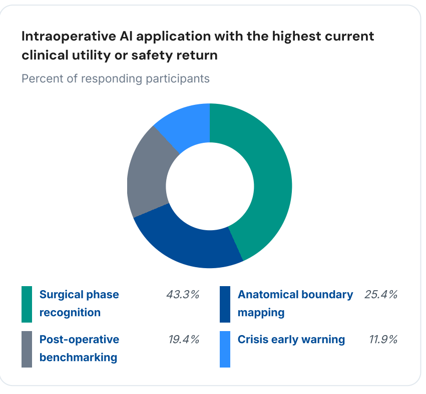
Surgical phase recognition leads at 43.3%, followed by real-time anatomical boundary mapping at 25.4%, automated post-operative benchmarking at 19.4%, and intraoperative crisis early warning at 11.9%.

INSIGHT

The leading use case is operational rather than autonomous. Surgeons appear to place the clearest value on AI that recognizes where the team is in the procedure without taking over the next technical action.

WHY IT MATTERS

The adoption sequence may begin with tasks that improve workflow or measurement and only later expand toward higher-stakes real-time recommendations.



Phase recognition can support logistics, documentation, room turnover, or team coordination while leaving the surgeon's tissue-level decision intact. Anatomical boundary mapping moves closer to the operative field and therefore asks for greater confidence in image quality and landmark recognition.

The lower selection of crisis early warning is notable because it is the most time-sensitive application in the list. A safety alert that arrives late, misreads anatomy, or generates too many false positives can create a different kind of cognitive load from a post-operative benchmark.

The pattern therefore separates technical novelty from practical value. In this cohort, the most useful AI is not necessarily the most autonomous AI.

SECTION 03

Human-AI disagreement is being treated as a verification problem

CONTEXT

The most revealing moment for intraoperative decision support is not when the algorithm agrees with the surgeon. It is when a hazard flag contradicts tactile feedback and clinical intuition during a complex dissection.

DATA

40.3% pause and obtain a secondary maneuver or alternative imaging, 29.9% continue based on manual intuition, 19.4% consult a senior colleague or assistant, and 10.4% change the surgical trajectory purely for defensive medicine and liability mitigation.

INSIGHT

The dominant response is neither surrender nor dismissal. It is escalation: create more evidence before deciding which signal to trust.

WHY IT MATTERS

The safest interface may be one that supports a clear disagreement pathway rather than one that simply produces more alerts.

This is a distinctly surgical form of human-in-the-loop behavior. In many diagnostic contexts, there is time to order another test. Intraoperatively, the surgeon may need an answer immediately, making the design of verification pathways part of the clinical utility of the system.

The 29.9% who continue on manual intuition also matter. They indicate that tactile feedback and accumulated operative experience remain strong sources of authority even in an AI-enabled environment.

At the same time, the 10.4% who alter the trajectory purely for defensive reasons show how legal context can influence technical behavior. An alert does not need to be clinically persuasive to change the operation if ignoring it feels legally exposed.

SECTION 04

Distorted anatomy is the real-world stress test for surgical computer vision

CONTEXT

Computer vision depends on visual patterns. Surgery frequently produces exactly the conditions that make those patterns unstable: inflammation, scarring, prior procedures, anatomical variation, bleeding, smoke, steam, and overlapping instruments.

DATA

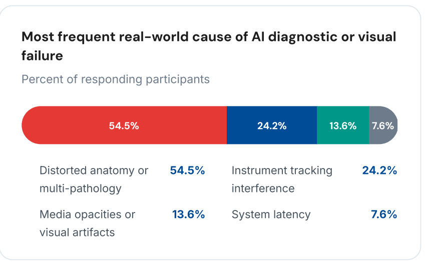
54.5% identify concomitant multi-pathology or distorted anatomy as the most frequent cause of AI diagnostic failure, followed by instrument tracking interference at 24.2%, media opacities and visual artifacts at 13.6%, and system latency at 7.6%.

INSIGHT

The hardest cases for the surgeon may also be the least representative cases for the model. That is a more consequential limitation than simple camera noise.

WHY IT MATTERS

Clinical confidence depends on how the system behaves outside ideal anatomy, especially when uncertainty is high and the potential consequence of a false landmark is significant.



The finding also changes the interpretation of alert fatigue. False positives may not arise only from a poorly calibrated threshold. They can arise because the anatomy itself no longer resembles the training distribution that underpins the model's segmentation.

Instrument tracking interference adds a second layer. The surgical field is dynamic, and object detection can be challenged by crossed instruments or occlusion. A model can be technically accurate in a clean frame and still become unreliable during the most active moments of a case.

FDA lifecycle guidance for AI-enabled medical device software emphasizes risk management across development and deployment. For intraoperative systems, this survey suggests that real-world robustness and uncertainty handling deserve the same attention as average performance metrics.

SECTION 05

Liability and automation bias sit on both sides of the surgeon-AI relationship

CONTEXT

Two different risks emerge when AI enters the OR. One is legal: what happens when an explicit warning is overridden and a complication follows? The other is cognitive: what happens when surgeons become too dependent on automated prompts?

DATA

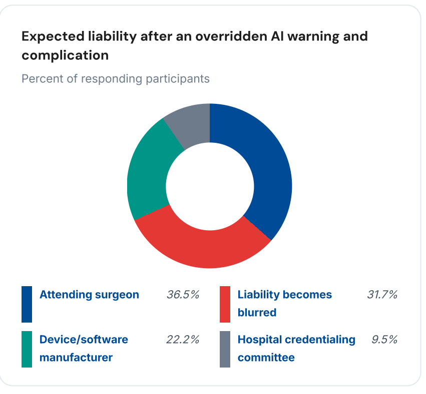
On liability, 36.5% place the bulk of responsibility on the attending surgeon, 31.7% expect blurred liability, 22.2% place it on the manufacturer, and 9.5% on the hospital credentialing committee. On automation bias, 43.5% identify trainee skill erosion as the greatest risk and 29.0% identify alert fatigue among experienced surgeons.

INSIGHT

The surgeon is caught between under-reliance and over-reliance. Ignore an alert and the digital record may matter; follow alerts too readily and clinical judgment or skill development may weaken.

WHY IT MATTERS

Safe adoption requires rules for both directions of error: when AI is ignored and when AI is trusted too much.



The liability split is revealing because no single non-surgeon actor dominates. The software manufacturer receives 22.2%, while the hospital receives 9.5%. Respondents therefore keep much of the accountability close to the operating surgeon, even when the algorithm generated the warning.

The automation-bias result adds a generational dimension without requiring a simplistic younger-versus-older narrative. 43.5% worry about trainees bypassing native spatial reasoning and tactile development, while 29.0% worry about experienced surgeons becoming desensitized to repeated false-positive alerts.

Both concerns point to the same design principle: an AI system can alter cognition even when it never performs an autonomous surgical step.

SECTION 06

The adoption bottleneck is as much governance and workflow as it is technology

CONTEXT

A technically capable AI tool still has to fit into the operating room's hardware, software, economics, privacy rules, credentialing, and training environment.

DATA

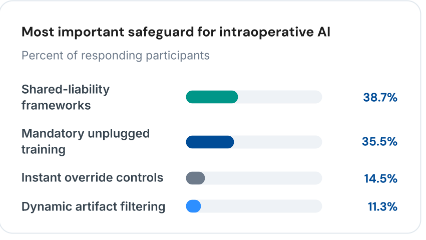
Integration/EHR friction and reimbursement limitations tie as the largest adoption barriers at 30.6% each, followed by data sovereignty and privacy at 25.8% and vendor lock-in at 12.9%. For safeguards, 38.7% choose shared-liability frameworks and 35.5% choose mandatory unplugged training.

INSIGHT

The two leading safeguards are not software features. They are an accountability framework and a training rule. The survey therefore treats governance as part of the clinical infrastructure of AI.

WHY IT MATTERS

Hospitals can purchase an algorithm without creating a trusted workflow. Sustainable adoption requires integration, funding, data governance, competency, and clear responsibility around use and override.



The barrier distribution is unusually even. No single obstacle overwhelms the others. Instead, workflow integration, reimbursement, privacy, and vendor constraints form a cluster, suggesting that deployment can fail at several institutional handoffs.

The safeguard results are more concentrated. Shared liability at 38.7% and unplugged training at 35.5% together account for 74.2% of responses. That concentration implies that surgeons want a defined professional operating model around the tool before they ask for more technical automation.

The future-role question points in the same direction. 35.5% expect routine components to become autonomous while complex reconstruction remains manual, 25.8% expect the surgeon to become an intraoperative supervisor, 29.0% expect the role to remain largely unchanged, and only 9.7% want AI confined to training and post-operative use. The leading future is therefore hybrid rather than absolute.

PRACTICE PATTERN

What may be sitting underneath the responses

These are interpretive patterns, not clinical recommendations. They describe tensions that appear when the survey findings are read together.

Adoption is task-specific

Maturity lens

Surgeons appear more open to AI that supports workflow or visible measurement than to broad autonomous intervention.

Verification protects trust

Disagreement lens

When machine and surgeon differ, additional evidence is the dominant response. Trust is procedural, not binary.

Complex anatomy is the edge case that matters most

Robustness lens

Distortion, scarring, inflammation, and overlapping pathology challenge the algorithm precisely when the surgical field is difficult.

Governance is part of the clinical device

Accountability lens

Liability, competency, override authority, reimbursement, integration, and data governance shape whether technical capability becomes clinical adoption.

CLOSING PERSPECTIVE

The future of intraoperative AI is likely to be negotiated one bounded task at a time

The survey does not support a simple story in which surgeons either trust AI or resist it. Instead, they draw boundaries around the kind of work the system is doing, the anatomy it is seeing, the consequence of being wrong, and the process available when its interpretation conflicts with their own.

That is why surgical phase recognition leads current utility while crisis early warning trails. It is why distorted anatomy dominates the failure question. It is why disagreement produces secondary evidence. And it is why shared liability and manual training outrank technical artifact filtering as safeguards.

The future-role responses reinforce that measured transition. The largest group expects a split model in which routine, repetitive components become autonomous while complex reconstruction remains manual. Another quarter foresee the surgeon as an intraoperative supervisor. Almost three in ten expect the role itself to remain largely unchanged. Only a small minority want AI restricted entirely to training and post-operative use.

The practical direction is therefore hybrid. AI can become more present in the OR without becoming the ultimate authority in the OR. Its clinical legitimacy will depend on whether it communicates uncertainty, survives difficult anatomy, integrates into existing systems, preserves manual expertise, and enters a governance model that makes responsibility understandable before something goes wrong.

CONCEPT FAQ

Clinical questions about intraoperative AI and surgical decision support

What is intraoperative AI?

Intraoperative AI refers to software that analyzes information during a procedure, such as live video, imaging, device data or workflow signals, to support tasks such as phase recognition, anatomical identification, risk detection or decision support while the operation is underway.

How is intraoperative AI different from robotic surgery?

Robotic surgery provides a mechanical platform through which a surgeon operates. AI is a computational layer that can interpret data, recognize patterns or generate prompts. The two can be integrated, but a robotic platform is not automatically an AI system.

Why can computer-vision AI struggle during real surgery?

Surgical video is highly variable. Blood, smoke, steam, lens contamination, instrument overlap, inflammation, scar tissue and altered anatomy can change the visual patterns available to an algorithm. Safe systems therefore need to communicate uncertainty and be tested under realistic operating conditions.

What is automation bias in the operating room?

Automation bias is the tendency to place too much weight on an automated recommendation. In surgery, it can appear as over-reliance on an alert, reduced independent checking or, over time, weaker development of manual pattern recognition if trainees rarely practise without digital assistance.

Should AI replace a surgeon's clinical judgment?

Current perioperative guidance frames AI as a support to clinical care rather than a substitute for professional judgment. Human oversight, clear intended use, training, governance and the ability to question or override the system remain central to safe integration.

What should hospitals evaluate before deploying AI in the operating room?

Evaluation should extend beyond model accuracy. Organizations need to consider patient safety, bias, privacy and security, staff education, workflow integration, clinical oversight, monitoring after deployment and clear governance for how AI outputs are used and documented.

Clinical context sources: AORN Guideline for Integration of Artificial Intelligence. These FAQs provide general educational context and do not replace specialty guidance or patient-specific clinical judgment.

SELECTED EXTERNAL CONTEXT

1. AORN, Guideline for Integration of Artificial Intelligence in Surgical Care, released June 2026
2. US FDA, Clinical Decision Support Software, Final Guidance, January 2026
3. US FDA, Good Machine Learning Practice for Medical Device Development: Guiding Principles
4. US FDA, Artificial Intelligence-Enabled Device Software Functions: Lifecycle Management and Marketing Submission Recommendations

MDForLives

The Global Pulse of Healthcare

Real Opinions to Insights. Validated by Data.