ISSN: 2641-6271

DOI: 10.33552/OAJAP.2026.08.000696

Open Access Journal Of Addiction and Psychology

Iris Publishers

Research Article

Copyright © All rights are reserved by Osvaldo Mazza MD PhD

# Drawing the Flight Path of Surgery: A Comparative Analysis of Aviation Flight Planning and Surgical Preoperative Optimization

**Osvaldo Mazza[1*], Marianna Mazza[2], Giuseppe Marano[3], Pietro Persiani[4] and Marco Crostelli[5]**

[1]*Trauma and Special Orthopaedic Surgey Unit, Ospedale Bambino Gesù, Romae, Italy; Dipartimento di scienze Anatomiche, Istologiche, Medico Legali e dell'Apparato Locomotore, La Sapienza University, Rome, Italy*

[2]*Marianna Mazza MD PhD Unit of Psychiatry, Fondazione Policlinico Universitario Agostino Gemelli IRCCS, Department of Neurosciences, Università Cattolica del Sacro Cuore, Italy*

[3]*Giuseppe Marano Unit of Psychiatry, Fondazione Policlinico Universitario Agostino Gemelli IRCCS, Department of Neurosciences, Università Cattolica del Sacro Cuore, Italy*

[4]*Pietro Persiani MD PhD, Dipartimento di Scienze Anatomiche, Istologiche, Medico Legali e dell'Apparato Locomotore, La Sapienza, University, Roma, Italy*

[5]*Marco Crostelli MD PhD, Spine Surgery Unit, Ospedale, Cristo Re, Italy*

***Corresponding author:** Osvaldo Mazza, Trauma and Special Orthopaedic Surgery Unit, Ospedale, Bambino Gesù, Dipartimento di Scienze Anatomiche, Istologiche, Medico Legali e dell'Apparato Locomotore, La Sapienza, University, Roma, Italy

**Received Date:** October 14, 2026

**Published Date:** November 02, 2026

## Abstract

**Background:** Preoperative planning in surgery and pre-flight planning in commercial and military aviation share a common structural architecture. Both domains require the reconciliation of incomplete information, finite resources and escalating environmental risk within compressed time frames, and both are vulnerable to the same class of failure: the cognitive overload of a single decision-maker operating inside a system whose complexity has outgrown individual human capacity. While much of the translational literature on aviation-derived safety has concentrated on intraoperative checklists and team training, the intellectual parallel between pre-flight planning and preoperative surgical strategy has been comparatively neglected.

**Objective:** To perform a structured comparative analysis of the phases of aviation flight planning and their direct counterparts in surgical preoperative optimization, and to evaluate the evidence supporting the transfer of aviation-derived planning and debriefing methodologies into surgical practice.

**Methods:** A narrative comparative review was conducted following a structured search of PubMed/MEDLINE, Scopus, Web of Science and the Cochrane Library, supplemented by regulatory and industry documentation from the Federal Aviation Administration and the World Health Organization, and by the principal biomedical engineering literature on computational surgical planning. Searches covered the period January 1990 to October 2026. One hundred and eight records were screened and 106 sources meeting pre-specified relevance criteria were retained. The two workflows were decomposed into seven functionally comparable domains, and each domain was analyzed for structural homology, available outcome evidence, and transferability constraints. Spinal instrumentation-specifically the accuracy of pedicle screw placement-was adopted as the principal worked example in the planning domains, replacing the laparoscopic cholecystectomy exemplar used in earlier formulations of this framework.

This work is licensed under Creative Commons Attribution 4.0 License

## Abstract

**Results:** Six of the seven domains demonstrated substantive structural homology: information gathering maps onto diagnostic imaging, three-dimensional reconstruction and patient-specific virtual planning; weight-and-balance and fuel planning map onto surgical armamentarium, implant sizing and blood product logistics; route and weather planning with alternate selection maps onto strategic surgical routing and pre-defined conversion trigger points; the decision funnel maps onto the contraction of intraoperative decision latitude as unrecognized deviation accumulates; emergency procedures and quick reference handbooks map onto standardized algorithms for major intraoperative complications; and pre-departure checklists map onto the WHO Surgical Safety Checklist. The seventh domain, systematic non-punitive post-procedure debriefing, is where the analogy becomes prescriptive rather than descriptive, and identifies the principal cultural divergence between the two industries. Evidence for the checklist domain is strong, with pooled data indicating reductions in postoperative morbidity and mortality in the order of one third. Evidence in the navigated instrumentation domain is quantitative and robust: pooled accuracy across 78 studies and 51,161 screws was 95.5 per cent for CT navigation, 93.1 per cent for freehand, 91.5 per cent for fluoroscopy-assisted and 90.5 per cent for robot-assisted placement, with the thoracic spine consistently the least accurate segment. Evidence for briefing, debriefing and team training is directionally positive but heterogeneous, and a cluster-randomized trial of aviation-based team training added to an existing checklist programmer found improvement in both arms without a between-arm difference.

**Conclusions:** The aviation flight-planning paradigm offers a coherent, operationally useful organizing framework for surgical preoperative optimization. Its value is greatest where it is applied at the level of the planning process rather than imported as specific artefacts or devices. Three constraints qualify the transfer: the irreducible biological unpredictability of the patient, the heterogeneity of the surgical team relative to a flight crew, and the absence in healthcare of a centralized, mandatory accident-investigation infrastructure. Adoption of a routine, non-punitive, multidisciplinary post-procedure debriefing model is the single most defensible and least costly extension of the analogy, and remains the least consistently implemented.


**Keywords:** Preoperative planning; patient safety; human factors; crew resource management; surgical checklists; clinical debriefing; cognitive load; high reliability organizations; aviation safety; surgical decision-making; pedicle screw accuracy; spinal instrumentation; intraoperative adverse events; quality improvement

## Introduction

Surgery and aviation are the two industrial-scale human activities in which a well-trained professional acts directly on a physical system whose failure can kill people, in real time, without the possibility of pausing to consult a colleague. The comparison between them is not a rhetorical device invented by health services researchers; it is a structural observation that has been made repeatedly, and productively, for more than two decades [1,2]. The translation of aviation-derived paradigms into clinical practice has been associated with measurable reductions in adverse events, most visibly through the World Health Organization Surgical Safety Checklist [3,4]. That translation, however, has been selective. Attention has concentrated overwhelmingly on the intraoperative phase: on the checklist, on the time-out, on team communication, on crisis resource management training. The pre-procedural phase-the interval in which a surgeon gathers information, weighs competing risks, allocates resources, selects an approach and commits to a strategy-has attracted far less comparative scrutiny, despite being structurally the closest analogue to flight planning, and arguably the phase in which the greatest volume of latent error is generated.

The scale of the problem justifies the effort. An estimated 313 million surgical procedures are performed worldwide each year, and at least 4.2 million people die within thirty days of an operation, roughly half of those deaths occurring in low- and middle-income countries [5,6]. In high-income settings, adverse events occur in approximately 14.4 per cent of surgical patients, and systematic review evidence indicates that between 38 and 52 per cent of surgical adverse events are potentially preventable [7,8]. The WHO estimates that up to one in four patients is harmed in hospital care in low- and middle-income countries, with 134 million adverse events annually contributing to around 2.6 million deaths [9]. These figures are no longer controversial. What remains contested is the causal structure: whether the residual burden of preventable surgical harm is best addressed by more rigorous execution of intraoperative processes, or by better design of the decision-making that precedes the incision.

This article argues for the latter. We propose that surgical preoperative optimization and aviation flight planning are homologous processes that have been independently optimized within two different industries, and that the differences between them are as instructive as the similarities. We decompose both workflows into seven functionally comparable domains, examine each for structural correspondence and for the strength of available outcome evidence, and then identify where the analogy breaks down. Throughout, spinal instrumentation is used as the principal worked example, because it is the surgical domain in which the mapping between pre-procedural plan and intraoperative execution is most literal and most precisely measurable: a planned trajectory in three-dimensional space, an instrumented deviation from it, and an outcome-neurological, mechanical or vascular-that follows directly from the magnitude and direction of that deviation. Finally, we argue that the most consequential and least implemented element of the aviation model is not the checklist but the mandatory, non-punitive, post-mission debriefing, and that its systematic adoption represents the next frontier in surgical quality improvement. In doing so we take seriously the substantial critical literature that has questioned how far the aviation metaphor can

be pushed in medicine, and we attempt to be precise about which parts of it survive scrutiny.

## Materials and Methods

This study is a narrative comparative review with a structured search strategy and pre-specified analytical domains. It is not a systematic review and does not claim the methodological guarantees of one; the objective was conceptual synthesis rather than the pooling of effect estimates, and the comparison of two industries with profoundly different evidence cultures does not admit a single unified quality appraisal instrument. Where quantitative findings are cited, they are reported as published, with the design and limitations of the source stated. A structured search of PubMed/MEDLINE, Scopus, Web of Science and the Cochrane Library was conducted for the period 1 January 1990 to 15 October 2026, with no language restriction beyond the requirement that an English-language abstract or full text be retrievable.

The search combined four concept blocks. The first addressed aviation operations and human factors, using terms including aeronautical decision making, DECIDE model, flight planning, fuel reserve, alternate aerodrome, NOTAM, weather briefing, quick reference handbook, crew resource management, and line-oriented flight training. The second addressed surgical and perioperative safety broadly, using terms including preoperative planning, surgical safety checklist, surgical debriefing, morbidity and mortality conference, crew resource management in surgery, non-technical skills, situational awareness, intraoperative adverse event, and failure to rescue. The third addressed spinal instrumentation specifically, using terms including pedicle screw, pedicle screw accuracy, pedicle screw misplacement, Gertzbein-Robbins classification, medial breach, lateral breach, triggered electromyography, intraoperative navigation, robot-assisted spine surgery, three-dimensional printed drill guide, and learning curve. The fourth addressed cognitive and organizational mechanisms, using terms including cognitive load, working memory, fixation error, plan continuation bias, automation complacency, shared mental model, high reliability organization, psychological safety, and speaking up.

Because aviation safety is documented substantially outside the peer-reviewed biomedical literature, the search was supplemented by primary regulatory and industry sources, including the Federal Aviation Administration Pilot's Handbook of Aeronautical Knowledge and Aviation Weather Handbook, the Skyboard aviation safety reference maintained by EUROCONTROL, and accident-analytic publications from the Embry-Riddle Aeronautical University scholarly repository [10–14]. Patient safety policy documentation was drawn from the WHO Global Patient Safety Action Plan 2021–2030 and the WHO Safe Surgery programmer pages [9,15]. Technical literature on computational surgical planning and on instrumentation accuracy was retrieved through the same databases.

Records were screened by title and abstract against four pre-specified inclusion criteria: the record addressed at least one aviation safety or flight-planning construct; the record addressed at least one surgical or perioperative construct; the record permitted a substantive inference about transferability, mechanism or outcome; and the record was a peer-reviewed study, systematic review, clinical guideline, regulatory document or authoritative industry reference. Conference abstracts without retrievable full text, opinion pieces without substantive argument, and commercial promotional material were excluded. Grey literature was excluded except where it represented the primary source documentation of a regulatory standard, in which case it was treated as a primary document rather than as evidence of effect.

One hundred and eight records were screened, and 106 sources were retained in the final synthesis. Because the review spans two evidence cultures with incompatible hierarchies of evidence, sources were categorized into three tiers for interpretive purposes rather than scored on a single scale. Tier 1 comprised randomized and cluster-randomized trials, prospective cohorts with independent outcome evaluation, and prospective diagnostic accuracy studies. Tier 2 comprised systematic reviews and meta-analyses, national audits and registry analyses, and large retrospective cohorts. Tier 3 comprised regulatory standards, industry guidance, narrative reviews of mechanism, and theoretical or conceptual contributions. No meta-analysis was performed and no formal risk-of-bias instrument was applied; the narrative reports the design of each cited study so that the reader can weigh it directly.

The comparative analytic framework was constructed iteratively. Aviation flight planning was decomposed from the FAA Pilot's Handbook of Aeronautical Knowledge, the Aviation Weather Handbook, and the Skyboard fuel-planning definitions into its functional stages: acquisition of aircraft status and of route and aerodrome information; weight and balance computation; fuel requirement computation including reserves and alternate fuel; route selection with continuous in-route weather evaluation; alternate aerodrome selection; the progressive narrowing of decision options as environmental risk increases; rehearsal of emergency procedures through memory items and quick reference handbooks; and pre-departure verification against standard checklists [10,11,13]. These stages were then interrogated against the surgical literature to identify the closest functional counterpart in preoperative optimization, whether or not the surgical literature describes the practice in aviation terms. Where a functional counterpart was identified, the domain was retained; where no credible counterpart existed, the absence was recorded as a finding rather than forced into the framework. Seven domains survived this process, and each is analyzed below for structural homology, mechanism, available outcome evidence, and transferability constraints.

For the worked example, spinal instrumentation was selected as the domain in which the plan–execution–outcome chain is most directly quantifiable. Three features determined the choice. First, the plan is explicit and geometric: a trajectory defined in three-dimensional space, with a measurable angular and linear deviation from the achieved instrument position. Second, deviation is graded by an internationally used ordinal instrument, the Gertzbein-Robbins classification, which assigns breaches of 0–2 mm, 2–4

mm and greater than 4 mm to ascending grades and permits pooling across studies. Third, the consequences of specific breach directions are anatomically predictable and clinically documented, which allows a directional error to be traced to a specific class of harm rather than to a generic complication. This replaces the laparoscopic conversion exemplar used in earlier formulations of this framework, which illustrated the same principle but did not permit the same precision of measurement.

Two analytical conventions are applied throughout. First, structural homology is distinguished from demonstrated effect: establishing that two processes are functionally analogous does not establish that importing the practice from one to the other improves outcomes, and we mark the distinction explicitly in each domain. Second, the analysis treats the aviation analogy as a source of hypotheses about system design rather than as a validated intervention, and the burden of proof is placed on the proposed transfer rather than on those who decline it.

## Results: Domain-by-Domain Comparison

The comparative decomposition yielded seven functional domains in which aviation flight planning and surgical preoperative optimization perform a structurally corresponding task Table 1, supplied as a separate PDF, summarizes the mapping together with the strength of the supporting evidence and the principal transferability constraint identified for each domain. The spinal instrumentation evidence base is reported in Table 2 and is carried through the discussion as the principal worked example.

The two tables below accompany the article "Drawing the Flight Path of Surgery: A Comparative Analysis of viation Flight Planning and Surgical Preoperative Optimisation". Table 1 maps the seven functional domains of aviation flight planning onto their surgical counterparts and grades the strength of the supporting evidence. Table 2 reports the quantitative accuracy of pedicle screw placement by technique, with the directional consequence of breach, and supports the spinal instrumentation worked example carried through the planning domains of the discussion (Table 1).

- **Key to the Evidence Strength Column**

Strong denotes a consistent effect across randomized, prospective or multi-cohort evidence. Moderate to strong denotes a consistent direction of effect with heterogeneous magnitude. Moderate denotes supportive but procedure-specific evidence. Directional denotes a plausible mechanism with limited outcome evidence. Grading follows the three evidence tiers described in Materials and Methods (section 2) **(Table 2)**.

- **Directional Consequence of Breach**

**Table 1:** Structural correspondence between aviation flight planning and surgical preoperative optimization.

| Aviation Domain | Surgical Counterpart | Evidence strength | Principal Constraint |
|---|---|---|---|
| Information gathering: aircraft status, NOTAMs, route limits | Diagnostic imaging, CT/MRI, 3D reconstruction, patient-specific virtual planning | Moderate to strong | Anatomy is not standardized; value of added imaging is procedure-specific |
| Weight, balance and fuel reserve computation | Instrumentation, implant sizing, blood ordering and reserve | Moderate to strong | Over-ordering wastes a shared resource unlike surplus fuel |
| Route and weather planning with alternate aerodrome selection | Strategic routing with pre-briefed conversion and bail-out trigger points | Moderate | Plan continuation bias is under-measured in surgery |
| The decision funnel and continuous ADM loop | Contraction of intraoperative decision latitude; curvature of the error–harm relationship | Moderate | No validated real-time instrument for surgical decision latitude |
| Emergency procedures, memory items, quick reference handbook | Cognitive aids and algorithms for major complications | Directional | Outcome transfer less well established than training effect |
| Pre-departure checklist verification | WHO Surgical Safety Checklist (Sign In, Time Out, Sign Out) | Strong | Communication items are the least reliably performed |
| Mandatory non-punitive post-mission debriefing | Routine post-procedure surgical debriefing | Directional | Schedule pressure; hierarchy limits speaking up |

**Table 2:** Accuracy of pedicle screw placement by technique, with the directional consequence of breach.

| Technique | Pooled fully-Interpedicular or Safe-Zone Accuracy | Source and Evidence Tier |
|---|---|---|
| Intraoperative CT navigation | 95.5% | Systematic review and meta-analysis, 78 studies, 51,161 screws [95]; Tier 2 |
| Freehand | 93.1% | As above [95]; Tier 2 |
| Fluoroscopy-assisted | 91.5% | As above [95]; Tier 2 |
| Robot-assisted | 90.5% | As above [95]; confirmatory umbrella review [96]; Tier 2 |
| Augmented reality navigation | Unacceptable breaches (GRS grade C–E): 0.7% vs 3.6% freehand (p = 0.013) | Comparative cohort [97]; Tier 2 |
| Patient-specific 3D-printed guides | Superior to freehand in spinal deformity; reduced total radiation dose | Prospective randomized clinical trial, Level I [98]; Tier 1 |
| Patient-specific 3D-printed guides (cadaveric) | Entry-point accuracy 0.76 ± 0.52 mm; angular accuracy 3.22 ± 2.34° | Cadaveric accuracy study [103]; Tier 2 |

Freehand breaches concentrated at L5 and were predominantly medial (n = 22)-that is, toward the spinal canal and the traversing nerve root. Augmented reality breaches concentrated at L4 and were predominantly lateral, while the overall rate of clinically unacceptable breach was significantly lower than freehand [97]. Medial breaches of less than 2 mm in the lumbar spine produced no neurological deficit in any study identified, across thirteen studies. Medial breaches of 2–4 mm increased the risk of neurological deficit, with one such case developing radiculopathy requiring surgical revision. Innon-lumbar segments, medial breach carries a materially higher risk of spinal cord injury [99,100]. Risk of vertebral artery injury in the sub axial cervical spine is governed by a measurable safety margin-the combined distance of the vertebral artery from the lateral and medial walls of the transverse foramen-and the decision to revise should rest on the patient's clinical condition and breach status, rather than on screw position alone [115]. Verification caveat: triggered electromyography has very high specificity but only fair sensitivity for detecting mispositioned screws, with a reported sensitivity of 43.4 per cent at a threshold below 5 mA. It functions as a warning indicator rather than as a navigation display, and should be briefed and interpreted accordingly [108,112]. Source: synthesis of the literature reviewed in the accompanying article. Breach grading follows the Gertzbein-Robbins classification (grade 0, screw within pedicle; grade 1, breach <= 2 mm; grade 2, breach > 2–4 mm; grade 3, breach > 4 mm). Reference numbers correspond to the reference list of the article.

Two features of the evidence base are worth stating before the domains are examined individually. The first is that the spinal instrumentation literature is quantitative in a way that most surgical planning literature is not: accuracy is expressed as a proportion of acceptable screws within a large denominator, permitting direct comparison between techniques. Pooled analysis across 78 studies, 7,858 patients and 51,161 screws reported fully-interpedicular accuracy of 95.5 per cent for intraoperative CT navigation, 93.1 per cent for freehand technique, 91.5 per cent for fluoroscopy-assisted placement and 90.5 per cent for robot-assisted placement, with thoracic screws associated with lower accuracy across techniques and robot assistance showing fewer thoracic breaches than freehand or fluoroscopy [95]. The second is that the relationship between the magnitude of deviation and the probability of harm is continuous rather than binary, and is stratified by both direction and vertebral level [99,100]. These two features together make spinal instrumentation the most rigorous available test case for the planning hypothesis advanced in this article.

## Discussion

### Information Gathering: The Flight Plan and the Diagnostic Reconstruction

A pilot's preparation begins with the assembly of an operational picture that is deliberately broader than the flight itself. Aircraft status and airworthiness, airman certification and currency, aerodrome conditions and notices to airmen, and airspace and route limitations are all integrated before any performance computation is attempted [10,11]. The distinctive feature of this process is not the volume of data but its standardization: the categories are fixed, the ordering is conventional, and the briefing is structured so that a lapse in one category is detectable by its omission from a fixed sequence. The system is designed so that the pilot does not have to remember what to look for in surgery the functional counterpart is the diagnostic and anatomical appraisal that precedes the incision, and in spinal instrumentation the correspondence is unusually literal because both professions are working from a three-dimensional spatial model of a structure they must traverse without error.

Computed tomography, magnetic resonance imaging and three-dimensional reconstruction act as the surgeon's spatial chart, and the evidence supporting them is now substantial. A systematic review of three-dimensional reconstruction in liver surgery encompassing 47 studies and 7,724 patients found that preoperative three-dimensional reconstruction was used for planning in 87.2 per cent of included studies and contributed to increased accuracy of resection [16,17]. Virtual surgical planning in orthognathic surgery has reached comparable maturity, with 125 of 133 assessable patients in one systematic review achieving a successful outcome against a 2 mm mean difference criterion [18]. Broader systematic review confirms that virtual, augmented and mixed reality occupy a genuine, if maturing, role in bridging conventional imaging and intraoperative execution [20]. In the spine, however, the information-gathering step has reached a degree of formalization that goes beyond imaging interpretation. Patient-specific three-dimensional virtual surgical planning translates the preoperative computed tomography into planned screw trajectories and then into a physical or digital guide that reproduces those trajectories in the operating theatre [101,102].

Cadaveric accuracy assessment of customized drill guides for pedicle and lateral mass screw insertion in the cervical and upper thoracic spine found an overall entry-point accuracy of 0.76 ± 0.52 mm and three-dimensional angular accuracy of 3.22 ± 2.34° [103]. A prospective randomized clinical trial with independent outcome evaluation-Level I evidence-compared freehand insertion with patient-specific three-dimensional printed guides in spinal deformity and found the guides more accurate, with reduced total radiation dose even when the low-dose preoperative planning computed tomography was taken into account [98]. In a clinical series of cervical pedicle screws placed with patient-specific guides, 107 of 115 screws (93.1 per cent) were completely within the peduncle, with mean sagittal and transverse angular deviations of 2.94° and 3.04° respectively [104]. The comparison sharpens a specific point about failure modes. The aviation literature is explicit that operating without adequate navigation instrumentation in instrument meteorological conditions produces spatial disorientation, and that the resulting accident rate is high enough to be a defining hazard of the environment [10].

The surgical equivalent is the surgeon who prepares a pedicle track from tactile feedback and a mental model rather than from a verified spatial representation. The consequence is quantifiable: pooled accuracy across 78 studies and 51,161 screws was 95.5 per cent for intraoperative CT navigation against 93.1 per cent for freehand and 91.5 per cent for fluoroscopy-assisted placement, and systematic review of prospective in vivo studies found the proportion of screws fully contained within the peduncle ranged from 69 to 94 per cent using freehand technique, from 28 to 85 per cent with fluoroscopy guidance, and from 89 to 100 per cent with CT navigation [95,105]. A meta-analysis of 3,442 patients and 43,305 screws reported average safe-zone accuracy of 91.4 per cent for freehand or fluoroscopy against 97.3 per cent for navigation [106]. The mechanism is consistent with the aviation account: it is not more data in the cockpit that prevents disorientation, it is reliable presentation of the spatial situation.

A transferability caution applies, and it is the same one that applies to the aviation case. Aviation information gathering is standardized across a global fleet because aircraft, procedures and reporting formats are standardized. Vertebral anatomy is not, and the value of additional imaging is segment-specific: the thoracic spine is consistently the least accurate segment across techniques, and robot assistance has been shown to reduce thoracic breaches relative to freehand and fluoroscopy [95]. The lesson to import is the discipline of a fixed and complete information set for each defined procedure, with explicit criteria for adequacy, rather than the adoption of any particular modality or device.

## Resource and Logistics Balancing: Weight, Fuel and the Surgical Instrumentalism

Aircraft performance is governed by weight and balance limitations, and the flight plan computes a fuel figure composed of taxi fuel, trip fuel, contingency fuel, alternate fuel and final reserve fuel [10,13]. The consequences of getting these numbers wrong are asymmetric and unforgiving: excess weight degrades climb gradient, while insufficient fuel produces starvation at the least convenient point of the flight. Final reserve fuel is defined as the fuel required to fly for thirty minutes above the alternate aerodrome, and the prohibition on planning a flight that infringes it is explicit [13]. The reserve exists because the plan is expected to be wrong.

The surgical counterpart is the preparation of the operative armamentarium: specialized instrumentation, implant sizing across the anticipated range, backup trays, consumable hemostatic agents, intraoperative imaging capability and blood products. The analogue of fuel reserve in this domain is blood availability and patient blood management, structured around three pillars: optimizing the patient's own red cell mass, minimizing blood loss, and optimizing physiological tolerance of anemia [23,24]. Preoperative optimization of hemoglobin and iron stores, and the deliberate timing of elective surgery to allow that optimization to take effect, are explicit recommendations in national guideline frameworks [25]. The maximum surgical blood ordering schedule is a procedure-specific instrument designed to order sufficient blood for 85 to 90 per cent of patients undergoing a given operation, and its implementation reduces cross-matched units, waste and cost while preserving safety [26,27]; analysis across 53,526 patients and 1,632 procedures estimated that eliminating unnecessary orders would free annual hospital charges in the order of $211,448 at a single institution [28].

In spinal instrumentation the reserve calculation has additional structure, because the instrumentation itself is the reserve. The requirement is not only that the planned implant be available across a range of diameters and lengths, but that the means of executing a trajectory change be present before the first screw is placed. The literature on revision thresholds is explicit that apparent malposition is commonly corrected during the operation, and that doing so extends operative duration, increases radiation time and tissue trauma, increases blood loss, and reduces pull-

out strength of the revised screw [107]. This is an argument for pre-briefed contingency capacity rather than for improvisation: if the decision to revise is to be taken expeditiously, the equipment, the neuromonitoring and the team's mutual understanding of the threshold must all be established before the relevant screw is placed.

Where the analogy diverges is in the cost of misjudging the other direction. Over-ordering blood wastes a scarce shared resource and, in low- and middle-income settings, can deprive another patient; excess fuel principally costs money. And unlike fuel, an implant removed after malposition cannot be returned to the reserve-the revision consumes it. This asymmetry argues for centralized, procedure-level protocols rather than individual surgeon discretion-precisely the structure that aviation has adopted and surgery largely has not.

## Strategic Routing and Environmental Risk: The Flight Path and the Screw Trajectory

Aviation flight planning mandates the selection of alternate destination aerodromes and the continuous evaluation of in-route weather [10-12]. The specification of an alternate is not a contingency in the colloquial sense; it is a mandatory component of the plan, briefed before departure, with its own approach minima, fuel availability and ground conditions assessed in advance [12]. The FAA Aviation Weather Handbook is explicit that a weather briefing is advisory in nature and that the final decision as to whether a flight can be conducted safely rests solely with the pilot, and that the go or no-go decision should be reached by weighing current and forecast conditions rather than by seeking an authoritative instruction [11]. The pilot is thus the responsible decision-maker, but the decision must be made against a structure that has already forced the identification of escape routes.

Spinal instrumentation is the clearest surgical instantiation of this domain, because the route and the structure being navigated are the same object. The planned trajectory is the flight path; the pedicle is the corridor; and the medially deviated screw is the aircraft that has drifted into terrain. The pre-briefed alternate takes a specific and measurable form: the threshold for abandoning a trajectory. The recognized proxies for that threshold are the triggered electromyography response and the intraoperative imaging check. Systematic review and meta-analysis of triggered electromyography found that a current threshold of 10–12 mA with a pulse duration of 300 μsec provided the most accurate testing parameters for detecting misplaced screws, and concluded that the technique has very high specificity but only fair sensitivity for detecting mispositioned pedicle screws [108]. Subsequent work has refined the interpretation: thresholds of 6 mA or below are considered indicative of breach, 6–8 mA a possible breach, above 8 mA generally safe, and above 12 mA contact with neural structure is rare [109]. A retrospective analysis of 8,584 pedicle screw stimulations found that screw stimulation at or below 8 mA served as an accurate rule-in test for postoperative neurological deficit, warranting intraoperative re-evaluation or replacement of the screw [110]. Earlier large-series work on 2,450 consecutively placed screws reached a compatible conclusion with a stricter cut-off: a positive response below 5.0 mA was highly specific for a medial breach but poorly sensitive, while responses at higher thresholds were associated with a high rate of false positives [111].

The critical feature for the aviation analogy is what triggered electromyography cannot do. It's very high specificity and only fair sensitivity means that a reassuring threshold does not establish correct placement; it establishes only the absence of a detected breach. In the language of the cockpit, the instrument is a warning light, not a navigation display. Systematic review of intraoperative neuromonitoring across 9,957 cases has accordingly questioned its standalone utility for lumbar pedicle screw placement, and reported that a threshold below 5 mA carried a sensitivity of only 43.4 per cent for detecting pedicle wall breaches [112]. The operational consequence is that the decision to continue a trajectory must rest on a pre-briefed combination of evidence -imaging check, stimulation threshold, tactile feedback and the surgeon's appraisal of bone quality- established before the case begins, rather than on the interpretation of a single intraoperative reading under time pressure.

The mechanism that both industries are defending against is a well-documented bias. Aviation calls it plan continuation error, or get-there-itis; accident analysis in general aviation has found that among environment-related fatal accidents, the decision to depart into forecast adverse weather was the single most prevalent subcategory, accounting for 56 per cent of that category in one analysis [14]. In spinal instrumentation the equivalent phenomenon is the decision to accept a track that tactile feedback or a borderline stimulation threshold has already flagged as doubtful, because the pedicle has been prepared, because extraction requires declaring that something has gone wrong, or because the instrumentation set is already mounted. The structural remedy in both industries is the same: the threshold must be chosen and briefed before the trajectory is started, when the decision is cheap and unemotional, because it will not be chosen well at the moment it becomes necessary.

## The Cognitive Continuum: The Decision Funnel and the Curvature of Harm

The decision funnel describes a dynamic cognitive model in which the available option set contracts as time passes and environmental hazard increases. In flight, a pilot approaching a convective system begins with a wide range of lateral and vertical deviations; as the aircraft closes on the hazard, the window for simple avoidance narrows until the pilot is confined to high-risk emergency man oeuvres. This is not merely a metaphor about shrinking options. It is a statement about the cost of late recognition: an identical hazard, recognized earlier, is cheap to avoid; recognized later, it is expensive to survive.

The aeronautical decision-making literature has responded with a formal continuous loop. The DECIDE model directs the pilot to Detect the change, Estimate the correction required, choose a desirable outcome, identify actions, Do the necessary actions, and evaluate the effect of the action, and it is explicitly a loop rather

a one-time preflight checklist [32,33]. The specific failure the loop is designed to catch is that the pilot may be running an outdated version of the problem: undetected, the deviation from plan proceeds uncorrected until the option set has already collapsed.

Spinal instrumentation offers an unusually precise instantiation of the funnel, because the relationship between the magnitude of deviation and the probability of harm is documented as a continuous function with a distinct threshold rather than as a binary event. A systematic review of the risk of neurological deficit in medially breached pedicle screws assessed by computed tomography found that none of the sub-2 mm medial breaches in lumbar vertebrae caused any neurological deficit across thirteen studies, while a medial breach of 2–4 mm increased the risk of neurological deficit, with one such case developing radiculopathy requiring surgical revision [99]. A separate review noted an incidence of nerve root injury risk at the 6 mm computed tomography deviation level [113]. In non-lumbar segments, medial breach carries a materially higher risk of spinal cord injury [99,100]. Directional analysis of a comparative cohort reinforces the point: freehand breaches concentrated at L5 and were predominantly medial, whereas augmented reality breaches concentrated at L4 and were predominantly lateral [97].

This is the funnel expressed anatomically. The first millimeter of medial deviation is tolerated, and correctly so; the second is where the risk begins to climb steeply. A system that reacts only when the screw is already in the canal has waited until the terminal part of the funnel, where the available corrective actions are extraction, revision with alternative instrumentation and management of an established deficit. A system that reacts in the first millimeter has the option of a trajectory change at negligible cost. The empirical foundation for the cognitive account is also now available: systematic review of objective cognitive workload assessment in surgery confirms that cognitive overload impairs clinical decision-making and performance, that up to one third of surgical errors can be attributed to fatigue and excessive workload, and that burnout is significantly associated with surgical errors [37,38]. Prospective assessment of surgeons working post-call has found significantly poorer non-technical skills ratings for situational awareness, decision-making, and communication and teamwork, even where discrete error counts did not rise [39,40]. Systematic review of cognitive support systems in the operating room indicates that such technologies reduce measured mental workload, but that their clinical evidence base remains immature and ergonomic limitations have hindered adoption [41,42].

The practical implication of the funnel is that the decisive intervention point is early, and that its target is recognition rather than rescue. The curvature of the harm function in spinal instrumentation makes this unusually concrete: the intervention sequence-tap, image, stimulate, rod-is itself a series of decision gates, and each gate that closes removes options. This is a substantially different proposition from improving performance in the terminal part of the funnel, and it explains why crisis management training, valuable as it is, cannot substitute for early recognition.

## Crisis Management and Contained Contingencies: Flight Emergencies and Iatrogenic Injury

Aviators mitigate in-flight crises by training intensively for worst-case scenarios and by relying on memory items and quick reference handbooks that prescribe an ordered sequence of actions under time pressure [10,43]. The design logic rests on two premises: that severe emergencies degrade the capacity for novel reasoning, and that the correct sequence of actions is known in advance and therefore should not have to be reconstructed in the moment.

Medical culture has historically treated complications as unpredictable anomalies rather than as expected structural events. The evidence no longer supports that framing, and in spinal instrumentation the taxonomy of harm is unusually well defined. A mispositioned screw may cause dural tear, nerve root irritation with radicular pain, motor or sensory dysfunction, vascular injury, cerebrospinal fluid leak, visceral injury, pedicle fracture and implant loosening [113,114]. The vascular dimension is segment-specific and has been experimentally characterized: cadaveric work on sub axial cervical pedicle screw malposition has validated a "safe zone" hypothesis defined by the combined distance of the vertebral artery from the lateral and medial walls of the transverse foramen, with a significant correlation between the safety margin and the intraformational area occupied by screws in causing vertebral artery damage [115]. That study concluded that the decision for revision surgery should be based on the patient's clinical condition and breach status-whether the vessel is displaced or injured-rather than on screw position alone [115]. This is a direct surgical analogue of the aviation principle that a warning indication must be interpreted in the context of the actual system state, not acted on reflexively.

Delayed presentations deserve particular emphasis because they are the surgical equivalent of the latent failure. One report documents a neurological deficit appearing six years after adolescent idiopathic scoliosis correction, caused by a medially misplaced thoracic pedicle screw [116]. A systematic review of mispositioned lumbosacral screws reported misplacement in the lumbosacral canal exceeding 8.7 per cent, with revision undertaken early in patients showing frank neurological compromise and conservative management attempted in those with milder findings [113]. The concept of failure to rescue formalizes the same insight at the hospital level: death after a complication is not simply the gravity of the complication but a property of the system's capacity to recognize and treat it, and between-hospital variation in failure to rescue after major surgery is substantial [49,50]. Systematic review of intraoperative adverse event grading has nonetheless documented a persistent lack of standardization, which limits comparison and impedes the very learning that aviation derives from classification [47,48].

The aviation-derived recommendation is therefore that major procedural complications should be treated as pre-rehearsed contingencies rather than as failures. In spinal surgery this means procedure-specific cognitive aids for defined catastrophic events-major vascular injury during anterior approaches, air embolism,

spinal cord monitoring signal loss, cerebrospinal fluid fistula- held in the operating theatre and rehearsed in situ. Simulation-based crisis resource management training has an established evidence base: deliberate practice with focused feedback is a requirement for non-technical skill acquisition, briefings that build shared mental models improve team performance during subsequent crises, and interprofessional crisis simulation has been shown to improve team coordination measured on validated instruments [51-53]. Evidence for transfer to patient outcomes is more limited. A national stepped-wedge cluster-randomized trial of in-situ simulation-based team training in New Zealand public hospitals could not attribute improved patient outcomes to the intervention, plausibly reflecting high baseline teamwork and surgical performance, dilution of effect over the intervention periods, and confounding by the COVID-19 pandemic [54].

The honest reading is that the mechanism is sound, the training technology is mature, and the outcome evidence is thinner than the enthusiasm for it. Rehearsed algorithms plausibly shorten response latency and reduce omission error during crises, which is what they are designed to do; whether that translates into reduced mortality at scale remains less well established than the aviation analogy implies.

## Operational Safeguards: Checklists and the Verification of Trajectory

Checklists in aviation are not memory tests. They are structural verification tools designed to catch human error before it translates into an accident, and the human factors literature on their design is explicit about the conditions under which they work: checklist content must be confined to items that are important, that are not already reliably performed, and that are not readily verifiable by other means [43]. The recognition that approximately 80 per cent of aviation accidents are related to human factors, with the largest concentrations in the landing and takeoff phases, explains why verification is concentrated at those phase boundaries [32].

The WHO Surgical Safety Checklist adapted this concept to surgery, dividing the operation into Sign In, Time Out and Sign Out, and its introduction was associated with reductions in both death and complications in a diverse group of hospitals [3]. In the original prospective study of 3,733 consecutive non-cardiac surgical patients, the rate of death fell from 1.5 per cent to 0.8 per cent and inpatient complications from 11.0 per cent to 7.0 per cent [3]. Subsequent cohort analysis in a single tertiary center covering 25,513 adult patients found a comparable reduction in in-hospital mortality [55], and a five-year implementation analysis demonstrated statistically significant reductions in mortality and length of admission [56]. Compliance, however, remains the binding constraint. Clinical audit of checklist use has repeatedly found that the items which promote communication between team members are the most poorly performed, and that the items most easily reduced to a tick-confirming identity and site-are the ones most reliably completed [57].

Spinal instrumentation adds a verification layer that has no direct aviation equivalent in the pre-departure checklist, but does have one in the post-departure cross-check: the intraoperative imaging check of screw position. Here the literature is explicit about the limits of the instrument. Triggered electromyography has very high specificity but only fair sensitivity [108], and a large-series review questioned its standalone utility for lumbar screw placement [112]. This is a genuine and constructive complication for the analogy: in aviation, a verification step that fails to detect a dangerous condition reliably would be redesigned or supplemented. The surgical answer has been to supplement the electrical check with an imaging check, which is precisely the redundancy principle underlying aviation's use of independent, non-overlapping verification methods.

The broader literature on whether aviation-derived training adds to the checklist alone is more sobering. A prospective parallel-group cluster-randomized trial across 31 hospitals, all of which had already implemented a surgical safety checklist a median of 19 months earlier, randomized operating room teams to an aviation-based crew resource management training programmer. The risk of major adverse events fell from 8.8 to 5.5 per cent in the intervention hospitals, but a comparable improvement occurred in the control arm, and the trial found no difference between arms [58,59]. An umbrella review of crew resource management in healthcare has concluded that the evidence base is heterogeneous and that the question of whether aviation-style team training changes outcomes is not settled [60,61]. Systematic review of operating room teamwork nonetheless supports the association between structured team training, briefing and debriefing, team familiarity and a positive safety climate on one hand and lower postoperative morbidity and fewer adverse events on the other, though effects on mortality are inconsistent [62].

Three conclusions follow. First, the checklist is a validated instrument and its adoption was the correct first move. Second, adding aviation-derived training on top of an established checklist programmer does not automatically produce additional outcome benefit. Third, the mechanism by which checklists work is communicative rather than procedural, and the communicative items are exactly the ones that get dropped when implementation is superficial. The WHO itself notes that the checklist is not intended to be comprehensive and encourages local modification [63], a point frequently overlooked when a service imports a fixed artefact.

## Closing the Loop: Debriefing, Hierarchy and the Learning Curve

The most significant cultural divergence between modern surgery and aviation lies in post-procedure analysis. Elite military aviation mandates a meticulous, non-punitive debriefing after every mission regardless of outcome, dissecting human factors, communication gaps and micro-errors to refine collective performance continuously. Surgical debriefing, by contrast, is frequently perfunctory or omitted entirely under schedule pressure unless a catastrophic outcome occurs, and the postoperative phase has been described in the literature as the forgotten phase of the surgical safety checklist [64,65]. This asymmetry matters because the debrief is the mechanism by which the other six domains are continuously recalibrated. Without it, a checklist becomes a ritual

a one-time preflight checklist [32,33]. The specific failure the loop is designed to catch is that the pilot may be running an outdated version of the problem: undetected, the deviation from plan proceeds uncorrected until the option set has already collapsed.

Spinal instrumentation offers an unusually precise instantiation of the funnel, because the relationship between the magnitude of deviation and the probability of harm is documented as a continuous function with a distinct threshold rather than as a binary event. A systematic review of the risk of neurological deficit in medially breached pedicle screws assessed by computed tomography found that none of the sub-2 mm medial breaches in lumbar vertebrae caused any neurological deficit across thirteen studies, while a medial breach of 2–4 mm increased the risk of neurological deficit, with one such case developing radiculopathy requiring surgical revision [99]. A separate review noted an incidence of nerve root injury risk at the 6 mm computed tomography deviation level [113]. In non-lumbar segments, medial breach carries a materially higher risk of spinal cord injury [99,100]. Directional analysis of a comparative cohort reinforces the point: freehand breaches concentrated at L5 and were predominantly medial, whereas augmented reality breaches concentrated at L4 and were predominantly lateral [97].

This is the funnel expressed anatomically. The first millimeter of medial deviation is tolerated, and correctly so; the second is where the risk begins to climb steeply. A system that reacts only when the screw is already in the canal has waited until the terminal part of the funnel, where the available corrective actions are extraction, revision with alternative instrumentation and management of an established deficit. A system that reacts in the first millimeter has the option of a trajectory change at negligible cost. The empirical foundation for the cognitive account is also now available: systematic review of objective cognitive workload assessment in surgery confirms that cognitive overload impairs clinical decision-making and performance, that up to one third of surgical errors can be attributed to fatigue and excessive workload, and that burnout is significantly associated with surgical errors [37,38]. Prospective assessment of surgeons working post-call has found significantly poorer non-technical skills ratings for situational awareness, decision-making, and communication and teamwork, even where discrete error counts did not rise [39,40]. Systematic review of cognitive support systems in the operating room indicates that such technologies reduce measured mental workload, but that their clinical evidence base remains immature and ergonomic limitations have hindered adoption [41,42].

The practical implication of the funnel is that the decisive intervention point is early, and that its target is recognition rather than rescue. The curvature of the harm function in spinal instrumentation makes this unusually concrete: the intervention sequence-tap, image, stimulate, rod-is itself a series of decision gates, and each gate that closes removes options. This is a substantially different proposition from improving performance in the terminal part of the funnel, and it explains why crisis management training, valuable as it is, cannot substitute for early recognition.

## Crisis Management and Contained Contingencies: Flight Emergencies and Iatrogenic Injury

Aviators mitigate in-flight crises by training intensively for worst-case scenarios and by relying on memory items and quick reference handbooks that prescribe an ordered sequence of actions under time pressure [10,43]. The design logic rests on two premises: that severe emergencies degrade the capacity for novel reasoning, and that the correct sequence of actions is known in advance and therefore should not have to be reconstructed in the moment.

Medical culture has historically treated complications as unpredictable anomalies rather than as expected structural events. The evidence no longer supports that framing, and in spinal instrumentation the taxonomy of harm is unusually well defined. A mispositioned screw may cause dural tear, nerve root irritation with radicular pain, motor or sensory dysfunction, vascular injury, cerebrospinal fluid leak, visceral injury, pedicle fracture and implant loosening [113,114]. The vascular dimension is segment-specific and has been experimentally characterized: cadaveric work on sub axial cervical pedicle screw malposition has validated a "safe zone" hypothesis defined by the combined distance of the vertebral artery from the lateral and medial walls of the transverse foramen, with a significant correlation between the safety margin and the intraformational area occupied by screws in causing vertebral artery damage [115]. That study concluded that the decision for revision surgery should be based on the patient's clinical condition and breach status-whether the vessel is displaced or injured-rather than on screw position alone [115]. This is a direct surgical analogue of the aviation principle that a warning indication must be interpreted in the context of the actual system state, not acted on reflexively.

Delayed presentations deserve particular emphasis because they are the surgical equivalent of the latent failure. One report documents a neurological deficit appearing six years after adolescent idiopathic scoliosis correction, caused by a medially misplaced thoracic pedicle screw [116]. A systematic review of mispositioned lumbosacral screws reported misplacement in the lumbosacral canal exceeding 8.7 per cent, with revision undertaken early in patients showing frank neurological compromise and conservative management attempted in those with milder findings [113]. The concept of failure to rescue formalizes the same insight at the hospital level: death after a complication is not simply the gravity of the complication but a property of the system's capacity to recognize and treat it, and between-hospital variation in failure to rescue after major surgery is substantial [49,50]. Systematic review of intraoperative adverse event grading has nonetheless documented a persistent lack of standardization, which limits comparison and impedes the very learning that aviation derives from classification [47,48].

The aviation-derived recommendation is therefore that major procedural complications should be treated as pre-rehearsed contingencies rather than as failures. In spinal surgery this means procedure-specific cognitive aids for defined catastrophic events-major vascular injury during anterior approaches, air embolism,

spinal cord monitoring signal loss, cerebrospinal fluid fistula- held in the operating theatre and rehearsed in situ. Simulation-based crisis resource management training has an established evidence base: deliberate practice with focused feedback is a requirement for non-technical skill acquisition, briefings that build shared mental models improve team performance during subsequent crises, and interprofessional crisis simulation has been shown to improve team coordination measured on validated instruments [51-53]. Evidence for transfer to patient outcomes is more limited. A national stepped-wedge cluster-randomized trial of in-situ simulation-based team training in New Zealand public hospitals could not attribute improved patient outcomes to the intervention, plausibly reflecting high baseline teamwork and surgical performance, dilution of effect over the intervention periods, and confounding by the COVID-19 pandemic [54].

The honest reading is that the mechanism is sound, the training technology is mature, and the outcome evidence is thinner than the enthusiasm for it. Rehearsed algorithms plausibly shorten response latency and reduce omission error during crises, which is what they are designed to do; whether that translates into reduced mortality at scale remains less well established than the aviation analogy implies.

## Operational Safeguards: Checklists and the Verification of Trajectory

Checklists in aviation are not memory tests. They are structural verification tools designed to catch human error before it translates into an accident, and the human factors literature on their design is explicit about the conditions under which they work: checklist content must be confined to items that are important, that are not already reliably performed, and that are not readily verifiable by other means [43]. The recognition that approximately 80 per cent of aviation accidents are related to human factors, with the largest concentrations in the landing and takeoff phases, explains why verification is concentrated at those phase boundaries [32].

The WHO Surgical Safety Checklist adapted this concept to surgery, dividing the operation into Sign In, Time Out and Sign Out, and its introduction was associated with reductions in both death and complications in a diverse group of hospitals [3]. In the original prospective study of 3,733 consecutive non-cardiac surgical patients, the rate of death fell from 1.5 per cent to 0.8 per cent and inpatient complications from 11.0 per cent to 7.0 per cent [3]. Subsequent cohort analysis in a single tertiary center covering 25,513 adult patients found a comparable reduction in in-hospital mortality [55], and a five-year implementation analysis demonstrated statistically significant reductions in mortality and length of admission [56]. Compliance, however, remains the binding constraint. Clinical audit of checklist use has repeatedly found that the items which promote communication between team members are the most poorly performed, and that the items most easily reduced to a tick-confirming identity and site-are the ones most reliably completed [57].

Spinal instrumentation adds a verification layer that has no direct aviation equivalent in the pre-departure checklist, but does have one in the post-departure cross-check: the intraoperative imaging check of screw position. Here the literature is explicit about the limits of the instrument. Triggered electromyography has very high specificity but only fair sensitivity [108], and a large-series review questioned its standalone utility for lumbar screw placement [112]. This is a genuine and constructive complication for the analogy: in aviation, a verification step that fails to detect a dangerous condition reliably would be redesigned or supplemented. The surgical answer has been to supplement the electrical check with an imaging check, which is precisely the redundancy principle underlying aviation's use of independent, non-overlapping verification methods.

The broader literature on whether aviation-derived training adds to the checklist alone is more sobering. A prospective parallel-group cluster-randomized trial across 31 hospitals, all of which had already implemented a surgical safety checklist a median of 19 months earlier, randomized operating room teams to an aviation-based crew resource management training programmer. The risk of major adverse events fell from 8.8 to 5.5 per cent in the intervention hospitals, but a comparable improvement occurred in the control arm, and the trial found no difference between arms [58,59]. An umbrella review of crew resource management in healthcare has concluded that the evidence base is heterogeneous and that the question of whether aviation-style team training changes outcomes is not settled [60,61]. Systematic review of operating room teamwork nonetheless supports the association between structured team training, briefing and debriefing, team familiarity and a positive safety climate on one hand and lower postoperative morbidity and fewer adverse events on the other, though effects on mortality are inconsistent [62].

Three conclusions follow. First, the checklist is a validated instrument and its adoption was the correct first move. Second, adding aviation-derived training on top of an established checklist programmer does not automatically produce additional outcome benefit. Third, the mechanism by which checklists work is communicative rather than procedural, and the communicative items are exactly the ones that get dropped when implementation is superficial. The WHO itself notes that the checklist is not intended to be comprehensive and encourages local modification [63], a point frequently overlooked when a service imports a fixed artefact.

## Closing the Loop: Debriefing, Hierarchy and the Learning Curve

The most significant cultural divergence between modern surgery and aviation lies in post-procedure analysis. Elite military aviation mandates a meticulous, non-punitive debriefing after every mission regardless of outcome, dissecting human factors, communication gaps and micro-errors to refine collective performance continuously. Surgical debriefing, by contrast, is frequently perfunctory or omitted entirely under schedule pressure unless a catastrophic outcome occurs, and the postoperative phase has been described in the literature as the forgotten phase of the surgical safety checklist [64,65]. This asymmetry matters because the debrief is the mechanism by which the other six domains are continuously recalibrated. Without it, a checklist becomes a ritual

a) Adopt a fixed, procedure-specific preoperative information set with explicit adequacy criteria, and require documented personal review of the imaging and of the planned trajectories by the operating surgeon.

b) Implement procedure-level maximum blood ordering schedules and structured preoperative coordination between surgical, anesthetic and transfusion services, mirroring the fuel-planning meeting logic.

c) Define and document the verification package for each instrumented procedure before the case begins: which check is used at which level, what its known sensitivity and specificity are, and what result constitutes a stop. Triggered electromyography should be understood and briefed as a highly specific but only fairly sensitive warning indicator, not as a navigation display [108,112].

d) Require a documented, pre-briefed trajectory-change criterion for each instrumented level, expressed in anatomical terms and in millimeters, and communicate it to the whole team before the first screw is placed.

e) Introduce a mandatory pre-incision briefing of no more than two minutes covering the plan, the anticipated difficulty, the levels of greatest risk and the specific findings the team should be watching for.

f) Place procedure-specific cognitive aids for defined catastrophic events in the operating theatre and rehearse them in situ rather than only in the simulation center.

g) Institute a routine postoperative debrief of no more than three minutes, held before room turnover, structured around what went well, what deviated from plan, and what should change next time.

h) Feed debriefs outputs into a registry-linked morbidity and mortality process with explicit follow-up on actions, so that debrief data become quality improvement input rather than commentary.

i) Instrument the pathway, not the intention: measure briefing compliance, debrief completion rate, trajectory-criterion documentation, revision and conversion rates, and Gertzbein-Robbins-graded breaches against pre-planned trajectories, and audit them as process measures alongside outcome measures.

Two implementation principles are drawn from the evidence and are worth stating because they are commonly violated. First, the mechanism of a checklist is communicative, and compliance audits that count ticked boxes while omitting team introduction and shared-risk discussion measure the wrong thing; the communication items are both the most valuable and the most fragile [57,67]. Second, local adaptation is not a concession but a condition of effectiveness. The WHO explicitly invites modification of its checklist to fit local practice [63], and the cluster-randomized evidence suggests that successful implementation of aviation-based team training requires adaptation of its principles to the surgical milieu rather than faithful transplantation [58,59]. Transferring a foreign artefact wholesale is the most reliable way to produce the appearance of compliance without the behavior.

### Future Directions

Four research priorities emerge, the fourth of which is specific to the worked example. The first is instrumented process measurement in the planning domain: the aviation comparison is supported largely by structural argument, and the planning domain lacks an instrument analogous to Cassandra. Deploying Cassandra-graded adverse event capture alongside documented planning artefacts would allow the central hypothesis of this article to be tested rather than asserted [44-46]. The second is the economics and feasibility of debriefing at scale in resource-constrained systems, since the objections raised in high-income settings concern time and throughput, and these constraints bind harder where the perioperative mortality burden is concentrated [5,6,9]. The third is the refinement of objective cognitive workload measurement, including through pupillometry and heart-rate variability; real-time feedback on overload that enables timely intraoperative adjustment is directly analogous to the workload-management function of crew resource management, and would convert the decision funnel from a conceptual model into a monitorable operational state [37,94].

The fourth priority is specific to spinal instrumentation and is, in our view, the most consequential. The literature currently reports accuracy as an aggregate proportion of acceptable screws. What it does not report, and what the aviation analogy demands, is the conditional question: given a documented deviation at a defined magnitude and direction, what was the probability and severity of harm, and what corrective action was taken at which point in the case? Systematic review of the risk of neurological deficit in medially breached screws has begun this work by stratifying by breach magnitude and vertebral level [99], and the cadaveric safe-zone work on vertebral artery injury has begun it for the vascular dimension [115]. What is needed is a prospective instrument that records, for every instrumented screw, the planned trajectory, the achieved trajectory, the direction and magnitude of any deviation, the verification findings at the time of decision, and the action taken-so that the relationship between deviation geometry, timing of recognition and clinical outcome becomes empirically estimable rather than inferred. That instrument would be the surgical equivalent of the flight data recorder, and without it the planning hypothesis of this article remains, precisely, a hypothesis.

## Conclusions

Preoperative surgical planning and pre-flight aviation planning are homologous processes, and the homology is structural rather than metaphorical. Both domains require the integration of incomplete information into a committed plan, the allocation of finite resources with asymmetric penalties for underestimation, the pre-selection of escape routes before the moment of need, the preservation of decision latitude by early recognition of deviation, verification against a standardized sequence, and the systematic conversion of individual experience into organizational learning. Six of the seven domains examined display substantive structural correspondence, and spinal instrumentation provides the closest

available instantiation of the analogy because the plan, the deviation and the consequence are all directly measurable.

Spinal instrumentation also demonstrates the analogy's value at its most concrete. Pooled accuracy across 78 studies and 51,161 screws ranges from 90.5 per cent for robot-assisted placement to 95.5 per cent for intraoperative CT navigation, with freehand at 93.1 and fluoroscopy-assisted at 91.5 per cent, and the thoracic spine is consistently the least accurate region [95]. The relationship between error and harm is continuous rather than binary: sub-2 mm medial breaches in the lumbar spine produced no neurological deficit in any study identified, while 2–4 mm medial breaches increased the risk of deficit, and non-lumbar medial breach carries a materially higher risk of spinal cord injury [99,100]. Freehand breaches concentrate at L5 and are predominantly medial-toward the canal and the traversing root [97]. This is the decision funnel expressed in anatomy, and it establishes that the decisive intervention point is early, geometric and measurable, not late and heroic.

The strength of the case for transfer is not uniform across domains. In information gathering, resource planning and checklists, the analogy is largely descriptive of practices surgery has already validated, with checklist implementation associated with reductions in morbidity and mortality of roughly one third and a consistent direction of effect across cohorts and national audits [3,55,56]. In strategic routing, the decision funnel and crisis management, the analogy is generative: it reframes existing but sporadic surgical practices as systematic operational requirements. In routine post-procedure debriefing, the analogy is prescriptive and identifies the largest remaining gap. The evidence for the debriefs specific value in a skills-learning domain is unusually direct, because instrumentation accuracy improves measurably with accumulated experience-from 5.1 to 2.0 per cent malposition across successive patient cohorts in navigated lumbar placement, with significant improvement after approximately thirty screws [118,119]. That improvement represents knowledge acquired at the cost of the earlier patients, and a contemporaneous structured debrief is the only mechanism that would preserve it.

Three constraints must qualify the conclusion. The irreducible biological and geometric unpredictability of the patient-reinforced, in this domain, by a tolerance for error sometimes smaller than the resolution of the instruments used to plan and verify it-means that some residual deviation reflects neither planning nor execution failure [99,121]. The heterogeneity of the surgical team relative to a two-person flight crew means that interventions designed for a cockpit operate under assumptions that do not hold in an operating theatre, with the interpretation of neuromonitoring by a person other than the operator being a specific instance [89]. And the absence in healthcare of a mandatory, independent, technically authoritative investigation function means that importing the debrief without the apparatus that gives a debrief consequence reproduces form without function [90,91].

Surgery should therefore not become aviation, and the frequent invocation of aviation as a rhetorical standard of perfection obscures more than it clarifies. What surgery can take from aviation is a discipline of planning: a fixed and complete information set for each defined procedure, an explicit resource and reserve calculation, a pre-briefed escape route measured in millimeters and degrees, an instrumented awareness of the narrowing decision window, a rehearsed response to the predictable catastrophe, documented knowledge of what each verification instrument can and cannot detect, and-above all-the routine, unhurried, blame-free examination of what actually happened. The first seven are largely within the reach of individual teams. The eighth requires a change in what operating theatres treat as the end of a case, and it is the one on which the durability of all the others depends.

## Acknowledgment

None.

## Conflict of Interest

No conflict of Interest.

## References

1. Haynes AB, Weiser TG, Berry WR, Lipsitz SR, Breizat AH, et al. (2009) A surgical safety checklist to reduce morbidity and mortality in a global population. N Engl J Med 360(5): 491-499.
2. Catchpole K, Mishra A, Handa A, McCulloch P (2008) Teamwork and error in the operating room: analysis of skills and roles. Ann Surg 247(4): 699-706.
3. World Health Organization. Surgical Safety Checklist. Geneva: WHO; 2009.
4. World Health Organization. Safe surgery. Geneva: WHO.
5. Meara JG, Leather AJM, Hagander L, Alkire BC, Alonso N, et al. (2015) Global Surgery 2030: evidence and solutions for achieving health, welfare, and economic development. Lancet 386(9993): 569-624.
6. Nepogodiev D, Martin J, Biccard B, Makupe A, Bhangu A (2019) Global burden of postoperative death. Lancet 393(10170): 401-403.
7. Anderson O, Davis R, Hanna GB, Vincent CA (2013) Surgical adverse events: a systematic review. Am J Surg 206(2): PP 253-62.
8. Gawande AA, Thomas EJ, Zinner MJ, Brennan TA (1999) The incidence and nature of surgical adverse events in Colorado and Utah in 1992. Surgery 126(1): 66-75.
9. World Health Organization (2021) Global Patient Safety Action Plan 2021-2030: towards eliminating avoidable harm in health care. Geneva: WHO.
10. Federal Aviation Administration (2023) Pilot's Handbook of Aeronautical Knowledge (FAA-H-8083-25C). Oklahoma City: FAA.
11. Federal Aviation Administration (2026) Aviation Weather Handbook (FAA-H-8083-28B). Oklahoma City: FAA.
12. Federal Aviation Administration. The weather plan you never write down: understanding flight plans, flight following, and search-and-rescue services. FAA Cleared for Takeoff blog.
13. SKYbrary Aviation Safety. Fuel-flight planning definitions. EUROCONTROL.
14. Boyd DD, Scharf MT, Cross JL (2023) Deficient aeronautical decision-making contributions to fatal general aviation accidents. Embry-Riddle Aeronautical University Scholarly Commons.
15. World Health Organization. Global patient safety action plan 2021-2030. Geneva: WHO.

1. Lang H, Radtke A, Liu C, Frühauf NR, Peitgen HO, et al (2005) 3D liver reconstruction and computer-aided tumor analysis for preoperative planning in liver surgery. Br J Surg 92(11): 1405-1412.
2. Boedecker C, Huettl F, Saalfeld P, Paschold M, Kneist W, et al. (2024) 3-D reconstruction in liver surgery: a systematic review. HPB (Oxford) 26(3): PP 317-329.
3. Stokbro K, Aagaard E, Torkov P, Bell RB, Thygesen T (2014) Virtual planning in orthognathic surgery. Int J Oral Maxillofac Surg 43(8): 957-965.
4. Pellino G, Keller DS, Sorgato N, Ruiz MG, Modesti M, et al. (2024) The use of mixed reality in the preoperative planning of colorectal surgery: preliminary experience with a narrative review. Cir Esp 102(7): 393-399.
5. Sinha A, Sathianarayanan S, Kulkarni S, Patil S, Kumar A, et al. (2026) The digital evolution of surgical planning: a systematic review of immersive and interactive technologies. Front Surg 13:1297911.
6. National Blood Authority (Australia) (2024) Patient blood management. Canberra: NBA.
7. National Blood Authority (Australia) (2012) Patient Blood Management Guidelines: Module 2-Perioperative. Canberra: NBA.
8. Hofmann A, Farmer S, Towler SC (2012) Strategies to preempt and reduce the use of blood products: an overview of the benefits and risks of patient blood management. Transfus Med Hemother 25(1): 66-73.
9. Khanna P, Kumar A, Garg S, Bhartia VK (2011) Implementation of maximum surgical blood ordering schedule and an improvement in transfusion practices of surgeons subsequent to intervention. Indian J Hematol Blood Transfus 27(4): 223-227.
10. Al-Riyami AZ, Al-Nomani M, Al-Kindi S, Al-Riyami R, Al-Hinai A, et al (2025) A maximum surgical blood ordering schedule: does it add value? Vox Sang 120(5): 443-451.
11. Frank SM, Oleyar MJ, Ness PM, Tobian AAR (2014) Reducing unnecessary preoperative blood orders and costs by implementing an updated institution-specific maximum surgical blood order schedule and a remote electronic blood release system. Anesthesiology 121(3): 501-509.
12. Yazdi AP, Bigdeli M, Sadeghi M, Ahmadi A, Soltani S (2021) Strategies for improvement of blood consumption management in the operating rooms: experts' suggestions. Int J Gen Med 14: 5523-5530.
13. Federal Aviation Administration (2023) Aeronautical decision-making. In: Pilot's Handbook of Aeronautical Knowledge, Chapter 2. Oklahoma City: FAA.
14. Federal Aviation Administration. Aeronautical decision-making: DECIDE model excerpt. FAA Safety Team.
15. Flin R, Youngson G, Yule S (2007) How do surgeons make intraoperative decisions? Qual Saf Health Care 16(3): 235-239.
16. Cristancho SM, Apramian T, Vanstone M, Lingard L, Ott M, et al. (2013) When surgeons face intraoperative challenges: a naturalistic model of surgical decision making. Am J Surg 205(2): 156-162.
17. Pauley K, Flin R, Yule S, Youngson G (2011) Surgeons' intraoperative decision making and risk management. Am J Surg 202(4): 375-381.
18. Mhawej R, El Helou M, Moutran A, Chahine R, Akl H, et al. (2025) Objective assessment of cognitive workload in surgery: a systematic review. Ann Med Surg 87(6): 3864-3877.
19. Al-Ghunaim TA, Johnson J, Biyani CS, Alshahrani KM, Dunning A, et al. (2022) Surgeon burnout, impact on patient safety and professionalism: a systematic review and meta-analysis. Am J Surg 224(1 Pt A): 228-238.
20. Sun EC, Kalmar AF, Lien JR, Kim SH, Gormley WB, et al. (2023) Impact of sleep deficiency on surgical performance: a prospective assessment. J Patient Saf 219(4): e60-e67.
21. Edwards MP, Slattery P, Tang J, Chan K, Ahmad M, et al. (2021) Sleepy surgeons: a multi-method assessment of sleep deprivation and performance in surgery. J Surg Res 268: 461-469.
22. Jayaraju S, Nandakumar S, Surendran S, Nair A, Menon R (2024) A review of cognitive support systems in the operating room. Surg Endosc 38(3): 1245-1256.
23. Al-Hiyari MS, Al-Sharif R, Al-Khateeb T, Al-Hadithi A, Al-Rawashdeh M (2022) The measurement of cognitive workload in surgery using pupil metrics: a systematic review and narrative analysis. World J Surg 46(10): 2320-3230.
24. Degani A, Wiener EL (1993) Cockpit checklists: concepts, design, and use. Hum Factors 35(2): 345-359.
25. Dell-Kuster S, Gomes NV, Gawria L, Aghlmandi S, Aduse-Poku M, et al. (2020) Prospective validation of classification of intraoperative adverse events (ClassIntra): international, multicentre cohort study. BMJ 370: m2917.
26. Kaafarani HMA, Mavros MN, Hwabejire J, Fagenholz P, Yeh DD, et al. (2014) Derivation and validation of a novel severity classification for intraoperative adverse events. J Am Coll Surg 218(6): 1120-1128.
27. Gawria L, Rosenthal R, van Goor H, Dell-Kuster S (2022) Classification of intraoperative adverse events in visceral surgery. Surgery 171(6):1502-1509.
28. Gawria L, Dell-Kuster S, Rosenthal R, van Goor H, Breitenstein S, et al. (2023) Appraisal of intraoperative adverse events to improve surgical quality. Br J Surg 110(6): 686-694.
29. Rosenthal R, Hoffmann H, Clavien PA, Bucher HC, Dell-Kuster S (2023) Severity grading systems for intraoperative adverse events: a systematic review of the literature and citation analysis. Ann Surg 278(2): 217-224.
30. Ghaferi AA, Birkmeyer JD, Dimick JB (2009) Complications, failure to rescue, and mortality with major inpatient surgery in Medicare patients. Ann Surg 250(6):1029-1034.
31. Kuo LE, Kaufman EJ, Hoffman RL, Pascual JL, Martin ND, et al. (2019) Failure to rescue in surgical patients: a review for acute care surgeons. J Trauma Acute Care Surg 87(6): 1431-1440.
32. Lorello GR, Cook DA, Johnson RL, Brydges R (2014) Simulation-based training in anaesthesiology: a systematic review and meta-analysis. Br J Anaesth 112(2): 231-245.
33. Weller JM, Nestel D, Marshall SD, Brooks PM, Conn JJ (2016) Improving procedural performance through warm-up and mental practice. Br J Anaesth 117(3): 282-285.
34. Chima AM, Molina A, Sathianarayanan S, Kurup V, Menon R (2025) Effect of interprofessional crisis simulation training in a non-operating room anesthesia setting on team coordination: a mixed methods study. J Clin Anesth 101: 111700.
35. Weller JM, Mahajan R, Fahey M, Fitzsimons J, Gargiulo D, et al. (2026) Effects of a national team training intervention for operating theatre teams on patient and staff outcomes: a stepped-wedge cluster-randomised trial and mixed-methods study. Br J Anaesth 136(4): 712-724.
36. van Klei WA, Hoff RG, van Aarnhem EEHL, Simmermacher RKJ, Regli LPE, et al. (2012) Effects of the introduction of the WHO Surgical Safety Checklist on in-hospital mortality: a cohort study. Ann Surg 255(1): 44-49.
37. De Jager E, McKenna C, Bartlett L, Gunnarsson R, Ho YH (2019) Implementation of the World Health Organization Surgical Safety Checklist at a tertiary hospital: a five-year review. World J Surg 43(2): 377-383.
38. Rai N, Ahmed R, Bhargava N, Sharma S, Gupta A, et al. (2019) Surgical safety checklist compliance: the clinical audit. Indian J Surg 81(5): 447-452.

39. Duclos A, Peix JL, Piriou V, Occelli P, Meyer M, et al. (2016) Cluster randomized trial to evaluate the impact of team training on surgical outcomes. Br J Surg 103(13): 1804-1814.

40. McCulloch P, Mishra A, Handa A, Dale T, Hirst G, et al. (2009) The effects of aviation-style non-technical skills training on technical performance and outcome in the operating theatre. Qual Saf Health Care 18(2): 109-115.

41. Gross B, Rusin L, Kiesewetter J, Zottmann JM, Fischer MR, et al. (2021) What do we really know about crew resource management in healthcare? An umbrella review on crew resource management and its effectiveness. J Patient Saf 17(8): e1704-e1722.

42. Buljac-Samardzic M, Doekhie KD, van Wijngaarden JDH (2020) Interventions to improve team effectiveness within health care: a systematic review of the past decade. Hum Resour Health 18(1): 2.

43. Kurmann A, Tschan F, Semmer NK, Hunziker S, Bolli WB, et al. (2026) The impact of operating room teamwork on patients outcomes: a systematic review. Am J Surg 242:116345.

44. World Health Organization. Surgical Safety Checklist (revised January 2009). Geneva: WHO; 2009.

45. Bartz-Kurycki MA, Anderson KT, Abraham JE, Masada KM, Wang KS, et al. (2017) Debriefing: the forgotten phase of the surgical safety checklist. J Pediatr Surg ;52(5): 866-872.

46. Rose MR, Kimbrough MK, Nolan HR, Day NE, Sexton KW, et al. (2024) Use of an electronic documentation module to improve surgical debriefing effectiveness. Am J Surg 236:115839.

47. Magill ST, Wang DD, Rutledge WC, Lau D, Berger MS, et al. (2017) Changing operating room culture: implementation of a postoperative debrief and improved safety culture. World Neurosurg 107: 597-603.

48. Kauffmann R, Butterworth S, Chen S, Lee C, LeBlanc V, et al. (2017) Effects of perioperative briefing and debriefing on patient safety: a prospective intervention study. BMJ Open 7(12): e018367.

49. Weigl M, Antoniadis S, Chiapponi C, Bruns CJ, Sexton JB (2021) Long-term effects of perioperative briefing and debriefing on team climate: a mixed-method evaluation study. J Multidiscip Healthc 14: 2541-2553.

50. Berenholtz SM, Schumacher K, Hayanga AJ, Simon M, Goeschel C, et al. (2009) Implementing standardized operating room briefings and debriefings at a large regional medical center. Jt Comm J Qual Patient Saf 35(8): 391-397.

51. Guevara OA, Weller JM, Wong DT, Fisher R, Cohen M, et al. (2022) The surgical debrief: just another checklist or an instrument to drive cultural change? Am J Surg 223(1): 123-128.

52. Ricci MA, Brumsted JR (2012) Crew resource management: using aviation techniques to improve operating room safety. Aviat Space Environ Med 83(4): 441-444.

53. Vashdi DR, Bamberger PA, Erez M (2013) Can surgical teams ever learn? The role of coordination, complexity, and transitivity in action team learning. Acad Manage J 56(4): 945-971.

54. Brindle ME, Henrich N, Foster A, Marks S, Rose M, et al. (2018) Implementation of surgical debriefing programs in large health systems: an exploratory qualitative analysis. BMC Health Serv Res 18(1): 210.

55. Etherington N, Wu M, Cheng-Boivin O, Larrigan S, Boet S (2019) Interprofessional communication in the operating room: a narrative review to advance research and practice. Can J Anaesth 66(10): 1251-1260.

56. Clark RR, Basta M, Liu Y, Balk EM, Behbehani S, et al. (2024) The art of speaking up: supporting a culture of safety in the OR. AORN J 120(3): 158-167.

57. Patni N, Bansal S, Sharma A, Gupta R, Mehra A, et al. (2019) Challenging authority and speaking up in the operating room environment: a narrative synthesis. Br J Anaesth 122(2): 233-244.

58. Kolbe M, Grande B, Spahn DR (2015) Briefing and debriefing during simulation-based training and beyond: content, structure, attitude and setting. Best Pract Res Clin Anaesthesiol 29(1): 87-96.

59. Gore DC (2006) National survey of surgical morbidity and mortality conferences. Am J Surg 191(5): 708-714.

60. Beaulieu-Jones BR, Wilson S, Howard DS, Kaczmarek N, Mahoney A, et al. (2023) Defining a high-quality and effective morbidity and mortality conference: a systematic review. JAMA Surg 158(12): 1328-1336.

61. Torres-Landa S, Farrell TM (2026) Evolution of the morbidity and mortality conference in surgery: from Codman to contemporary practice. Am Surg 92(9): 2243-2248.

62. Hutter MM, Rowell KS, Devaney LA, Sokal SM, Warshaw AL, et al. (2006) Identification of surgical complications and deaths: an assessment of the traditional surgical morbidity and mortality conference compared with the American College of Surgeons-National Surgical Quality Improvement Program. J Am Coll Surg 203(5): 618-624.

63. Kim S, Johnson A, Patel R, Nguyen T, Okafor C, et al. (2021) Morbidity and mortality conferences in general surgery: a narrative systematic review. Can J Surg 64(2): E211-E221.

64. Beaulieu-Jones BR, Marchese M, Monaghan A, Kulkarni S, Sandhu A, et al. (2023) Defining a high-quality and effective mortality and morbidity conference: a systematic review. J Am Coll Surg 237(3): 553-563.

65. Ridelberg M, Roback K, Nilsen P (2025) High reliability organizations and healthcare safety outcomes on patients and staff: scoping review. BMC Health Serv Res 25(1): 1304.

66. Provost SM, Lanham HJ, Leykum LK, McDaniel RR, Pugh J (2025) High reliability organization (HRO) principles and patient safety. Rockville: AHRQ PSNet.

67. Beaulieu-Jones BR, Rios-Doria E, Kulkarni S, Sandhu A, Monaghan A, et al. (2026) Reframing the surgical morbidity and mortality conference. J Pediatr Surg 61(4): 161982.

68. Oliver D (2018)Are comparisons between acute healthcare and the aviation industry invidious? BMJ 361:k2203.

69. Rogers J, Gaba D (2011) Have we gone too far in translating ideas from aviation to patient safety? BMJ 342:c7309.

70. Kapur N, Parand A, Soukup T, Reader T, Sevdalis N (2016) Aviation and healthcare: a comparative review with implications for patient safety. JRSM Open 7(1): 2054270415616548.

71. Thomas EJ (2006) Aviation safety methods: quickly adopted but questions remain. Rockville: AHRQ PSNet.

72. Macrae C, Vincent C (2014) Learning from failure: the need for independent safety investigation in healthcare. J R Soc Med 107(11): 439-443.

73. Gogalniceanu P, Calder F, Callaghan C, Sevdalis N, Mamode N (2021) Surgeons are not pilots: is the aviation safety paradigm relevant to modern surgical practice? J Surg Educ 78(5): 1393-1399.

74. (2023) Aviation safety: near misses between commercial flights have increased over the past decade. New York Times.

75. Dias RD, Zenati MA, Stevens RH, Gabany JM, Yule SJ (2022) Assessing team situational awareness in the operating room via computer vision. IEEE CogSIMA PP: 160-164.

76. Perdomo-Pantoja A, Ishida W, Zygourakis C, Holmes C, Iyer RR, et al. (2019) Accuracy of current techniques for placement of pedicle screws in the spine: a comprehensive systematic review and meta-analysis of 51,161 screws. World Neurosurg 126: 664-678.

77. Du YX, Wang Y, Liu H, Zhou Y, Zhang X, et al. (2026) Robot-assisted pedicle screw placement in spinal surgery: an updated systematic umbrella review and meta-meta-analysis of comparisons against freehand and navigated techniques 11(1): 57-69.

1. (2025) Accuracy of augmented reality navigation versus freehand pedicle screw placement: rates of clinically unacceptable breach.
2. Farshad M, Betz M, Farshad-Amacker NA, Moser M, Spirig JM, et al. (2019) Pedicle screw insertion with patient-specific 3D-printed guides based on low-dose CT scan is more accurate than free-hand technique in spine deformity patients: a prospective, randomized clinical trial. Eur Spine J 28(7): 1712-1723.
3. Rahmad Mulyadi, Witantra Dhamar Hutami, Kevin Dilian Suganda, Dhiya Farah Khalisha (2024)Risk of neurologic deficit in medially breached pedicle screws assessed by computed tomography: a systematic review. Asian Spine J 18(6): 903-912.
4. Mohamed El-Meshtawy, Moataz Abdelraheem Ahmed, Ibrahim El Sayed Abdellatif Abuomira, Amr Abdelhalem Amr, Mohamed A A Ibrahim (2025) Validity of a new scoring system for assessment and decision guidance of misplaced pedicular screws. SICOT J 11: 24.
5. Peter A J Pijpker, Joep Kraeima, Max J H Witjes, DL Marinus Oterdoom, Rob A Vergeer, et al. (2021) Accuracy of patient-specific 3D-printed drill guides for pedicle and lateral mass screw insertion: an analysis of 76 cervical and thoracic screw trajectories. Spine (Phila Pa 1976) 46(3): E160-E168.
6. J Cool, J van Schuppen, M A de Boer, B J van Royen (2021) Accuracy assessment of pedicle screw insertion with patient-specific 3D-printed guides through superimpose CT analysis in thoracolumbar spinal deformity surgery. Eur Spine J 30(11): 3216-3224.
7. Peter A J Pijpker, Joep Kraeima, Max J H Witjes, DL Marinus Oterdoom, Maarten H Coppes, et al. (2019) Accuracy assessment of pedicle and lateral mass screw insertion assisted by customized 3D-printed drill guides: a human cadaver study. Neurosurg 196(1): 94-102.
8. Nicola Marengo, Stefano Colonna, Enrico Lo Bue, Alessandro Pesaresi, Ayoub Saaid, et al. (2025) Cervical pedicle screw placement with patient-specific 3D-printed guides: accuracy and safety in a clinical experience 34(4): 1248-1255.
9. Gelalis ID, Paschos NK, Pakos EE, Politis AN, Arnaoutoglou CM, et al. (2012) Accuracy of pedicle screw placement: a systematic review of prospective in vivo studies comparing free hand, fluoroscopy guidance and navigation techniques. Eur Spine J 21(2): 247-255.
10. Aoude AA, Fortin M, Figueiredo R, Jarzem P, Ouellet J, et al. (2015) Methods to determine pedicle screw placement accuracy in spine surgery: a systematic review. Eur Spine J 24(5): 990-1004.
11. Giuseppe F Papalia, Gianluca Vadalà, Fabrizio Russo, Gianmarco Marcello, Niccolò Nardi, et al. (2024) Higher accuracy and better clinical outcomes in navigated thoraco-lumbar pedicle screw fixation versus conventional techniques: a systematic review and meta-analysis. Spine (Phila Pa 1976) 49(19): 1370-1380.
12. Anthony L Mikula, Seth K Williams, Paul A Anderson (2016) The use of intraoperative triggered electromyography to detect misplaced pedicle screws: a systematic review and meta-analysis. J Neurosurg Spine 24(4): 624-635.
13. Hamid Abbasi, Dominic J Moore, Mohadese Rajaeirad, Jiawen Zhan (2021) Screw stimulation thresholds for neuromonitoring in minimally invasive oblique lateral lumbar interbody fusion (OLLIF): a correlational study. Cureus 16(6): e62859.
14. Samyuktha R Melachuri, Manasa K Melachuri, Katherine Anetakis, Donald J Crammond, Jeffrey R Balzer, et al. (1976) Diagnostic accuracy of thresholds less than or equal to 8 mA in pedicle screw testing during lumbar spine procedures to predict new postoperative lower extremity neurological deficits. Spine 46(12): E139-E145.
15. S Parker, Anubhav G Amin, SH Farber, M McGirt, D Sciubba, et al. (2011) Ability of electromyographic monitoring to determine the presence of malpositioned pedicle screws in the lumbosacral spine: analysis of 2,450 consecutively placed screws. J Neurosurg Spine15(2): 130-135.
16. Remi M Ajiboye, Stephen D Zoller, Anthony Doro, Zachary D Burke, William Sheppard, et al. (2017) The utility of intraoperative neuromonitoring for lumbar pedicle screw placement is questionable: a review of 9,957 cases. Spine (Phila Pa 1976) 42(17): E1026-E1031.
17. JY Du, JS Wu, ZQ Wen, XJ Lin (2016) Treatment strategies for early neurological deficits related to malpositioned pedicle screws in the lumbosacral canal. Bone Joint Res 5(2): 46-51.
18. Jinshan Tang, Ziqiang Zhu, Tao Sui, Dechao Kong, Xiaojian Cao (2014) Position and complications of pedicle screw insertion with or without image-navigation techniques in the thoracolumbar spine: a meta-analysis of comparative studies. J Biomed Res 28(3): 228-239.
19. Özcan Sönmez, Yahya Güvenç, Buket Gedik, Kadir Kolçak (2026) Risk of vertebral artery injury associated with subaxial cervical pedicle screw malposition: cadaveric evidence and clinical relevance.
20. Sudhir Suggala, Garrett A Dyess, Olivier Darbin, Richard Philip Menger (2024) Delayed neurological deficit due to a medially misplaced thoracic pedicle screw during adolescent idiopathic scoliosis correction: a complication six years in the making. Spine Deform 12(4): 311-324.
21. Xiaobang Hu, Isador H Lieberman (2014) What is the learning curve for robotic-assisted pedicle screw placement in spine surgery? Clin Orthop Relat Res 472(6): 1831-1836.
22. Siddiqui MI, Wallace DJ, Salazar LM, Vardiman AB (2019) Robot-assisted pedicle screw placement: learning curve experience. World Neurosurg 130: e417-e422.
23. Martin Wood (2014) The surgical learning curve and accuracy of minimally invasive lumbar pedicle screw placement using CT-based computer-assisted navigation plus continuous electromyography monitoring: a retrospective review of 627 screws in 150 patients. Int J Spine Surg 8: 27.
24. Joseph Maalouly, Mehul Sarkar, John Choi (2021) Retrospective study assessing the learning curve and the accuracy of minimally invasive robot-assisted pedicle screw placement during the first 41 robot-assisted spinal fusion surgeries. J Spine Surg 7(2): 184-189.
25. Pedicle screw placement accuracy in robot-assisted spinal fusion in a multicenter study. Boise State University Mechanical Engineering Faculty Publications and Presentations.
26. Mirza Pojskić, Miriam Bopp, Omar Alwakaa, Christopher Nimsky, Benjamin Saß (2025) Comparative analysis of the pedicle screw accuracy, screw revision and loosening rate and radiation exposure of robotic-guided, intraoperative CT-navigation guided, and fluoroscopy guided placement technique. Brain Spine 4(6): 105899.