



# Critical views for safe surgical phase progression in endoscopic endonasal transsphenoidal pituitary adenoma resection: modified Delphi consensus

Tjasa Zaletel<sup>1,2</sup> · Danyal Z. Khan<sup>2,3</sup> · Anjana Wijekoon<sup>3,4</sup> · Zhehua Mao<sup>3,4</sup> · Joao Paulo Almeida<sup>5</sup> · Anouk Borg<sup>2</sup> · Jonathan Chainey<sup>6</sup> · Michael D. Cusimano<sup>7</sup> · Daniel A. Donoho<sup>8,9</sup> · Neil Dorward<sup>2</sup> · Juan Carlos Fernandez-Miranda<sup>10</sup> · Giorgio Fiore<sup>11</sup> · Theofanis Giannis<sup>2</sup> · Alfonso Lagares Gomez-Abascal<sup>12,13,14</sup> · Lauren Harris<sup>2</sup> · Abhiney Jain<sup>2,3</sup> · Ruth Lau<sup>15</sup> · Sacit B. Omay<sup>16</sup> · Igor Paredes<sup>12,13,14</sup> · Daniel Prevedello<sup>17</sup> · Gabriel Zada<sup>18</sup> · Danail Stoyanov<sup>3,4</sup> · Sophia Bano<sup>3,4</sup> · Hani J. Marcus<sup>2,3</sup>

Received: 1 December 2025 / Accepted: 4 January 2026  
© The Author(s) 2026

## Abstract

**Purpose** Endonasal transsphenoidal surgery is the gold-standard for pituitary adenoma resection, yet no intraoperative framework exists to confirm safe phase progression. Inspired by the Critical View of Safety in laparoscopic cholecystectomy and engineering “phase-gate” process, we propose the Critical Views for Phase Progression (CVPPs) – a set of visual cues confirming phase objectives and safe phase progression. Designed to be clinically relevant and machine-readable, CVPPs aim to support training and future AI-driven guidance systems.

**Methods** A three-round modified Delphi process was conducted involving 15 pituitary surgery experts from 13 centres across Europe and North America. CVPPs for the naso-sphenoid, sellar, and closure phases were classified as “Essential”, “Desirable” or “Not Necessary”. Consensus required  $\geq 70\%$  agreement. A local validation study was subsequently performed involving six experts who reviewed 15 intraoperative video clips and rated their confidence to proceed, which was compared against the predefined reference derived from the finalised CVPPs.

**Results** Consensus identified essential and desirable CVPPs across all three phases for both micro- and macroadenoma variants, reflecting differences in exposure goals and surgical risk. Validation demonstrated high concordance between participant ratings and predefined references. Discrepancies arose only in a minority of intentionally incomplete (“unsafe”) views and were attributable to contextual misinterpretation of short video segments, rather than disagreement with the CVPP framework.

**Conclusion** This international, multicentre consensus is the first to define CVPPs. By standardising intraoperative visual benchmarks, CVPPs can enhance training, mitigate risks, and provide a foundation for future AI-driven guidance systems capable of real-time anatomical annotation and decision support.

**Keywords** Consensus · CVPP · Delphi · Endoscopic endonasal · Endoscopic transsphenoidal surgery · Pituitary adenoma

## Introduction

Endoscopic endonasal transsphenoidal surgery has become the gold-standard surgical approach for pituitary adenoma resection, offering a minimally invasive route and

reduced morbidity compared to transcranial approaches [1–5]. Despite its widespread adoption, significant variability exists in the performance of the procedure and the resulting outcomes [1, 6–8]. Our group has recently developed a consensus-derived workflow analysis framework

Tjasa Zaletel and Danyal Z Khan contributed equally to this work.

Extended author information available on the last page of the article

for pituitary surgery, which breaks the surgery down into phases (naso-sphenoid, sellar, closure) and their constituent steps, in order to objectively characterise operations [1]. However, the applicability of this framework to improve the quality of surgery intraoperatively in real-time has not yet been realised.

In other fields, workflow analysis has been applied intraoperatively via the Critical View of Safety (CVS) concept [9–12]. This establishes key criteria that must be present to progress safely from one surgical phase to the next. First described by Strasberg in 1995 as a target identification method, CVS in laparoscopic cholecystectomy (LC) promotes the recognition of the cystic duct and artery before proceeding with division to minimise the risk of iatrogenic bile duct injury [13]. It is now considered one of the most important critical factors for overall safety during LC by the Society of American Gastrointestinal and Endoscopic Surgeons expert group [14]. CVS has been used in audit, quality measures, outcome predictions [15–17], and more recently to develop real-time artificial intelligence (AI) models to validate CVS achievement during elective LC [18–21].

We sought to adapt a similar approach to pituitary surgery and align it to previous workflow analyses across all surgical phase transitions. We introduce the concept of “Critical Views for Phase Progression (CVPP)”, a set of visual cues that represent completion of the surgical objective of that phase and readiness to proceed to the next phase. While inspired by the CVS concept, CVPPs aim to define minimum visual criteria for safe phase progression in a complex, multi-phase procedure and, therefore, do not intend to capture full operative nuance. This approach also draws inspiration from engineering principles, particularly the phase-gate process (“waterfall model”), a widely used model in systems engineering and project management to evaluate progress and assess readiness at predefined checkpoints before proceeding to subsequent phases [22–25]. By adapting this method to the surgical workflow, each surgical stage is treated as a distinct phase with a “gate” requiring confirmation of essential visual cues and satisfactory conditions before proceeding (Fig. 1).

This verification process serves multiple purposes. Firstly, it ensures adequate exposure and anatomical orientation, thereby mitigating intraoperative risks (e.g., inadvertent neurovascular injury). Secondly, it enhances surgical standardisation, facilitating surgical training and reducing variability. Finally, beyond its immediate clinical application, it establishes a foundation for future AI-driven intraoperative guidance, as these defined visual cues can serve as training benchmarks for AI systems to assist surgeons in real-time and guide safe progression to the next phase through automated anatomical annotation

of the surgical field to improve consistency, efficiency and outcomes.

This study aims to establish an international expert consensus on the visual cues required for safe progression through each surgical phase of endonasal transsphenoidal pituitary adenoma resection and to validate whether these consensus-defined CVPPs reliably capture the visual information needed for safe intraoperative decision-making across different clinical cases and operating surgeons. This is the first such analysis to define phase-specific criteria for all phase transitions within a single operation, providing a standardised framework for intraoperative decision-making.

## Methods

### Overview

This study employed a modified Delphi process [26] to establish consensus on the CVPPs in endonasal transsphenoidal pituitary adenoma resection. These phase-specific visual cues serve as objective intraoperative benchmarks, ensuring that essential procedural goals (e.g., anatomical identification, haemostasis, structural integrity) are met before transitioning to the next phase. Essential cues were defined as mandatory criteria that must be visualised for safe progression, whereas desirable cues are identified as beneficial for surgical precision but not mandatory for phase transition. We aim for these criteria to be both clinically relevant and machine-readable, facilitating their integration into surgical practice and future AI-driven intraoperative guidance systems. For consistency in workflow annotation, the nasal and sphenoid components were combined into a single naso-sphenoid phase to minimise short transitions, simplify workflow segmentation, and facilitate consistency for data structuring and future model training. This study focused exclusively on standard endoscopic transsphenoidal pituitary adenoma resections, excluding transcavernous and other extended endonasal approaches.

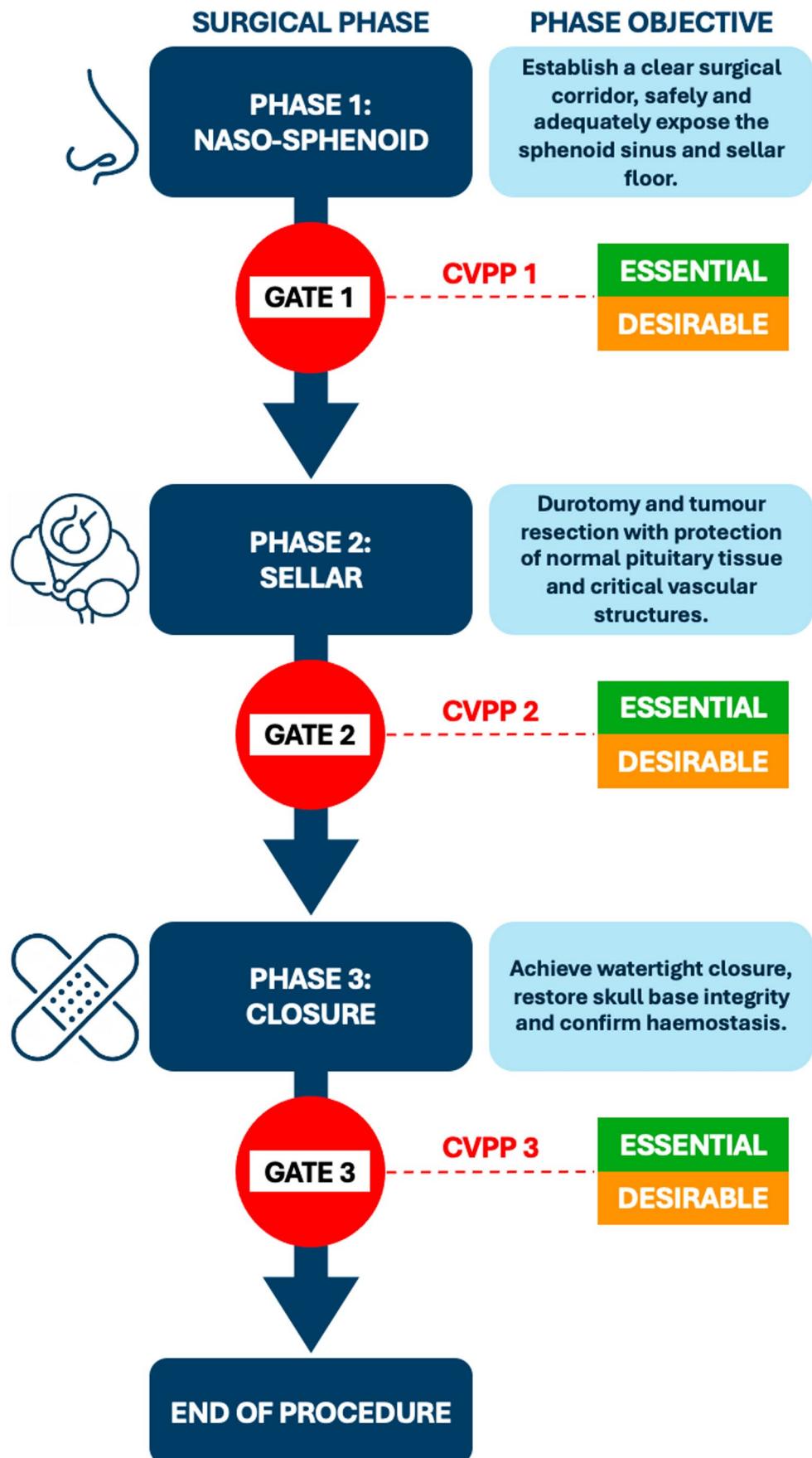
### Modified Delphi process and sampling

An initial set of CVPPs and Delphi questions was developed by a local steering committee consisting of three researchers, including a consultant neurosurgeon.

### Participant selection

A panel of international expert members was selected from our established collaborative pituitary surgery networks

**Fig. 1** Modified “Phase-Gate” Model for Endoscopic Transsphenoidal Pituitary Adenoma Surgery. Inspired by the Phase-Gate Process (Waterfall Model) used in systems engineering, this modified framework structures the surgical workflow into three distinct phases: Naso-Sphenoid, Sellar, and Closure. Each phase has a clearly defined objective, and progression between phases is governed by three Critical Views for Phase Progression (CVPP 1–3), which act as decision checkpoints (“gates”) requiring confirmation of a set of visual cues before safely progressing to the next surgical phase. These cues are categorised as essential (mandatory for safe transition) and desirable (helpful for enhanced safety or precision but not mandatory), and indicate that the surgical objective of each phase has been successfully achieved, ensuring systematic intraoperative validation, minimising risk, and enhancing surgical standardisation



based on their experience and expertise in pituitary and skull base surgery, with additional experience in surgical data science (computer vision [CV] and AI). Eligible participants included expert neurosurgeons (i.e., consultants) with active roles in surgical training, research, and guideline development. This ensured that the resulting framework was both clinically grounded and adaptable to future digital applications. Participants were recruited through direct email invitations [27]. Each round included participants from the previous rounds with additional new experts joining as the study expanded. The documents were presented using Microsoft Word (Version 16.94, Microsoft, Washington, USA) and administered using Google Forms (Google LLC, California, USA).

### Consensus round 1

The initial CVPPs were distributed to a small group of experts ( $n=9$ ). For each phase, example intraoperative images were provided to illustrate a potential view at the end of that phase, serving as a reference to facilitate the consensus process (Appendix A). Experts were asked a series of questions seeking to assess the accuracy and completeness of visual cues within each CVPP; this included classifying each visual cue as essential, desirable, or not necessary, and providing additional input on missing elements or suggesting modifications for evaluation in subsequent rounds (Appendix B).

According to the Delphi technique, circulation and iterative revision of the visual cues within each CVPP were repeated until data saturation was achieved, i.e., until all experts were satisfied that the CVPPs were complete and accurate. Round 1 was repeated two times, occurring over three weeks (March–April 2025). The final agreement thresholds were pre-defined as  $\geq 70\%$  for inclusion and  $\geq 85\%$  for exclusion (i.e., rated “not necessary”). Contentious cues ( $< 70\%$  agreement) were considered in subsequent rounds with clarifying examples and guidance. Any additional suggestions were added if (i) in-scope and (ii) not duplicate.

### Consensus round 2

Refined CVPPs were then distributed to a larger group ( $n=11$ ; 9 from prior round, 2 new). Experts were asked to assess the revised CVPPs and expand the defined domains to cover possible global variations in practice. This round was repeated until all the experts agreed that the visual cues reflected the operative practice, and there were no additional suggestions from the participant group. The final inclusion was based on the above-defined consensus thresholds. As with Round

1, additional suggestions were reviewed and added if appropriate. Round 2 was repeated twice, occurring over a three-week period (April–May 2025).

### Consensus round 3

Round 3 was conducted with an expanded international expert panel ( $n=15$ ; 11 from previous rounds, 4 new) and focused exclusively on contentious visual cues (i.e.,  $< 70\%$  agreement), along with any new cues proposed by the panel. Visual cues that had already achieved consensus (i.e.,  $\geq 70\%$  agreement) in prior rounds were fixed and excluded from re-rating. A summary of results from previous rounds was shared with the panel, clearly indicating which visual cues had already reached consensus and which required re-rating. In addition, clarifying examples were added to cues with a high degree of variability to guide consistent interpretation. Round 3 was completed over an eight-week period (July–September 2025), finalising consensus on all remaining contentious visual cues.

### Data collection and analysis

Participant demographics collected included training grade, years of experience in transsphenoidal surgery, and associated institution. The Delphi surveys regarding the CVPPs consisted of Likert-scale items (essential, desirable, not necessary) along with open-ended fields for qualitative input (additional suggestions or comments). Content analysis was used to analyse free-text responses, which included removing out-of-scope suggestions, grouping similar suggestions together, and comparing them to existing data points.

### Validation study

A local validation study was conducted to evaluate whether the finalised CVPPs reliably capture the visual information for safe intraoperative decision-making. The validation was performed using a structured survey (Qualtrics XM, Utah, USA) administered via email to a network of neurosurgeons performing endoscopic pituitary surgery ( $n=6$ ) at a tertiary neurosurgical centre in the UK, including board-certified surgeons ( $n=3$ ), those undertaking a dedicated pituitary fellowship ( $n=2$ ), and a resident with interest in skull base surgery ( $n=1$ ). The inclusion of surgeons at various stages of training and independent practice was intentional, ensuring that the validation reflected those most likely to utilise and benefit from the CVPP framework in real-world practice.

Participants were shown a total of 15 short ( $\sim 10$  s) intraoperative video clips, representing potential endoscopic views at the end of each surgical phase. Clips were chosen

instead of static frames to allow assessment of dynamic cues (e.g., CSF leakage, arachnoid movement), and were intentionally short to isolate end-phase visual information and minimise reliance on procedural context. For each clip, experts rated their confidence to proceed using three categories (“clearly safe”, “just adequate – proceed with caution”, “not safe”). Prior to assessment, participants were provided with the finalised CVPP framework and a brief guidance on its intended use. The responses were automatically compared against a predefined reference classification set by our study group; where discrepancies occurred, participants were prompted to provide their reasoning and feedback. Clips were purposively selected to represent both complete (“safe”) and incomplete (“unsafe”) end-phase views. As a guide, our study group classified views missing any essential visual cue as “unsafe”, those containing all essential cues as “just adequate”, and those showing all essential plus at least one desirable cue as “clearly safe” to proceed. For bilateral structures (e.g., carotid prominences), inadequate visualisation of either side was treated as failure to meet the essential criterion. For analysis, ratings were collapsed into a binary outcome (i.e., “safe to proceed” vs. “do not proceed” i.e., combining “clearly safe” and “just adequate” categories) to better reflect real-world decision-making. Responses were analysed for concordance to assess the reliability and face validity of the CVPP framework for real-world intraoperative application.

## Ethics

This study involved no patient data, participation was voluntary and all participant responses were fully anonymised. All video clips were sourced from our institutional research repository of routinely collected operative recordings, with prior patient consent for research use. The topic was non-sensitive, and data collection did not involve human subjects or clinical interventions; therefore, formal ethical approval was not required in accordance with our local institutional policy (see UCL exemptions).

## Results

### General

Response rate remained high across Delphi rounds (89% in Round 1, 100% in Round 2, 87% in Round 3) with no significant attrition across Delphi rounds. Overall, 15 expert panel members from 13 centres across five countries participated: UK ( $n=3$ ), USA ( $n=6$ ), Canada ( $n=2$ ), Spain ( $n=3$ ), and Italy ( $n=1$ ).

### Final CVPP framework

Consensus on CVPPs was progressively refined across three iterative rounds through quantitative rating, structured feedback, and re-assessment of contentious items. Round 1 generated predominantly qualitative input, leading to structural refinements with subdivision of naso-sphenoid and sellar phases into macro- and microadenoma cases to reflect differing exposure objectives between tumour types. Subsequent rounds focused on resolving variability by standardising definitions, providing additional guidance, and clarifying examples. The final set of consensus-defined CVPPs is presented in Table 1.

### Naso-sphenoid phase

Following feedback from Round 1 on exposure sufficiency, the naso-sphenoid phase was subdivided into macro- and microadenoma tumour types to account for differences in surgical objectives. The view at the end of naso-sphenoid phase comprised 7 visual cues for macroadenomas (4 essential, 3 desirable) and 7 for microadenomas (5 essential, 2 desirable), focusing on midline orientation and critical neurovascular landmarks.

Nearly all visual cues in this phase reached consensus. Of note, optic nerve prominences (ONPs) in macroadenomas, and the clival recess and tuberculum sellae in microadenomas, fell just below the predefined consensus threshold, with responses split between “Essential” and “Desirable” categories. Following internal discussion within the study group, these cues were classified in alignment with the classifications used in the alternate tumour subtype to maintain cross-phase consistency and support logical continuity for future model development. Therefore, ONPs were categorised as “Desirable”, and clival recess and tuberculum sellae as “Essential” for both tumour types. For the cavernous sinus, inter-rater variability was resolved when the panel agreed that visualising its margins holds particular importance in invasive microadenomas extending laterally into the medial compartment, leading to consensus.

### Sellar phase

The sellar phase generated the most detailed discussion and greatest inter-rater variability, particularly concerning the degree of tumour resection required before progressing to closure. The final CVPPs were defined separately for macro- and microadenomas, with consensus achieved on 6 visual cues (5 essential, 1 desirable) for each. A small number of cues (i.e., normal pituitary tissue and CSF leak±arachnoid tear detection in macroadenomas) initially fell below the

**Table 1** CVPPs across surgical phases in endoscopic transsphenoidal pituitary adenoma resection

Phase	Tumour type	Visual cues	Desirable
		Essential	
Naso-Sphenoid	Macroadenoma	<ul style="list-style-type: none"> <li>• Sellar bulge</li> <li>• Carotid protuberances</li> <li>• Clival recess</li> <li>• Tuberculum sellae</li> </ul>	<ul style="list-style-type: none"> <li>• Opticocarotid recess</li> <li>• Optic nerve prominences</li> <li>• Planum sphenoidale</li> <li>• Opticocarotid recess</li> <li>• Optic nerve prominences</li> </ul>
	Microadenoma	<ul style="list-style-type: none"> <li>• Sellar bulge</li> <li>• Carotid protuberances</li> <li>• Clival recess</li> <li>• Tuberculum sellae</li> <li>• Cavernous sinus margins (<i>if applicable</i>)</li> </ul>	
Sellar	Macroadenoma	<ul style="list-style-type: none"> <li>• Sellar fossa clearance of tumour</li> <li>• Sellar floor* visualisation &amp; clearance</li> <li>• Arachnoid dome descent and suprasellar compartment clearance</li> <li>• Residual tumour identification (<i>if goal debulking, then acceptable presence; if goal resection, then absence</i>)</li> <li>• CSF leak±arachnoid tear detection</li> </ul>	<ul style="list-style-type: none"> <li>• Normal pituitary tissue identification</li> </ul>
	Microadenoma	<ul style="list-style-type: none"> <li>• Sellar fossa clearance of tumour</li> <li>• Sellar floor* visualisation &amp; clearance</li> <li>• Residual tumour identification (<i>if goal debulking, then acceptable presence; if goal resection, then absence</i>)</li> <li>• CSF leak±arachnoid tear detection</li> <li>• Normal pituitary tissue identification</li> </ul>	<ul style="list-style-type: none"> <li>• Medial wall of cavernous sinus inspection</li> </ul>
Closure	Macroadenoma & Microadenoma	<ul style="list-style-type: none"> <li>• Haemorrhage control</li> <li>• No active CSF leak</li> <li>• Stability of repair construct</li> <li>• Vascularised flap viability (<i>if applicable</i>)</li> </ul>	<ul style="list-style-type: none"> <li>• Repair construct pulsatility (<i>if applicable</i>)</li> </ul>

\* This refers to the surgical floor at the inferior aspect of the resection cavity, which approximately represents the posterior wall of the sella extending to the dorsum sellae anatomically

inclusion threshold with the ratings split between “Essential” and “Desirable” categories. After internal group discussion, both were included in the “Essential” category to maintain consistency between tumour subtypes and to reflect their clinical importance in confirming adequate decompression and haemostasis.

### Closure phase

Consensus for the closure phase was achieved rapidly with a low degree of inter-rater variability. The CVPPs for this phase were the same for macro- and microadenomas, comprising 5 visual cues (4 essential, 1 desirable) focusing on adequate haemostasis, watertight closure, and construct stability. Early variability regarding flap viability was resolved by clarifying that this cue applies when a vascularised flap is used in reconstruction, achieving full consensus in the final round.

The final set of CVPPs represent a unified international consensus on essential and desirable visual cues for safe phase progression in endoscopic pituitary adenoma resection. Figure 2 illustrates their integration within the modified phase-gate framework, showing (A) schematic depictions alongside (B) real-life annotated intraoperative

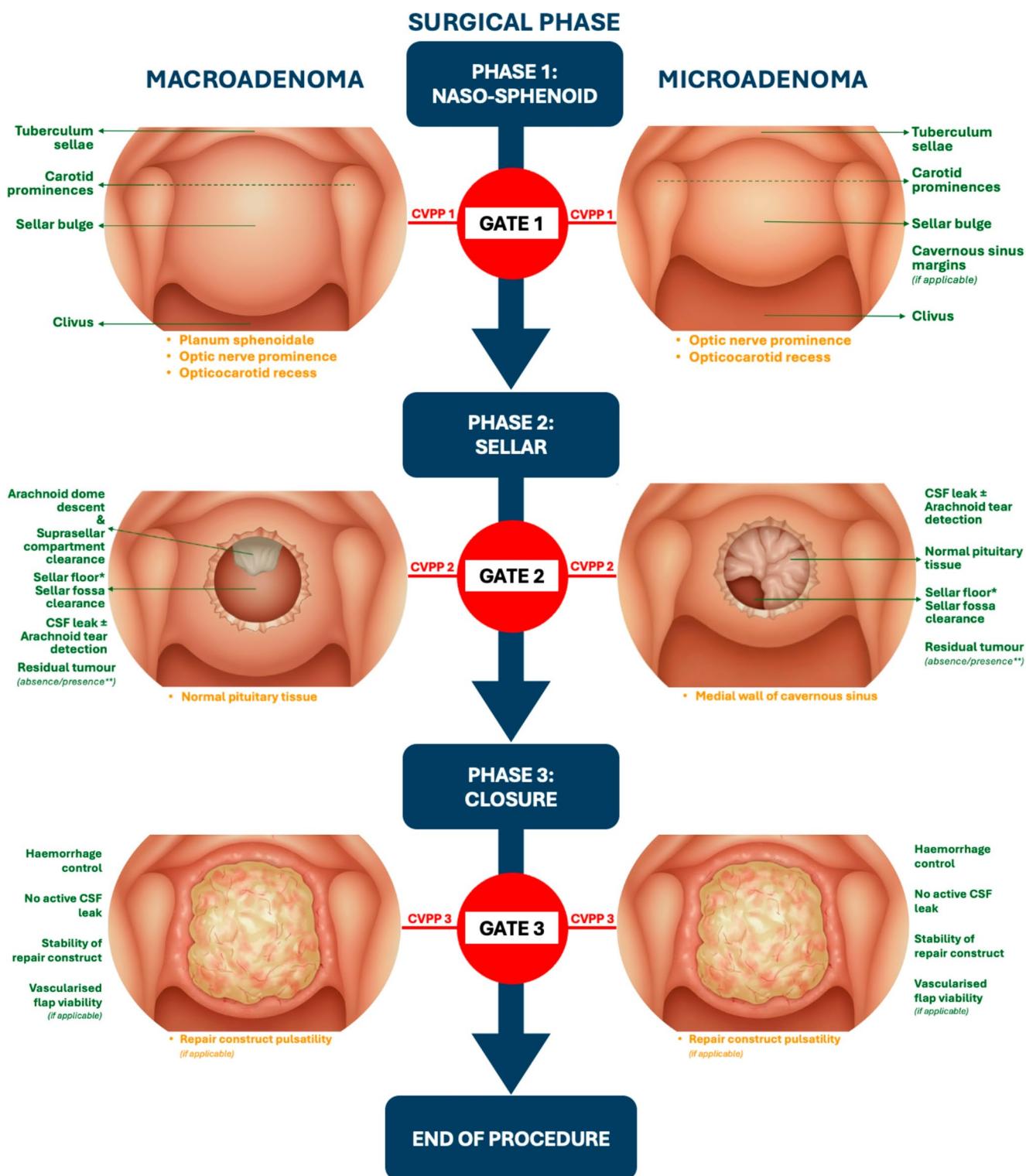
examples that demonstrate how these critical views appear under operative conditions.

### Validation study

A total of 90 responses were collected, which were compared to our predefined reference classification. A summary of participant agreement and qualitative feedback in cases of discordance is presented in Table 2.

Participant ratings showed high concordance with the predefined reference, with a mean agreement of 94% across all clips. Agreement was complete for all but one “safe” clip (i.e., 9 out of 10), confirming that the CVPPs accurately reflected intraoperative readiness to proceed. The single instance of disagreement occurred when one participant rated a view as unsafe due to perceived incomplete visualisation of the carotid protuberances and requested additional navigation data, reflecting caution rather than disagreement with the CVPPs.

Across all subphases, five clips represented intentionally “unsafe to proceed” views with lower agreement levels. Two instances of partial discordance were noted, with 2 out of 6 participants (33%) rating the view as “safe”; both involved microadenoma cases with incomplete carotid



**Fig. 2** Modified “Phase-Gate” Model with Visual Cues for Each CVPP in Endoscopic Transsphenoidal Pituitary Adenoma Surgery. The diagrams with illustrated (A) and intraoperative endoscopic (B) example views summarise essential (green) and desirable (yellow) visual cues defining safe transition between the naso-sphenoid, sellar, and closure phases as defined by the modified Delphi consensus process. Separate depictions for macroadenoma and microadenoma reflect phase-

specific exposure objectives. \* This refers to the surgical floor at the inferior aspect of the resection cavity, which approximately represents the posterior wall of the sella extending to the dorsum sellae anatomically. \*\* If goal debulking, then acceptable presence; if goal resection, then absence. **A** Schematic illustration of the modified phase-gate CVPP framework. **B** Annotated intraoperative endoscopic example views of CVPPs

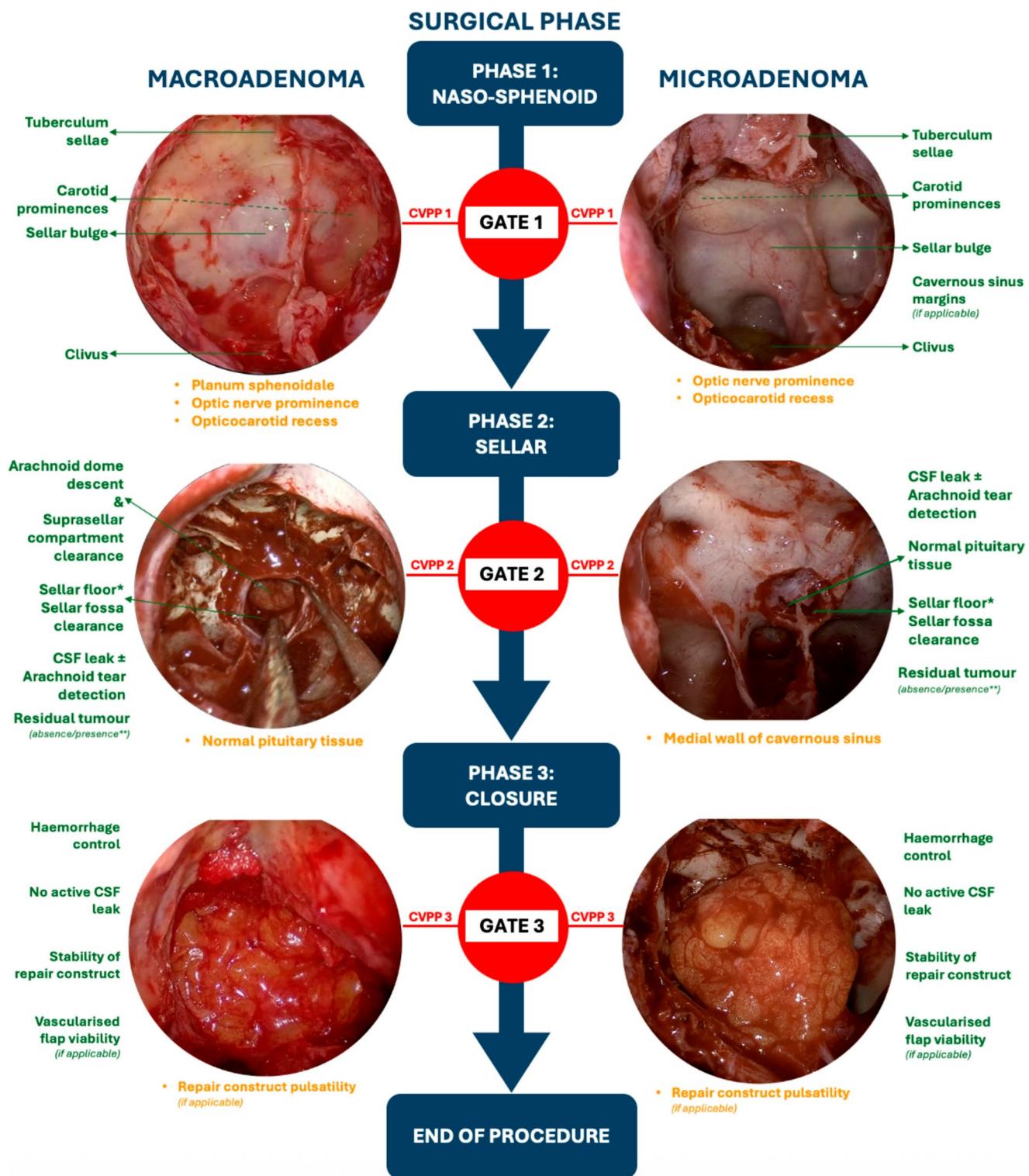


Fig. 2 (continued)

(naso-sphenoid phase) or sellar floor visualisation (sellar phase). Qualitative feedback suggested these discrepancies reflected uncertainty about surgical stage – specifically,

whether clips represented exposure or resection endpoints, rather than the CVPP criteria.

**Table 2** Validation study results - participant agreement with reference classification with summary of feedback for discordant cases

PHASE	TUMOUR TYPE	No. of clips	Reference classification	Agreement with reference (%)	False-proceed cases (n, %)	False-hold cases (n, %)	Qualitative feedback for discordance
<i>Naso-sphenoid</i>	<i>Macro</i>	3	2 safe; 1 unsafe	94%	0	1 (8%)	Limited carotid visualisation
	<i>Micro</i>	3	2 safe; 1 unsafe	89%	2 (33%)	0	Adequate exposure with endoscope repositioning
<i>Sellar</i>	<i>Macro</i>	3	2 safe; 1 unsafe	100%	0	0	
	<i>Micro</i>	3	2 safe; 1 unsafe	89%	2 (33%)	0	One participant acknowledged misjudgement; another noted further resection is required
<i>Closure</i>	<i>Macro &amp; Micro</i>	3	2 safe; 1 unsafe	100%	0	0	

## Discussion

### Principal findings

This study is the first to define “Critical Views for Phase Progression (CVPPs)” and apply them in the context of endoscopic transsphenoidal pituitary adenoma resection using an international, expert-driven consensus process. Building on our previously published operative workflow for this procedure [1], the current Delphi analysis shifts the focus from task sequencing and how surgery is performed to phase-specific visual confirmation criteria and how readiness to progress is determined. It introduces a novel framework aimed at supporting safe intraoperative decision-making and enhancing standardisation across surgical practice.

Using a modified Delphi methodology, we identified a core set of phase-specific essential and desirable visual cues that define safe progression between major surgical phases (i.e., naso-sphenoid, sellar, closure). Our proposed CVPP framework adapts principles from surgical safety, systems engineering, and workflow science to define what surgeons must visualise before advancing to the next stage of surgery [22, 23, 28]. The agreed CVPPs can be used for education (e.g., operative video annotation), surgical skills assessment, clinical auditing, and the development of surgical safety checklists, models, and simulators [21, 29–32].

This work is a product of an international, multicentre consensus, capturing variations in surgical practice and training backgrounds. Given this breadth of representation, achieving a structured framework and maintaining consistent terminology proved challenging and required several iterative refinements across multiple Delphi rounds. Early discussions highlighted variability in exposure standards and anatomical emphasis among experts, depending on the operative contexts (e.g., tumour size, invasiveness), which prompted the creation of separate CVPPs for macro- and microadenoma cases. Additional complexity arose because

the study introduced a novel concept that had not been previously defined in the literature. Iterative refinement across Delphi rounds ensured experts shared a consistent understanding of the terminology and definitions, using examples to resolve interpretative discrepancies and accommodate the full spectrum of operative practice. This ultimately produced a unified framework that is both methodologically consistent and flexible enough to reflect real-world variability in surgical decision-making within a shared safety structure.

The validation of the CVPP framework confirmed high overall concordance between expert judgement and the pre-defined reference classifications, supporting its face validity and clinical reliability. While lower levels of agreement were observed for intentionally incomplete or “unsafe” views, qualitative feedback suggested that these discrepancies in the decision to proceed arose from a contextual misinterpretation of the short video clips (i.e., a lack of procedural context) rather than disagreement with the visual cues themselves. This confirms that CVPPs reliably capture critical visual information that surgeons use, often intuitively, to judge intraoperative readiness to proceed. It is important to note that the present study was designed as a face-validity assessment of the CVPP framework rather than an educational intervention and future work is needed to evaluate whether the structured pre- and post-educational exposure improves overall performance.

### Findings in the context of existing literature

The concept of visually confirming safety before procedural progression is not new to surgery but has rarely been formalised beyond specific specialities. The most established example is Strasberg’s “Critical View of Safety” in LC [9, 11, 13, 16], which significantly reduced iatrogenic bile duct injury through direct visual confirmation of key anatomical landmarks (hepatocystic triangle, cystic plate,

cystic duct and artery) before transection of the cystic duct and artery [10, 33–35]. Subsequent studies have validated the utility of CVS as both a teaching and safety tool, demonstrating how structured visual verification at critical transition points can improve outcomes, enable standardised analysis of intraoperative performance, and guide the development of intelligent systems [20, 35–38]. Audit-based initiatives using intraoperative photo and video documentation of CVS have shown that satisfactory visual confirmation strongly correlates with lower rates of post-operative complications, establishing CVS adherence as a measurable quality indicator in institutional and national audits [35, 38]. In surgical education, structured curricula have been shown to improve surgeons' attainment of CVS [36], while studies also highlight that self-assessment often overestimates performance [38], highlighting the need for objective feedback and standardised assessment frameworks. Building on these efforts, Rios et al. recently introduced the Cholec80-CVS – an open dataset providing expert-annotated operative footage for the development and benchmarking of CV models capable of automatically detecting anatomical landmarks, validating CVS achievement, and flagging unsafe dissection zones in real time [20]. These developments demonstrate how the CVS framework has evolved from a visual safety principle to a foundation for quality improvement, training, and AI-driven surgical guidance – all directly translatable to the proposed CVPP framework. However, while CVPPs draw inspiration from the CVS, the two frameworks differ fundamentally in scope and intent. The CVS principle focuses on a single anatomical juncture with a clearly defined adverse outcome, therefore its strength lies in the simplicity and specificity of the intervention. The CVPP framework expands on this by embedding visual confirmation into every major surgical phase transition, establishing a continuous system of structured checkpoints that collectively define safe operative progression within a complex, context-dependent procedure. As a result, a direct one-to-one translation of the CVS concept is neither feasible nor intended. The CVPP framework deliberately operates at a higher level of abstraction, defining minimum visual criteria that indicate readiness to progress between surgical phases rather than attempting to encode the full nuance of operative technique or expert judgement. This abstraction is intentional and enables complex surgical workflows to be structured around predefined safety checkpoints without constraining case-specific strategy.

Beyond its surgical roots, the CVPP framework draws inspiration from engineering risk management principles. In systems engineering, the phase-gate (“waterfall”) process is a technique in which a project is divided into distinct stages (“phases”), separated by decision points (“gates”)

[22, 23]. This provides a structured mechanism for assessing readiness and mitigating risk by ensuring that each developmental stage is advanced only once specific, objectively verifiable criteria have been met [22, 23, 39, 40]. Applying this philosophy to the operative environment, the CVPP framework introduces a similar mechanism of safety gating within surgery, integrating a proactive form of hazard control into the operative workflow – one that can be standardised, auditable, and amenable to future digital translation. Engineering-derived frameworks are increasingly recognised in healthcare for their ability to standardise complex processes, reduce cognitive load, and minimise error propagation – all of which align with the principles of modern surgical safety systems [41–44]. This cross-disciplinary translation situates the CVPP framework within a growing movement towards structured, data-driven, cognitively intuitive surgery. Efforts to standardise operative workflows and develop machine-readable frameworks have been increasingly more common across various surgical fields – from minimally invasive general surgery to robotic procedures [45–49]. In this context, CVPPs represent a natural progression from our previously published workflow, translating implicit decision-making into explicit and quantifiable data.

Although applied to pituitary surgery in this study, the CVPP model is inherently procedure-agnostic. Its principles (i.e., structured progression, predefined safety criteria, and visual confirmation) are universally applicable across surgical fields where phase-based transitions occur (e.g., robotic-assisted surgery). The framework's modularity allows local or speciality-specific visual criteria to be defined, validated, and incorporated into training, quality improvement systems and safety checklists. As such, CVPPs offer a scalable foundation for intraoperative safety validation and future automation across diverse operative settings.

Finally, a defining feature of the CVPP framework is its machine-readable structure. By formalising phase completion as a finite set of visually verifiable cues, CVPPs provide ground truth suitable for training CV models in phase recognition, anatomical landmark detection, and automated safety validation. This positions CVPPs as a bridge between clinical reasoning and machine-readable safety benchmarks, shifting surgical AI models from purely descriptive workflow and anatomy recognition toward explicit safety verification, expanding existing workflow and anatomy-recognition efforts in endoscopic transsphenoidal surgery [50–52].

## Clinical implications

CVPPs are intentionally designed to raise the floor, not the roof of surgical performance. Their primary aim is,

therefore, to define the minimum visual criteria a surgeon would reasonably expect to satisfy before phase progression, rather than to modify expert surgical technique. By making these safety-critical visual objectives explicit, reproducible, and assessable, CVPPs may provide a structured framework for use in training, audit, and structured assessment purposes. Clinically, the CVPP framework provides a reproducible tool for intraoperative verification and surgical audit with the potential to standardise intraoperative decision-making (e.g., akin to WHO Surgical Safety Checklist [28]). From a safety perspective, embedding visual cues into intraoperative checklists could reinforce visual verification as a formalised step before proceeding. Educationally, CVPPs transform tacit visual information into teachable and assessable benchmarks for surgical trainees, enabling structured feedback and objective assessment of competency-based progression. Evidence from CVS studies [36, 53] suggests that integrating CVPPs into surgical curricula or video-based learning may enhance recognition of the anatomical and procedural hallmarks underpinning safe phase completion, bridging the gap between observation and independent practice. The local validation study intentionally included surgeons at various stages of training and independent practice, including fellows and residents, who stand to benefit most from structured visual frameworks.

Beyond immediate clinical use, the annotated nature of CVPPs provides a foundation for future integration of AI and CV tools capable of automatically recognising critical anatomical views and phase transitions. As demonstrated in recent work on automated CVS detections, models trained on labelled operative footage can reliably detect safety-critical views in real time [18, 19, 21]. Within endoscopic transsphenoidal surgery specifically, recent advances in step segmentation and anatomical recognition provide the technical foundation for such systems [54, 55]. CVPP-labelled datasets could supply the phase-completion criteria missing from current models, linking visual understanding to actionable intraoperative decision support.

## Strengths & limitations

A major strength of this study was its international scope, which captured practices from multiple institutions and diverse surgical backgrounds. The expert panel, drawn from our established international research network, brought recognised expertise in pituitary and skull base surgery, ensuring that the framework reflected a broad range of operative practices and anatomical interpretations. While the panel spanned Europe and North America, the absence of low- and middle-income country representation means that our results may not fully reflect global practices. Furthermore,

the framework was intentionally designed to capture the minimum, safety-critical visual cues applicable across a wide range of cases – an approach that prioritises reproducibility and consistency, but does not attempt to capture case-specific anatomical variation (e.g., sphenoid pneumatisation patterns, skull base defects) or patient-specific surgical objectives. Consequently, experienced surgeons are unlikely to derive additional operative guidance from these cues – a limitation that reflects the intended scope of the framework.

Beyond this, some methodological limitations should also be acknowledged. First, the validation study was conducted within a single tertiary centre and involved a small cohort, which may limit the generalisability of our findings. Second, while the CVPPs formalise the visual readiness for phase progression, they do not account for dynamic elements of intraoperative decision-making (e.g., instrument handling, navigation data). Third, the nasal and sphenoid phases were combined into a single naso-sphenoid phase to facilitate future digital annotation. While this improved phase stability for data analysis, it may limit the granularity of workflow characterisations for studies focusing on the early exposure phase. Fourth, regarding the validation study, the use of short, decontextualised video clips may have contributed to isolated instances of interpretive variability. In addition, formal inter-rater reliability statistics (e.g., Cohen's  $\kappa$ ) were not calculated, as the sample size and purposive clip selection could render such metrics unstable and potentially misleading.

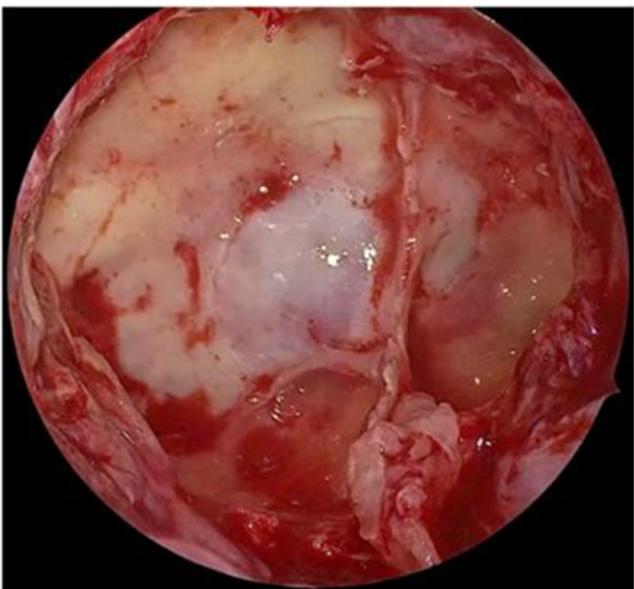
Ultimately, this framework provides a foundation for a universal phase-gate approach to intraoperative decision-making. Future work should focus on expanding the validation study across multiple centres, including pre- and post-educational intervention designs, and prospective integration of CVPPs into intraoperative workflows to assess their impact on surgical outcomes, learning curves and efficiency. As annotated datasets expand, CVPPs can serve as structured ground truth for CV models capable of real-time phase recognition, automated safety validation, and performance benchmarking.

## Conclusion

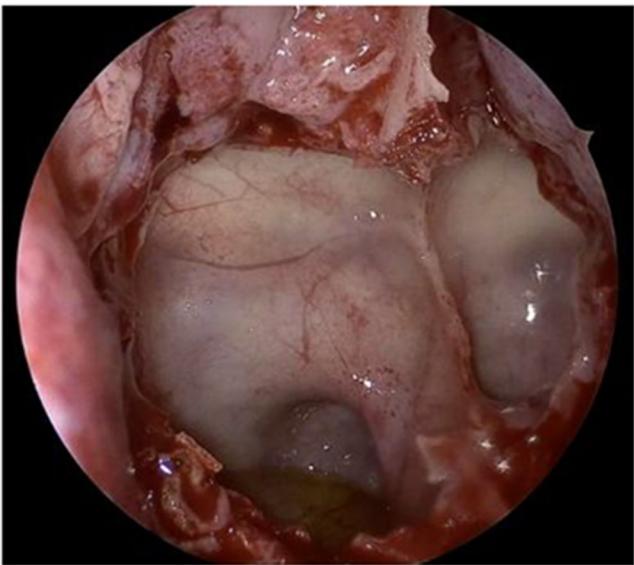
Through an international expert consensus process, this study establishes the first framework of “Critical Views for Phase Progression” and applies it to endoscopic transsphenoidal pituitary surgery. The CVPP framework bridges engineering-derived phase-gate principles with surgical safety, introducing a structured and verifiable approach to intraoperative decision-making and offering a scalable foundation for training, audit, and future digital integration.

**Appendix A: example intraoperative images illustrating a potential view at the end of each surgical phase for macro- and microadenomas**

(A) Potential intraoperative view at the end of naso-sphenoid phase - macroadenoma



(B) Potential intraoperative view at the end of naso-sphenoid phase - microadenoma



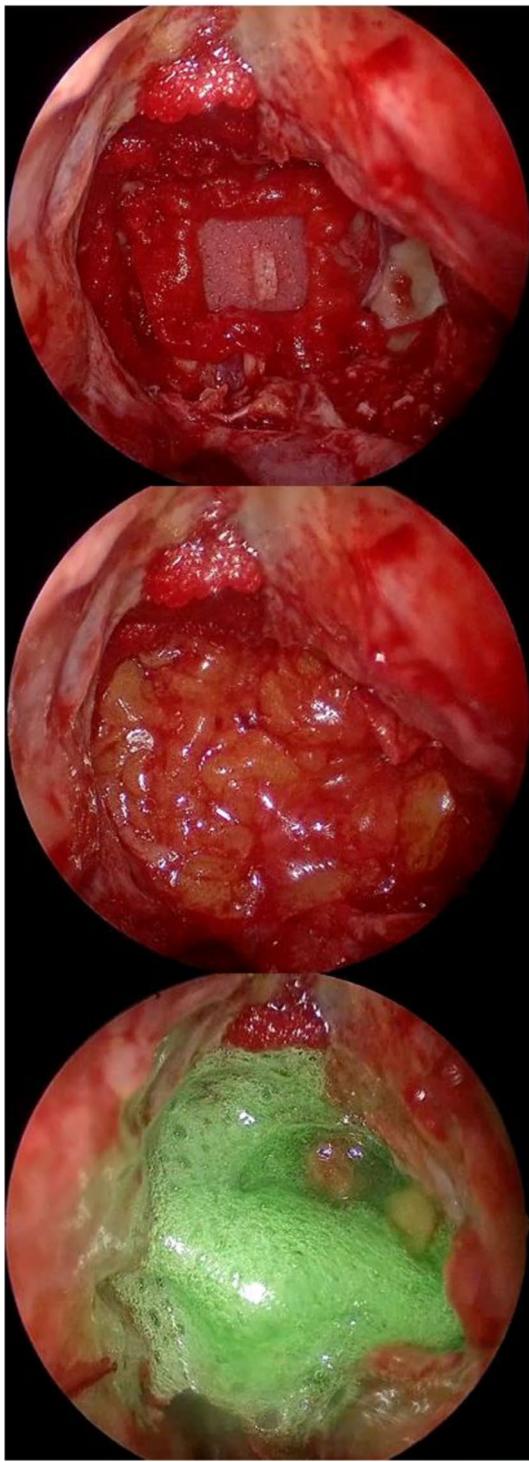
(C) Potential intraoperative view at the end of sphenoid phase - macroadenoma



(D) Potential intraoperative view at the end of sphenoid phase - microadenoma



(E) Potential intraoperative views at the end of closure phase



## Appendix B: guidance questions to experts during each consensus round

### Phase 1: Naso-sphenoid phase

*Q1:* When establishing critical view for safe progression at the end of naso-sphenoid phase, how would you classify the importance of visualising the following landmarks for safe progression to the sellar phase?

Visual Cue	Essential	Desirable	Not Necessary
Carotid protuberances	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Optic nerve prominences	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Sellar bulge or sellar floor	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Clival recess	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Tuberculum sellae	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Planum sphenoidale	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Cavernous sinus margins	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

*Q2:* Are there any additional landmarks you believe should be considered **essential** to visualise at the end of the naso-sphenoid phase?

*Q3:* Are there any additional landmarks you believe should be considered **desirable** to visualise at the end of the naso-sphenoid phase?

### Phase 2: Sellar phase

*Q4:* When establishing critical view for safe progression at the end of sellar phase, how would you classify the importance of visualising the following landmarks for safe progression to the closure phase?

Visual Cue	Essential	Desirable	Not Necessary
Normal pituitary tissue	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Pituitary stalk	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Arachnoid dome descent	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Floor of sella	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Sellar fossa clearance	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Suprasellar compartment clearance	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Medial wall of cavernous sinus	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Optic chiasm	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
CSF leak detection +/- arachnoid tear detection	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

*Q5:* Are there any additional landmarks you believe should be considered **essential** to visualise at the end of the sellar phase?

*Q6:* Are there any additional landmarks you believe should be considered **desirable** to visualise at the end of the sellar phase?

### Phase 3: Closure phase

*Q7:* When establishing critical view for safe progression at the end of closure phase, how would you classify the importance of visualising the following landmarks?

Visual Cue	Essential	Desirable	Not Necessary
Haemorrhage control	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
No active CSF leak	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Stability of repair construct	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Vascularised flap viability	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

*Q8:* Are there any additional landmarks you believe should be considered **essential** to visualise at the end of the closure phase?

*Q9:* Are there any additional landmarks you believe should be considered **desirable** to visualise at the end of the closure phase?

### General questions

*Q10:* Do you have any suggestions for improving the identification or visualization of landmarks during surgery?

### Abbreviations

AI	artificial intelligence
CSF	cerebrospinal fluid
CV	computer vision
CVPP	critical view for phase progression
CVS	critical view of safety
LC	laparoscopic cholecystectomy
ONP	optic nerve prominences
UK	United Kingdom
USA	United States of America

**Acknowledgements** HJM is supported by the NIHR Biomedical Research Centre at University College London. DZK is supported by an NIHR Doctoral Fellowship and a Google PhD Fellowship. MDC receives grant funding unrelated to this project from the Canadian Institutes of Health Research, the Public Health Agency of Canada, Physician Services Inc and Ondine Biomedical.

**Author contributions** Study conception and methodology was led by HJM, DZK, TZ, AW, ZM, PM, DS, and SB. Investigation, data curation and resources were performed by JPA, AB, JC, MC, DAD, ND, JCFM, GF, TG, ALGA, LH, RL, SBO, IPS, DP, and GZ. Material preparation, data collection and analysis were performed by TZ and DZK. The first draft of the manuscript was written by TZ, DZK,

and HJM. All authors reviewed and edited subsequent versions of the manuscript. All authors read and approved the final manuscript.

**Funding** No specific funding was received for this piece of work.

**Data availability** Available upon reasonable request.

### Declarations

**Ethical approval** Ethical approval and informed consent were unnecessary due to the nature of the study (consensus process amongst health care professionals). Participation in this study was voluntary.

**Competing interests** HJM is an employee of and has shares in Panda Surgical. DS is an employee of Digital Surgery, Medtronic and has shares in Panda Surgical, Odin Vision, EnAcuity and Helico Medical. JPA is a consultant for Stryker. MDC serves on the Pfizer Medical Advisory Board.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

### References

1. Marcus HJ, Khan DZ, Borg A, Buchfelder M, Cetas JS, Collins JW, Dorward NL, Fleseriu M, Gurnell M, Javadpour M, Jones PS, Koh CH, Layard Horsfall H, Mamelak AN, Mortini P, Muirhead W, Oyesiku NM, Schwartz TH, Sinha S, Stoyanov D, Syro LV, Tsermoulias G, Williams A, Winder MJ, Zada G, Laws ER (2021) Pituitary society expert Delphi consensus: operative workflow in endoscopic transsphenoidal pituitary adenoma resection. Pituitary 24(6):839–853
2. Alexander TD, Chitguppi C, Collopy S, Murphy K, Lavergne P, Patel P, Khouri T, Rabinowitz M, Nyquist GG, Rosen MR, Farrell CJ, Karsy M, Evans JJ (2022) Surgical outcomes of endoscopic transsphenoidal pituitary adenoma resection in elderly versus younger patients. J Neurol Surg Part B Skull Base 83(4):405–410
3. Ahn Y, Lee S, Shin DW (2024) Learning curve for endoscopic transsphenoidal surgery: a systematic review and Meta-Analysis. World Neurosurg 181:116–124
4. Namvar M, Iranmehr A, Fathi MR, Sadrhosseini SM, Tabari A, Shirzad N, Zeinalizadeh M (2022) Complications in endoscopic endonasal pituitary adenoma surgery: an institution experience in 310 patients. J Neurol Surg Part B Skull Base 84(3):255–265
5. Trimpou P, Backlund E, Ragnarsson O, Skoglund T, Hallén T, Gudnadottir G, Carlqvist J, Farahmand D (2022) Long-Term outcomes and complications from endoscopic versus microscopic transsphenoidal surgery for cushing's disease: A 15-Year Single-Center study. World Neurosurg 166:e427–e434
6. Lucas JW, Zada G (2012) Endoscopic surgery for pituitary tumors. Neurosurg Clin N Am 23(4):555–569

7. van Furth WR, de Vries F, Lobatto DJ, Kleijwegt MC, Schutte PJ, Pereira AM, Biermasz NR, Verstegen MJT (2020) Endoscopic surgery for pituitary tumors. *Endocrinol Metab Clin North Am* 49(3):487–503
8. Khan DZ, Koh CH, Das A, Valetopolou A, Hanrahan JG, Horsfall HL, Baldeweg SE, Bano S, Borg A, Dorward NL, Olukoya O, Stoyanov D, Marcus HJ (2024) Video-Based performance analysis in pituitary Surgery—Part 1: surgical outcomes. *World Neurosurg* 190:e787–e796
9. Strasberg SM (2017) A perspective on the critical view of safety in laparoscopic cholecystectomy. *Ann Laparosc Endosc Surg* [Internet]. May 10 [cited 2025 Feb 21];2(5). Available from: <http://ales.amegroups.org/article/view/3940>
10. Antonopoulou MI, Manatakis DK (2022) Critical view of safety in laparoscopic cholecystectomy: A word of caution in cases of aberrant anatomy. *Surg J* 8(3):e157–e161
11. Montalvo-Javé EE, Contreras-Flores EH, Ayala-Moreno EA, Mercado MA (2022) Strasberg's critical view: strategy for a safe laparoscopic cholecystectomy. *Euroasian J Hepato-Gastroenterol* 12(1):40
12. Dermody SM, Shuman AG (2022) Defining critical portions of surgery. *Ann Surg* 276(2):213–214
13. Pucher PH, Brunt LM, Fanelli RD, Asbun HJ, Aggarwal R (2015) SAGES expert Delphi consensus: critical factors for safe surgical practice in laparoscopic cholecystectomy. *Surg Endosc* 29(11):3074–3085
14. The SAGES Cholecystectomy Program - Strategies for Minimizing Bile Duct Injuries [Internet]. SAGES. [cited 2025 Mar 6]. Available from: <https://www.sages.org/safe-cholecystectomy-program/>
15. Grütter AAJ, Daams F, Bonjer HJ, van Duijvendijk P, Tuynman JB, Jilesen A, Blomberg B, Berndsen B, de Betue C, Henneman D, Sloothaak D, Wassenaar E, Bruns E, Westerduin E, van Nieuwenhoven EJ, Frans F, Hoexum F, Prakken F, Musters G, Jalalzadeh H, Ebben H, Palamba HW, Atema J, Posthuma J, Dikken J, Govaert J, van Hilst J, ten Brinke J, Volders J, de Leur K, Govaert K, van der Werf L, Busweiler L, Goessens M, -van Bolster M, van Dorp M, Ayez N, Karthaus N, Moerbeek P, Johannesma P, Coelen RJ, Blok R, Bakx R, Mekke S, Gans S, van Dijk S, van der Ploeg S, Poels T, Coblijn U, Alberts V, Weeda V, Alberda W, Lastdrager W, Issa Y (2024) Video-based surgical quality assessment collaborators. surgical quality assessment of critical view of safety in 283 laparoscopic cholecystectomy videos by surgical residents and surgeons. *Surg Endosc* 38(7):3609–14
16. Manatakis DK, Papageorgiou D, Antonopoulou MI, Stamos N, Agalianos C, Ivros N, Davides D, Pechlivanides G, Kyriazanos I (2019) Ten-year audit of safe Bail-Out alternatives to the critical view of safety in laparoscopic cholecystectomy. *World J Surg* 43(11):2728–2733
17. Onoe S, Maeda A, Takayama Y, Fukami Y, Kaneoka Y (2017) A preoperative predictive scoring system to predict the ability to achieve the critical view of safety during laparoscopic cholecystectomy for acute cholecystitis. *HPB* 19(5):406–410
18. Leifman G, Golany T, Rivlin E, Khouri W, Assalia A, Reissman P (2024) Real-time artificial intelligence validation of critical view of safety in laparoscopic cholecystectomy. *Intell-Based Med* 10:100153
19. Petracchi EJ, Olivieri SE, Varela J, Canullan CM, Zandalazini H, Ocampo C, Quesada BM (2024) Use of artificial intelligence in the detection of the critical view of safety during laparoscopic cholecystectomy. *J Gastrointest Surg Off J Soc Surg Aliment Tract* 28(6):877–879
20. Ríos MS, Molina-Rodríguez MA, Londoño D, Guillén CA, Sierra S, Zapata F, Giraldo LF (2023) Cholec80-CVS: an open dataset with an evaluation of strasberg's critical view of safety for AI. *Sci Data* 10(1):194
21. Mascagni P, Vardazaryan A, Alapatt D, Urade T, Emre T, Fiorillo C, Pessaux P, Mutter D, Marescaux J, Costamagna G, Dallemande B, Padov N (2022) Artificial intelligence for surgical safety: automatic assessment of the critical view of safety in laparoscopic cholecystectomy using deep learning. *Ann Surg* 275(5):955–961
22. Petersen K, Wohlin C, Baca D (2009) The Waterfall Model in Large-Scale Development. In: Bomarius F, Oivo M, Jaring P, Abrahamsson P, editors. *Product-Focused Software Process Improvement* [Internet]. Berlin, Heidelberg: Springer Berlin Heidelberg; [cited 2025 Feb 21]. pp. 386–400. (Lecture Notes in Business Information Processing; vol. 32). Available from: [http://link.springer.com/https://doi.org/10.1007/978-3-642-0215-2\\_7\\_29](http://link.springer.com/https://doi.org/10.1007/978-3-642-0215-2_7_29)
23. Thesing T, Feldmann C, Burchardt M (2021) Agile versus waterfall project management: decision model for selecting the appropriate approach to a project. *Procedia Comput Sci* 181:746–756
24. (PDF) Open innovation and the stage-gate process: a revised model for new product development. ResearchGate [Internet] (2024) Oct 22 [cited 2025 Feb 21]; Available from: [https://www.researchgate.net/publication/259729298\\_Open\\_Innovation\\_and\\_the\\_Stage-Gate\\_Process\\_A\\_Revised\\_Model\\_for\\_New\\_Product\\_Development](https://www.researchgate.net/publication/259729298_Open_Innovation_and_the_Stage-Gate_Process_A_Revised_Model_for_New_Product_Development)
25. Cooper RG (1990) Stage-gate systems: a new tool for managing new products. *Bus Horiz* 33(3):44–54
26. Okoli C, Pawlowski SD (2004) The Delphi method as a research tool: an example, design considerations and applications. *Inf Manage* 42(1):15–29
27. Consortium CRANIAL, Khan DZ, Marcus HJ, Bandyopadhyay S, Schroeder BE, Patel V, O'Donnell A, Neurology, Group NI, Collaborative BNTR, Giamouriadis A, Bhatt P, Ram B, Varma A, Weir P, Hanna B, Hirst TC, McAlevey P, Paluzzi A, Tsermoulias G, Ahmed S, Soon WC, Chowdhury YA, Abualsaud S, Mahmood S, Naik P, Haiderkhan Z, Al-Mahfoudh R, Perera A, Rus M, Williams A, Hand C, Abhinav K, Cernei C, Dilnawaz A, Mannion R, Santarius T, Tysome J, Sharma R, Koliag AG, Donnelly N, Patel V, Venkatesh A, Hayhurst C, Mohamed A, Stew B, Merola J, Ooi SZY, Kamei M, Khan MH, Abrar S, McKeon C, McSweeney D, Javadpour M, Lacy P, Murray D, Roman E, Hossain-Ibrahim K, Ross P, Bennett D, McSorley N, Hounat A, Statham P, Hughes M, Hamdan A, Scott C, Moudgil-Joshi J, Bahl A, Bjornson A, Gatt D, Phillips N, Kalra N, Bautista M, Shirazi S, Gilkes CE, Millward CP, Ali AM, Paraskevopoulos D, Bal J, Matloob S, Lobo R, Mendoza N, Nair R, Dalton A, Nadig A, Hernandez L, Thomas N, Maratos E, Shapey J, Al-Barazi S, Mirza AB, Okasha M, Malhotra PS, Ahmed R, Dorward NL, Grieve J, Sayal P, Choi D, Cabriol I, Horsfall HL, Pollock J, Shoakazemi A, MacCormac O, Amirthalingam GNK, Martin A, Stapleton S, Hogg F, Richardson D, Gnanalingham K, Pathmanabhan O, Fountain DM, Bhalla R, Hannan CJ, Chadwick A, Jenkins A, Nicholson C, Shumon S, Youssef M, Allison C, Dow G, Robertson I, Glancz LJ, Sitaraman M, Kumaria A, Bagchi A, Cudlip S, Halliday J, Piper RJ, Boukas A, Amrouche M, Veljanoski D, Muquit S, Edlmann E, Maripi H, Wang Y, Hossain M, Alalade A, Maroof S, Patkar P, Sinha S, Mirza S, Henderson D, Khan MS, Mathad N, Hempenstall J, Neurosurgery DW, Marwaha P, Shaw S, Solomou G, Shrestha A, Fraser A, Hirst T, Chowdhury Y, Bilal S, Wildman J, Venkatesh A, Babu P, Carey C, Bathuni RR, Hossain-Ibrahim K, Brennan JN, Bjornson A, Ktayen H, Trichinopoly ST, Matloob S, Nadig A, Okasha M, Khan D, Shoakazemi A, Hogg F, Sobawale S, Suliman A, Kumaria A, Piper R, Owen W, Edlmann E, Sartaj A, Goacher E, Strachan E, Solomou G CSF rhinorrhoea after endonasal intervention to the skull base (CRANIAL): a multi-centre prospective observational study. *Front Oncol* [Internet]. 2023 Jan 4 [cited 2025 Mar 6];12. Available from: <https://doi.org/10.3389/fonc.2022.1049627>

28. Barimani B, Ahangar P, Nandra R, Porter K (2020) The WHO surgical safety checklist: A review of outcomes and implementation strategies. *Perioper Care Oper Room Manag* 21:100117

29. Lecuyer G, Ragot M, Martin N, Launay L, Jannin P (2020) Assisted phase and step annotation for surgical videos. *Int J Comput Assist Radiol Surg* 15(4):673–680

30. Lalys F, Jannin P (2014) Surgical process modelling: a review. *Int J Comput Assist Radiol Surg* 9(3):495–511

31. Collins JW, Levy J, Stefanidis D, Gallagher A, Coleman M, Cecil T, Ericsson A, Mottrie A, Wiklund P, Ahmed K, Pratschke J, Casali G, Ghazi A, Gomez M, Hung A, Arnold A, Dunning J, Martino M, Vaz C, Friedman E, Baste JM, Bergamaschi R, Feins R, Earle D, Pusic M, Montgomery O, Pugh C, Satava RM (2019) Utilising the Delphi process to develop a Proficiency-based progression Train-the-trainer course for robotic surgery training. *Eur Urol* 75(5):775–785

32. Mascagni P, Alapatt D, Murali A, Vardazaryan A, Garcia A, Okamoto N, Costamagna G, Mutter D, Marescaux J, Dallemande B, Padov N (2025) Endoscapes, a critical view of safety and surgical scene segmentation dataset for laparoscopic cholecystectomy. *Sci Data* 12(1):331

33. The SAGES Cholecystectomy Program - Strategies for Minimizing Bile Duct Injuries [Internet]. SAGES. [cited 2025 Nov 4]. Available from: <https://www.sages.org/safe-cholecystectomy-program/>

34. Strasberg SM (2019) A three-step conceptual roadmap for avoiding bile duct injury in laparoscopic cholecystectomy: an invited perspective review. *J Hepato-Biliary-Pancreat Sci* 26(4):123–127

35. Terho P, Sallinen V, Lampela H, Harju J, Koskenvuo L, Mennula P (2021) The critical view of safety and bile duct injuries in laparoscopic cholecystectomy: a photo evaluation study on 1532 patients. *HPB* 23(12):1824–1829

36. Wong HJ, Kojima Y, Su B, Attaar M, Wu H, Campbell M, Kuchta K, Linn JG, Denham W, Haggerty SP, Ujiki MB (2022) Long-term retention after structured curriculum on attainment of critical view of safety during laparoscopic cholecystectomy for surgeons. *Surgery* 171(3):577–583

37. Quarto G, Fornicola A (2025) Using artificial intelligence to enable safer laparoscopic cholecystectomy: A step toward a standardized critical view of safety and reduced bile duct injuries. *Ann Ital Chir* 96.

38. Athanasiadis DJ, Makhecha K, Blundell N, Mizota T, Anderson-Montoya B, Fanelli RD, Scholz S, Vazquez R, Gill S, Stefanidis D (2025) How accurate are surgeons at assessing the quality of their critical view of safety during laparoscopic cholecystectomy? *J Surg Res* 305:36–40

39. Conrad E, Misener S, Feldman J (2010) Chapter 9 - Domain 8: Application development security. In: Conrad E, Misener S, Feldman J, editors. *CISSP Study Guide* [Internet]. Boston: Syngress; [cited 2025 Nov 4]. pp. 329–69. Available from: <https://www.sciencedirect.com/science/article/pii/B9781597495639000093>

40. Jazayeri M (2003) Software Engineering. In: Meyers RA, editor. *Encyclopedia of Physical Science and Technology* (Third Edition) [Internet]. New York: Academic Press; pp. 1–14. Available from: <https://www.sciencedirect.com/science/article/pii/B0122274105008565>

41. Carayon P, Hundt AS, Alvarado CJ, Springman S, Borgsdorf A, Jenkins L Implementing a Systems Engineering Intervention for Improving Safety in Outpatient Surgeries. In: Henriksen K, Battles JB, Marks ES, Lewin DI, editors. *Advances in Patient Safety: From Research to Implementation (Volume 3: Implementation Issues)* [Internet]. Rockville (MD): Agency for Healthcare Research and Quality (US); 2005 [cited 2025 Nov 4]. (Advances in Patient Safety). Available from: <http://www.ncbi.nlm.nih.gov/books/NBK20560/>

42. Karsh B, Holden RJ, Alper SJ, Or CKL (2006) A human factors engineering paradigm for patient safety: designing to support the performance of the healthcare professional. *Qual Saf Health Care* 15(Suppl 1):i59–65

43. Clarkson J, Bogle D, Dean J, Tooley M, Trewby J, Vaughan L, Adams E, Dudgeon P (2017 Aug) Nicola Platt, Philippa Shelton. Engineering better care a systems approach to health and care design and continuous improvement. Royal Academy of Engineering

44. Gurses AP, Ozok AA, Pronovost PJ (2012) Time to accelerate integration of human factors and ergonomics in patient safety. *BMJ Qual Saf* 21(4):347–351

45. Demir KC, Schieber H, Weise T, Roth D, May M, Maier A, Yang SH (2023) Deep learning in surgical workflow analysis: A review of phase and step recognition. *IEEE J Biomed Health Inf* 27(11):5405–5417

46. Mezzina M, De Backer P, Vercauteren T, Blaschko M, Mottrie A, Tuytelaars T (2025) Surgeons versus computer vision: a comparative analysis on surgical phase recognition capabilities. *Int J Comput Assist Radiol Surg* 20(6):1283–1291

47. Garrow CR, Kowalewski KF, Li L, Wagner M, Schmidt MW, Engelhardt S, Hashimoto DA, Kennett HG, Bodenstedt S, Speidel S, Müller-Stich BP, Nickel F (2021) Machine learning for surgical phase recognition: A systematic review. *Ann Surg* 273(4):684

48. Liao W, Zhu Y, Zhang H, Wang D, Zhang L, Chen T, Zhou R, Ye Z Artificial intelligence-assisted phase recognition and skill assessment in laparoscopic surgery: a systematic review. *Front Surg* [Internet]. 2025 Apr 11 [cited 2025 Nov 4];12. Available from: <https://www.frontiersin.org/journals/surgery/articles/https://doi.org/10.3389/fsurg.2025.1551838/full>

49. Li P, Shu X, Feng CM, Feng Y, Zuo W, Tang J (2025) Surgical video workflow analysis via visual-language learning. *Npj Health Syst* 2(1):5

50. Khan DZ, Luengo I, Barbarisi S, Addis C, Culshaw L, Dorward NL, Haikka P, Jain A, Kerr K, Koh CH, Horsfall HL, Muirhead W, Palmisciano P, Vasey B, Stoyanov D, Marcus HJ (2021) Automated operative workflow analysis of endoscopic pituitary surgery using machine learning: development and preclinical evaluation (IDEAL stage 0). *J Neurosurg* 137(1):51–58

51. Das A, Khan DZ, Hanrahan JG, Marcus HJ, Stoyanov D (2023) Automatic generation of operation notes in endoscopic pituitary surgery videos using workflow recognition. *Intell-Based Med* 8:100107

52. Das A, Khan DZ, Williams SC, Hanrahan JG, Borg A, Dorward NL, Bano S, Marcus HJ, Stoyanov D (2023) A Multi-task Network for Anatomy Identification in Endoscopic Pituitary Surgery. In: Greenspan H, Madabhushi A, Mousavi P, Salcudean S, Duncan J, Syeda-Mahmood T, Taylor R, editors. *Medical Image Computing and Computer Assisted Intervention – MICCAI 2023* [Internet]. Cham: Springer Nature Switzerland; [cited 2025 Dec 26]. pp. 472–82. (Lecture Notes in Computer Science; vol. 14228). Available from: [https://doi.org/10.1007/978-3-031-43996-4\\_45](https://doi.org/10.1007/978-3-031-43996-4_45)

53. Nakazato T, Su B, Novak S, Deal SB, Kuchta K, Ujiki M (2020) Improving attainment of the critical view of safety during laparoscopic cholecystectomy. *Surg Endosc* 34(9):4115–23

54. Khan DZ, Valetopoulou A, Das A, Hanrahan JG, Williams SC, Bano S, Borg A, Dorward NL, Barbarisi S, Culshaw L, Kerr K, Luengo I, Stoyanov D, Marcus HJ (2024) Artificial intelligence assisted operative anatomy recognition in endoscopic pituitary surgery. *NPJ Digit Med* 7:314

55. Mao Z, Das A, Islam M, Khan DZ, Williams SC, Hanrahan JG, Borg A, Dorward NL, Clarkson MJ, Stoyanov D, Marcus HJ, Bano S (2024) PitSurgRT: real-time localization of critical anatomical structures in endoscopic pituitary surgery. *Int J Comput Assist Radiol Surg* 19(6):1053–1060

## Authors and Affiliations

Tjasa Zaletel<sup>1,2</sup> · Danyal Z. Khan<sup>2,3</sup> · Anjana Wijekoon<sup>3,4</sup> · Zhehua Mao<sup>3,4</sup> · Joao Paulo Almeida<sup>5</sup> · Anouk Borg<sup>2</sup> · Jonathan Chainey<sup>6</sup> · Michael D. Cusimano<sup>7</sup> · Daniel A. Donoho<sup>8,9</sup> · Neil Dorward<sup>2</sup> · Juan Carlos Fernandez-Miranda<sup>10</sup> · Giorgio Fiore<sup>11</sup> · Theofanis Giannis<sup>2</sup> · Alfonso Lagares Gomez-Abascal<sup>12,13,14</sup> · Lauren Harris<sup>2</sup> · Abhiney Jain<sup>2,3</sup> · Ruth Lau<sup>15</sup> · Sacit B. Omay<sup>16</sup> · Igor Paredes<sup>12,13,14</sup> · Daniel Prevedello<sup>17</sup> · Gabriel Zada<sup>18</sup> · Danail Stoyanov<sup>3,4</sup> · Sophia Bano<sup>3,4</sup> · Hani J. Marcus<sup>2,3</sup>

✉ Danyal Z. Khan  
d.khan@ucl.ac.uk

<sup>1</sup> Department of Clinical and Movement Neurosciences, University College London, London, UK

<sup>2</sup> Department of Neurosurgery, National Hospital for Neurology and Neurosurgery, London, UK

<sup>3</sup> The UCL Hawkes Institute, University College London, London, UK

<sup>4</sup> Department of Computer Science, University College London, London, UK

<sup>5</sup> Department of Neurosurgery, IU Health, Indiana University, Indianapolis, IN, USA

<sup>6</sup> Division de neurochirurgie, Département de chirurgie, Université de Montréal, Montréal, Canada

<sup>7</sup> Division of Neurosurgery, Department of Surgery, University of Toronto, Toronto, ON, Canada

<sup>8</sup> Division of Neurosurgery, Children's National Hospital, Washington, DC, USA

<sup>9</sup> Department of Neurosurgery, George Washington University School of Medicine & Health Sciences, Washington, DC, USA

<sup>10</sup> Stanford University School of Medicine, 213 Quarry Road, Palo Alto, USA

<sup>11</sup> Unit of Neurosurgery, Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico, Milan, Italy

<sup>12</sup> Instituto de Investigación Sanitaria Hospital 12 de Octubre (imas12), Madrid, Spain

<sup>13</sup> Neurosurgery Department, Hospital Universitario 12 de Octubre, Madrid, Spain

<sup>14</sup> Department of Surgery, Faculty of Medicine, Universidad Complutense de Madrid, Madrid, Spain

<sup>15</sup> Department of Neurosurgery, Hospital Universitari Joan XXIII, Universitat Rovira i Virgili, Tarragona, Spain

<sup>16</sup> Department of Neurosurgery, Yale School of Medicine, New Haven, Connecticut, USA

<sup>17</sup> Department of Neurosurgery, The Ohio State University and Wexner Medical Centre, Columbus, OH, USA

<sup>18</sup> Department of Neurological Surgery, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA