

ORIGINAL RESEARCH

Open access

Temporal Neural Network for Forecasting Operating Room Procedure Duration Using Surgical Procedure Codes, Surgeon-Specific Historical Performance, Anesthesia Records, Patient Comorbidity Profiles, and Real-Time Intraoperative Event Logs

Ahmed Mansour^{1*}, Omar Saeed¹

Abstract

Operating room procedure duration is central to scheduling efficiency, resource utilisation, staff coordination, and perioperative safety. Inaccurate forecasts can create idle capacity, overtime, delayed starts, cancellations, and avoidable strain across surgical services. Conventional estimates often rely on surgeon judgment, historical averages, or static regression models. These approaches do not fully account for evolving intraoperative conditions or the temporal structure of surgical progress once a case has begun. This article develops a conceptual temporal neural network for forecasting procedure duration before incision and updating the expected completion time during surgery. The model is designed to combine static case information with sequential intraoperative event logs. The proposed architecture uses procedure codes, surgeon-specific historical performance, anesthesia records, patient comorbidity profiles, and timestamped operative events. A recurrent neural network based on LSTM or GRU principles would update a remaining-time distribution as new events occur. Conceptually, the model could provide an initial duration estimate and then refine that estimate as the case progresses. For example, it would be expected to revise completion time upward when a laparoscopic case converts to an open procedure or when unexpected bleeding is recorded. A temporal duration-forecasting model could improve operating room coordination, reduce avoidable waiting, and support more responsive perioperative decision-making. Its value would depend on careful validation, workflow integration, and transparent communication to clinical teams.

Keywords Operating room forecasting, Procedure duration, Temporal neural network, Perioperative informatics, Surgical workflow, Anesthesia records

*Correspondence:

Ahmed Mansour
ahmed.mansour@gmail.com

¹ Department of Digital Health Systems and Informatics, Faculty of Medicine, Cairo University, Cairo, Egypt

Introduction

Operating room time is a high-cost, clinically critical resource because each scheduled case depends on coordinated availability of surgeons, anesthesiologists, nurses, equipment, recovery beds, and downstream inpatient capacity. Inaccurate procedure-duration estimates can propagate through the surgical day, contributing to overtime, delayed starts, underused block time, avoidable cancellations, and staff fatigue. Prior work on operating room efficiency and surgical cost

has shown that case-duration uncertainty is not merely a prediction problem but a system-level constraint on throughput and resource allocation [1-3]. For this reason, duration forecasting should be treated as a perioperative informatics problem embedded in operational decision-making rather than as a standalone statistical exercise.

Traditional duration estimates often rely on surgeon-entered values, historical averages, or linear modelling strategies that assume past case lengths can be summarised adequately before the procedure begins. Regression-based approaches have improved on simple averages, but they remain limited when procedure codes, patient condition, surgeon pace, and intraoperative deviations interact in nonlinear ways [4-6]. Scheduling methods that optimise staffing or block allocation may still inherit forecast error when the underlying case-length estimate is static and insensitive to the unfolding surgical course [7]. These limitations motivate models that can begin with preoperative information yet remain responsive after incision.

To clarify the conceptual distinction between traditional static estimation methods and the proposed dynamic framework, the structural comparison is presented in **Table 1**.

Table 1. Structural Comparison between Static and Temporal Approaches to Operating Room Duration Forecasting

Dimension	Conventional Static Models	Proposed Temporal Neural Network
Prediction Timing	Preoperative only	Preoperative + continuous intraoperative updates
Data Utilisation	Aggregated historical averages or regression inputs	Integrated static features + sequential intraoperative event streams
Representation of Surgical Process	Fixed, non-evolving estimate	Dynamic, time-evolving representation of surgical progression
Handling of Unexpected Events	Not explicitly incorporated	Explicitly integrated as high-impact sequential signals
Model Structure	Linear regression or static ML models	Recurrent architecture (LSTM/GRU) with hidden state updates
Personalisation	Limited or indirect	Explicit via surgeon-specific embeddings and historical patterns
Uncertainty Quantification	Often absent or limited	Probabilistic output with prediction intervals
Responsiveness During Surgery	None	Event-driven real-time forecast revision
Workflow Integration	Scheduling-focused only	Integrated preoperative and intraoperative decision support
Interpretability Context	Retrospective or static	Context-aware explanations linked to intraoperative events

Machine learning has increasingly been applied to surgical case-duration prediction, including models that use clinical and nonclinical features to estimate operative time or total operating room use. Studies in surgical time prediction, robotic surgery, arthroplasty, and general operating room utilisation show that machine learning could capture richer feature interactions than rule-based or linear methods [8-11]. However, many of these approaches remain primarily preoperative or static, meaning that they do not fully exploit the information contained in intraoperative milestones, anesthesia timing,

or evolving workflow events. A model for live duration forecasting should therefore integrate both baseline case features and the temporal sequence generated during the operation itself.

The central thesis of this Models/Deep Learning article is that a temporal neural network could forecast operating room procedure duration more adaptively by fusing static features with real-time intraoperative event streams. The static component would encode procedure codes, surgeon history, anesthesia plan, and comorbidity profile, while the temporal component would revise the forecast as operative milestones or adverse events are logged. Deep surgical workflow models, including recurrent, temporal convolutional, and attention-based architectures, demonstrate that surgical processes can be represented as ordered event or phase sequences rather than as fixed preoperative attributes alone. Such a framework would be expected to support both initial scheduling and intraoperative coordination through continuously updated estimates of remaining time.

Background

OR scheduling and the importance of duration prediction

Operating room scheduling requires matching uncertain surgical demand with constrained personnel, rooms, instruments, anesthesia resources, recovery capacity, and turnover processes. Poor duration prediction can create both under-utilisation, in which expensive rooms sit idle, and over-utilisation, in which later cases are delayed or pushed into overtime. Operations-oriented studies of workload planning and second-shift staffing indicate that schedule quality depends strongly on credible estimates of future case duration and day-of-surgery progression [2, 7]. Machine learning approaches to case-duration prediction and daily caseload forecasting therefore address a core operational problem rather than a narrow technical task [1, 12].

Surgery coding and case complexity

Surgical procedure codes provide a structured description of the planned intervention and can serve as a compact representation of expected operative work. Yet codes alone cannot capture anatomical difficulty, adhesions from prior operations, surgeon technique, equipment availability, or unexpected intraoperative findings. Case-duration models that use procedure descriptors have shown that coded procedural information can be valuable, but it must be combined with other perioperative variables to represent individual case complexity more fully [4, 8, 13]. Hierarchical grouping of rare procedures would therefore be important for a temporal neural network because the same coded procedure may have substantially different duration trajectories in different clinical contexts.

Surgeon-specific performance and anesthesia factors

Individual surgeon pace, case selection, team familiarity, and procedural learning can influence operative time even when nominal procedure type is similar. Prior models of surgical duration and robot-assisted surgery suggest that case length is shaped by provider- and procedure-specific patterns that are not reducible to diagnosis or operation code alone [5, 6]. Anesthesia information also matters because induction, airway management, maintenance, emergence, and medication-related events affect total room occupation and may interact with surgical complexity [14, 15]. A duration model should therefore distinguish pure operative time from broader procedure-related and anesthesia-related intervals when generating forecasts for scheduling decisions.

Patient comorbidity as a duration modifier

Patient comorbidity can modify procedure duration by increasing technical difficulty, anesthesia complexity, positioning requirements, vascular access needs, or the likelihood of intraoperative instability. Models using preoperative patient factors for procedures such as total knee arthroplasty illustrate how patient characteristics can inform expectations about operative time and perioperative resource use [11]. General duration-prediction studies similarly support the inclusion of demographic, clinical, and procedural variables as joint predictors rather than treating the surgical procedure as the only

determinant of time [8, 9]. For a temporal neural network, comorbidity profiles would provide the initial risk context against which later intraoperative events are interpreted.

Real-time intraoperative events and dynamic forecasting

Intraoperative event logs can encode milestones such as incision, laparoscopic access, specimen removal, closure, conversion to open surgery, bleeding events, device problems, or delayed counts. Surgical phase-recognition literature demonstrates that operative workflows contain temporal structure that can be learned from sequential data, including video-derived phases, workflow states, and action transitions [16-19]. Although event logs are less visually rich than surgical video, they are more directly available in many perioperative information systems and can support live updating of remaining-time predictions. A dynamic forecasting model should therefore treat each logged event as evidence that revises the expected completion trajectory rather than as a retrospective annotation only.

Model Development Overview

High-level predictive pipeline

The proposed predictive pipeline begins before the procedure with static inputs that generate a baseline duration forecast for scheduling and preparation. As the case proceeds, timestamped intraoperative events are sequentially passed into the temporal component, which updates the hidden state and revises the remaining-time distribution. This design follows the broader shift from static surgical prediction toward models that represent operations as temporally ordered processes, as seen in recurrent and temporal convolutional workflow-recognition systems [17, 18, 20]. The intended output is not a definitive completion time but a continuously revisable forecast suitable for decision support.

The overall hierarchical structure of the proposed temporal neural network, including static feature encoding, sequential event integration, and dynamic forecast updating, is illustrated in **Figure 1**.

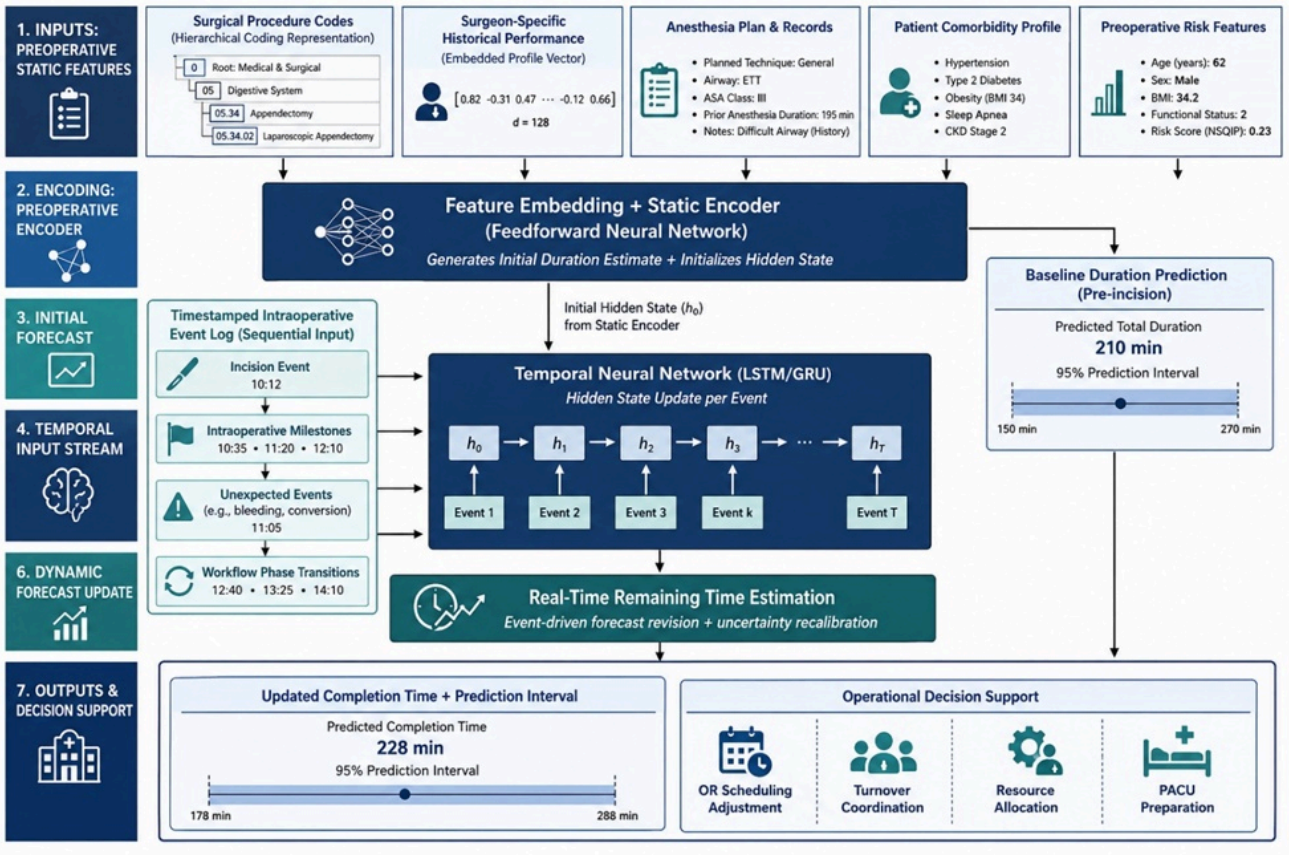


Figure 1. Hierarchical Architecture of a Temporal Neural Network for Dynamic Operating Room Procedure Duration Forecasting

Core inputs (static and dynamic)

Static inputs would include procedure codes, surgeon identifier, surgeon-specific historical speed metrics, anesthesia type, induction-related features, patient comorbidity vector, and relevant preoperative risk descriptors. Dynamic inputs would include timestamped intraoperative events such as instrument count, laparoscopic access, unexpected bleeding, specimen removal, conversion to open procedure, and closure-related milestones. Machine learning studies of operating room duration show the value of structured preoperative and perioperative variables, while workflow models show that sequential surgical information can be represented as an evolving signal [1, 8, 21, 22]. Combining these sources would allow the model to express both expected case complexity and observed progression.

Design principles

The model should be real-time updatable, robust to missing or delayed intraoperative entries, personalised to surgeon-patient-procedure combinations, and calibrated to express uncertainty rather than only a point estimate. It should also separate preoperative scheduling utility from intraoperative control-desk utility, because the same forecast may be used differently before room entry, after incision, and near closure. Literature on surgical artificial intelligence emphasises that model outputs must be clinically interpretable, workflow-aware, and validated in realistic operating room conditions before use in decision support [23-25]. Accordingly, this architecture is framed as a conceptual model that should be evaluated prospectively rather than as a completed performance claim.

Data Sources and Feature Engineering

Preoperative static features

Preoperative features would be extracted from scheduling systems, surgical coding records, anesthesia documentation, and pre-anesthesia assessment. Procedure codes could be mapped to hierarchical categories, surgeon identifiers linked to moving historical summaries for similar cases, anesthesia plans encoded by airway and technique, and comorbidity measures represented through structured clinical profiles. Prior case-duration studies indicate that surgical time prediction benefits from combining procedure, provider, patient, and operational variables rather than relying on any single category alone [4, 5, 9, 11]. These features would form the baseline context from which the temporal model begins forecasting before intraoperative events are available.

The integration of heterogeneous perioperative data sources and their distinct functional roles within the temporal model are systematically outlined in **Table 2**.

Table 2. Integrated Feature Domains and Their Functional Roles in Temporal Operating Room Duration Forecasting

Feature Domain	Data Source	Representation Strategy	Functional Role in Model	Temporal Interaction
Procedure Codes	Surgical scheduling systems	Hierarchical categorical embeddings	Encodes baseline procedural complexity	Provides initial context for all subsequent updates
Surgeon-Specific Performance	Historical case databases	Learned embedding vectors with moving averages	Captures individual operative pace and variability	Modulates interpretation of intraoperative progress
Anesthesia Records	Anesthesia information systems	Structured categorical + continuous variables	Represents perioperative workflow and physiological constraints	Influences both baseline estimate and late-stage progression
Patient Comorbidity Profile	Preoperative assessment records	Multi-dimensional clinical feature vector	Defines baseline risk and procedural difficulty	Conditions response to intraoperative deviations
Intraoperative Event Logs	Real-time OR documentation systems	Sequential timestamped event encoding	Provides dynamic evidence of surgical progression	Drives hidden state updates and forecast revision
Temporal Features (Elapsed Time)	Derived from timestamps	Continuous normalized variables	Tracks progression relative to expected workflow	Enables detection of delays or accelerations
Rare Event Signals	Event log anomalies	Sparse high-weight encoded features	Identifies deviations (e.g., bleeding, conversion)	Expands uncertainty and shifts duration trajectory

Intraoperative event log representation

Intraoperative event logs would be parsed as a sequence of discrete event types with associated timestamps aligned to the operating room clock. Each event would be represented by its type, elapsed time since case start, elapsed time since the previous event, and optionally its relationship to expected surgical phase. Work on surgical workflow recognition has shown that temporal ordering, phase transitions, and milestone progression are central to understanding operative

processes, whether learned from video or structured activity streams [17, 19, 22, 26]. In this model, event logs would provide the online evidence needed to update the remaining-duration forecast as the procedure unfolds.

Feature alignment and normalisation

Feature alignment would require synchronising preoperative scheduling data, anesthesia records, patient assessments, and intraoperative event timestamps into a single case timeline. Categorical variables such as procedure group, surgeon, anesthesia type, and event type could be encoded as embeddings, while continuous variables such as elapsed time, body mass index, and historical duration summaries would be scaled consistently. Rare procedures could be grouped through hierarchical coding structures to reduce sparsity while preserving clinically meaningful distinctions [10, 13]. This approach would support generalisation across common and uncommon operations without treating every procedure code as an unrelated category.

Temporal Neural Network Architecture

Preoperative encoding and initial forecast

The preoperative encoder would combine learned embeddings for categorical variables with continuous clinical and operational features in a feed-forward neural component. This component would generate an initial duration estimate and initialise the hidden state used by the recurrent portion of the model. Artificial neural network studies of surgical and anesthesia duration prediction provide a conceptual basis for nonlinear preoperative encoding, while operating room machine learning studies support the use of multi-source structured predictors [6, 14, 27]. The initial forecast would serve as the model's best estimate before direct evidence of intraoperative progress is available.

Recurrent core for sequential event integration

The recurrent core would use LSTM or GRU logic to ingest intraoperative events in temporal order, updating the hidden representation each time a new event is logged. This design would allow earlier events to shape the interpretation of later milestones, which is important because the meaning of a delay depends on where the case is in its operative trajectory. Recurrent convolutional and temporal surgical workflow models demonstrate that surgical processes can be learned as ordered sequences with evolving state, while attention-regularised transformer approaches suggest that critical events may be weighted differently depending on context [17, 20, 28]. In the proposed model, the recurrent state would translate event history into a revised estimate of time remaining.

Output: time-remaining distribution and prediction intervals

The output layer would produce a distribution of remaining time rather than only a single predicted endpoint, allowing the forecast to communicate uncertainty to scheduling teams. This could be implemented conceptually through quantile-based outputs or a parametric duration distribution, enabling prediction intervals for downstream operational decisions. Probabilistic forecasting work in surgical duration prediction supports the importance of uncertainty-aware estimates, while continuous real-time prediction research indicates that forecasts can be revised as the procedure advances [8, 21]. Such an output would be expected to help teams distinguish between cases that are predictably near completion and cases whose remaining duration remains uncertain.

Handling Real-Time Intraoperative Updates

Event-driven forecasting triggers

The model would recompute the duration forecast whenever a new intraoperative milestone or event is entered into the operative record. Events such as incision, trocar placement, specimen extraction, closure start, or room-exit preparation

would serve as temporal anchors indicating how far the case has progressed relative to expected workflow. Surgical phase-recognition models show that operative progress can be represented as a sequence of recognisable states, while real-time duration prediction work suggests that forecasts should be updated as the case unfolds rather than fixed at the start [18, 21, 29]. This event-driven approach would allow the forecast to reflect the current surgical phase instead of relying only on preoperative expectations.

Adapting to unexpected events

Unexpected intraoperative events would be treated as high-information signals that may substantially alter the expected remaining duration. For example, conversion from laparoscopic to open surgery, unanticipated adhesions, major bleeding, device malfunction, or additional dissection could shift the hidden state toward a longer and more uncertain completion trajectory. Video-based and workflow-based surgical artificial intelligence studies show that models can recognise procedural deviations, safety-relevant states, and operative actions that may influence downstream progress [30–32]. Attention mechanisms could further help the model assign greater weight to rare but consequential events when revising the forecast [28].

Managing streaming data latency and irregular sampling

Intraoperative event streams are likely to be irregular because not every milestone is documented immediately, and some entries may be charted retrospectively. The model should therefore represent elapsed time explicitly and maintain a provisional forecast when expected events are delayed or missing. Temporal convolutional and recurrent surgical workflow approaches are relevant because they are designed to reason over ordered but variably spaced procedural information [17, 20, 22]. A robust forecasting system would distinguish between true procedural delay and documentation latency while continuing to provide uncertainty-aware updates.

Surgeon-Specific Calibration and Interpretability

Personalisation via surgeon embeddings

A surgeon embedding would allow the model to learn systematic differences in operative pace, case-selection patterns, team coordination, and sensitivity to complexity. This personalisation would not imply that faster or slower surgeons are better or worse, but rather that individual historical patterns can improve case-specific forecasting when used carefully. Studies of machine learning for case duration and robot-assisted procedures support the inclusion of surgeon- and procedure-level predictors, while research on robotic surgery diffusion suggests that technology adoption and surgeon familiarity can shape operative workflow over time [5, 6, 33]. The embedding should therefore be calibrated as a forecasting feature, not as a performance ranking tool.

Explainable forecast updates for the OR team

For clinical acceptance, the model should provide human-readable explanations when it revises predicted completion time. A forecast update might state that remaining time increased because the case converted to open surgery, because closure had not begun when expected, or because anesthesia emergence preparation was delayed. Reviews of surgical artificial intelligence emphasise that transparency, workflow fit, and careful interpretation are essential when models are introduced into operating room environments [23–25]. Explanations should therefore accompany numerical forecasts so that clinicians can assess whether the update is plausible in the current clinical context.

OR Workflow Integration and Decision Support

Real-time dashboard for OR control desks

The model could be integrated into a real-time dashboard showing current case phase, predicted end time, uncertainty range, and confidence status for each active room. Such a display would support decisions about calling the next patient, preparing instruments, coordinating anesthesia availability, and anticipating recovery-room demand. Operating room efficiency studies indicate that case-duration prediction has operational value only when linked to day-of-surgery coordination, staffing, and schedule management [1, 2, 12]. The dashboard should therefore present forecasts as actionable operational signals rather than as isolated machine-learning outputs.

Interaction with scheduling systems and turnover coordination

Forecasts could be passed to scheduling and perioperative coordination systems to adjust expected start times for later cases and align turnover resources. If the model predicts a delayed finish, housekeeping, transport, anesthesia teams, and post-anesthesia care units could be alerted earlier; if completion appears imminent, the next case pathway could be accelerated. Studies on operating room cost, staffing, and delay factors show that prediction errors affect not only the current room but also the broader perioperative system [3, 7, 15]. Integration should therefore focus on safe coordination, avoiding disruptive overreaction to small forecast changes while responding to clinically meaningful shifts.

Evaluation Strategy

Predictive accuracy metrics

The model should be evaluated against surgeon estimates, historical averages, regression models, and static machine-learning baselines at preoperative and intraoperative checkpoints. Accuracy metrics could include mean absolute error, root mean squared error, and percentage error, but the evaluation should report them only in a formal empirical study rather than in this conceptual model article. Prior machine-learning and regression studies provide relevant comparator classes for assessing whether a temporal neural network adds value beyond static prediction [1, 4, 8, 9]. Evaluation should also examine whether forecast quality changes meaningfully after key events such as incision, midpoint milestones, and closure start.

Dynamic forecast reliability

Dynamic reliability should be assessed by determining whether uncertainty narrows appropriately as the procedure progresses and whether prediction intervals remain calibrated across case types. Distributional scoring approaches would be appropriate because a duration model that communicates uncertainty can be more useful than one that reports only a point estimate. Probabilistic surgical forecasting and real-time neural prediction studies support the importance of evaluating both the expected duration and the reliability of the forecast distribution [8, 21]. Reliability assessment should also examine whether rare events produce appropriately wider uncertainty rather than falsely precise estimates.

Operational impact assessment

Operational evaluation should test whether dynamic forecasts would improve scheduling decisions, turnover coordination, and resource allocation compared with usual estimates. Simulation of surgical schedules could assess how updated predictions might affect under-utilisation, over-utilisation, overtime risk, delayed starts, and recovery-room coordination without claiming prospective benefit before deployment. Work on surgical scheduling, operating room economics, and daily caseload prediction provides a foundation for evaluating such system-level consequences [2, 3, 7, 12]. Because biomedical artificial intelligence benchmarks can be sensitive to design choices, operational evaluation should be interpreted cautiously and validated across realistic institutional workflows [34].

Limitations

Data granularity and intraoperative log completeness

A central limitation is that many operating rooms do not maintain structured, timely, and standardised intraoperative event logs. Some events may be missing, entered late, or documented inconsistently across teams, which could make the temporal signal noisy and difficult to interpret. Surgical workflow research using video and structured phase data often relies on richer or more curated annotations than are available in routine perioperative information systems [19, 26, 29]. Therefore, the model's usefulness would depend on the quality, completeness, and timeliness of local documentation practices.

Surgeon-specific and institutional variability

A surgeon-specific model may learn institutional routines, staffing patterns, room practices, technology availability, and documentation habits that do not transfer cleanly to other hospitals. This risk is especially important when procedure mix, anesthesia workflow, robotic platform adoption, and perioperative staffing models differ across sites. Reviews of surgical artificial intelligence caution that models should not be assumed to generalise without external validation, workflow assessment, and careful recalibration [23-25]. A deployment strategy would therefore need to include site-specific monitoring and safeguards against encoding unfair or misleading provider-level interpretations.

Conclusion

A temporal neural network for operating room procedure-duration forecasting would combine preoperative case context with real-time evidence from the unfolding operation. It would begin with a baseline estimate and then revise expected remaining time as milestones, delays, and unexpected events are recorded.

The main strength of this approach is its ability to fuse procedure codes, surgeon-specific history, anesthesia information, patient comorbidity profiles, and intraoperative event streams into a single dynamic forecasting framework. By producing updated uncertainty estimates, the model could support more nuanced decisions than a fixed scheduled duration.

Important challenges remain before such a model could be used safely in live perioperative operations. These include incomplete event documentation, institutional variation, model calibration, user trust, alert fatigue, and integration with scheduling systems.

Prospective pilot studies in high-volume surgical suites would be needed to determine whether dynamic duration forecasts improve operating room efficiency and team acceptance. Such studies should evaluate not only predictive accuracy but also whether forecasts are understood, trusted, and used appropriately by perioperative teams.

Acknowledgements

None

Conflict of interest

None

Financial support

None

Ethics statement

None

Received: 29 Jan 2022

Revised: 10 Mar 2022

Accepted: 29 Apr 2022

Published online: 20 July 2022

Rights and permissions

Open Access The author(s) retain copyright. This article is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License. It may be shared and adapted for non-commercial purposes with appropriate attribution, an indication of changes, and distribution of adaptations under the same license. Third-party material may be subject to separate terms identified in its credit line. View the license at <https://creativecommons.org/licenses/by-nc-sa/4.0/>.

References

- Bartek MA, Saxena RC, Solomon S, Fong CT, Behara LD, Venigandla R, et al. Improving operating room efficiency: machine learning approach to predict case-time duration. *J Am Coll Surg.* 2019;229(4):346-54.
<https://doi.org/10.1016/j.jamcollsurg.2019.06.004>.
- Guetter CR, Williams BJ, Slama E, Arrington A, Henry MC, Möller MG, et al. Greening the operating room. *Am J Surg.* 2018;216(4):683-8.
<https://doi.org/10.1016/j.amjsurg.2018.07.033>.
- Childers CP, Maggard-Gibbons M. Understanding costs of care in the operating room. *JAMA Surg.* 2018;153(4):e176233.
<https://doi.org/10.1001/jamasurg.2017.6233>.
- Edelman ER, Van Kuijk SM, Hamaekers AE, De Korte MJ, Van Merode GG, Buhre WF. Improving the prediction of total surgical procedure time using linear regression modeling. *Front Med (Lausanne).* 2017;4:85.
<https://doi.org/10.3389/fmed.2017.00085>.
- Zhao B, Waterman RS, Urman RD, Gabriel RA. A machine learning approach to predicting case duration for robot-assisted surgery. *J Med Syst.* 2019;43(2):32.
<https://doi.org/10.1007/s10916-018-1140-4>.
- Tuwatananurak JP, Zadeh S, Xu X, Vacanti JA, Fulton WR, Ehrenfeld JM, et al. Machine learning can improve estimation of surgical case duration: a pilot study. *J Med Syst.* 2019;43(3):44.
<https://doi.org/10.1007/s10916-019-1185-z>.
- Burns ML, Saager L, Cassidy RB, Mentz G, Mashour GA, Kheterpal S. Association of anesthesiologist staffing ratio with surgical patient morbidity and mortality. *JAMA Surg.* 2022;157(9):807-15.
<https://doi.org/10.1001/jamasurg.2022.2804>.
- Jiao Y, Sharma A, Ben Abdallah A, Maddox TM, Kannampallil T. Probabilistic forecasting of surgical case duration using machine learning: model development and validation. *J Am Med Inform Assoc.* 2020;27(12):1885-93.
- Martinez O, Martinez C, Parra CA, Rugeles S, Suarez DR. Machine learning for surgical time prediction. *Comput Methods Programs Biomed.* 2021;208:106220.
<https://doi.org/10.1016/j.cmpb.2021.106220>.
- Chu J, Hsieh CH, Shih YN, Wu CC, Singaravelan A, Hung LP, et al. Operating room usage time estimation with machine learning models. *Healthcare (Basel).* 2022;10(8):1518.

<https://doi.org/10.3390/healthcare10081518>.

Abbas A, Mosseri J, Lex JR, Toor J, Ravi B, Khalil EB, et al. Machine learning using preoperative patient factors can predict duration of surgery and length of stay for total knee arthroplasty. *Int J Med Inform.* 2022;158:104670.

<https://doi.org/10.1016/j.ijmedinf.2021.104670>.

Hassanzadeh H, Boyle J, Khanna S, Biki B, Syed F. Daily surgery caseload prediction: towards improving operating theatre efficiency. *BMC Med Inform Decis Mak.* 2022;22(1):151.

<https://doi.org/10.1186/s12911-022-01890-z>.

Yuniartha DR, Masruroh NA, Herliansyah MK. An evaluation of a simple model for predicting surgery duration using a set of surgical procedure parameters. *Inform Med Unlocked.* 2021;25:100633.

<https://doi.org/10.1016/j.imu.2021.100633>.

Huang L, Chen X, Liu W, Shih PC, Bao J. Automatic surgery and anesthesia emergence duration prediction using artificial neural networks. *J Healthc Eng.* 2022;2022(1):2921775.

Pappada SM, Papadimos TJ, Khuder S, Mack ST, Beachy PZ, Casabianca AB. Contributing factors to operating room delays identified from an electronic health record: a retrospective study. *Anesthesiol Res Pract.* 2022;2022(1):8635454.

Celentano V, Smart N, Cahill RA, McGrath JS, Gupta S, Griffith JP, et al. Use of laparoscopic videos amongst surgical trainees in the United Kingdom. *Surgeon.* 2019;17(6):334-9.

<https://doi.org/10.1016/j.surge.2018.10.004>.

Jin Y, Dou Q, Chen H, Yu L, Qin J, Fu CW, et al. SV-RCNet: workflow recognition from surgical videos using recurrent convolutional network. *IEEE Trans Med Imaging.* 2018;37(5):1114-26.

<https://doi.org/10.1109/TMI.2017.2787659>.

Czempiel T, Paschali M, Keicher M, Simson W, Feussner H, Kim ST, et al. TeCNO: surgical phase recognition with multi-stage temporal convolutional networks. In: *Med Image Comput Comput Assist Interv. Cham: Springer; 2020.* p. 343-52.

https://doi.org/10.1007/978-3-030-59710-8_34.

Kitaguchi D, Takeshita N, Matsuzaki H, Takano H, Owada Y, Enomoto T, et al. Real-time automatic surgical phase recognition in laparoscopic sigmoidectomy using the convolutional neural network-based deep learning approach. *Surg Endosc.* 2020;34(11):4924-31.

<https://doi.org/10.1007/s00464-019-07249-4>.

Lea C, Flynn MD, Vidal R, Reiter A, Hager GD. Temporal convolutional networks for action segmentation and detection. In: *Proc IEEE Conf Comput Vis Pattern Recognit; 2017.* p. 156-65.

<https://doi.org/10.1109/CVPR.2017.25>.

Jiao Y, Xue B, Lu C, Avidan MS, Kannampallil T. Continuous real-time prediction of surgical case duration using a modular artificial neural network. *Br J Anaesth.* 2022;128(5):829-37.

<https://doi.org/10.1016/j.bja.2022.01.025>.

Ramesh S, Dall'Alba D, Gonzalez C, Yu T, Mascagni P, Mutter D, et al. Multi-task temporal convolutional networks for joint recognition of surgical phases and steps in gastric bypass procedures. *Int J Comput Assist Radiol Surg.* 2021;16(7):1111-9.

<https://doi.org/10.1007/s11548-021-02388-3>.

Garrow CR, Kowalewski KF, Li L, Wagner M, Schmidt MW, Engelhardt S, et al. Machine learning for surgical phase recognition: a systematic review. *Ann Surg.* 2021;273(4):684-93.

<https://doi.org/10.1097/SLA.0000000000004425>.

Hashimoto DA, Rosman G, Rus D, Meireles OR. Artificial intelligence in surgery: promises and perils. *Ann Surg.* 2018;268(1):70-6.

<https://doi.org/10.1097/SLA.0000000000002693>.

Birkhoff DC, van Dalen AS, Schijven MP. A review on the current applications of artificial intelligence in the operating room. *Surg Innov.* 2021;28(5):611-9.

<https://doi.org/10.1177/15533506211004539>.

Huauilmé A, Sarikaya D, Le Mut K, Despinoy F, Long Y, Dou Q, et al. Micro-surgical anastomose workflow recognition challenge report. *Comput Methods Programs Biomed.* 2021;212:106452.

<https://doi.org/10.1016/j.cmpb.2021.106452>.

Chen X, Huang L, Liu W, Shih PC, Bao J. Automatic surgery duration prediction using artificial neural networks. In: *Proc Int Conf Comput Sci Appl Eng*; 2021. p. 1-6.

<https://doi.org/10.1145/3483181.3483190>.

Czempiel T, Paschali M, Ostler D, Kim ST, Busam B, Navab N. OPERA: attention-regularized transformers for surgical phase recognition. In: *Med Image Comput Comput Assist Interv. Cham: Springer*; 2021. p. 604-14.

https://doi.org/10.1007/978-3-030-87234-2_58.

Yengera G, Mutter D, Marescaux J, Padoy N. Less is more: surgical phase recognition with less annotations through self-supervised pre-training of CNN-LSTM networks. *arXiv.* 2018;arXiv:1805.08569.

Jin A, Yeung S, Jopling J, Krause J, Azagury D, Milstein A, et al. Tool detection and operative skill assessment in surgical videos using region-based convolutional neural networks. In: *Proc IEEE Winter Conf Appl Comput Vis*; 2018. p. 691-9.

<https://doi.org/10.1109/WACV.2018.00081>.

Khan DZ, Luengo I, Barbarisi S, Addis C, Culshaw L, Dorward NL, et al. Automated operative workflow analysis of endoscopic pituitary surgery using machine learning: development and preclinical evaluation (IDEAL stage 0). *J Neurosurg.* 2022;137(1):51-8.

<https://doi.org/10.3171/2021.2.JNS204366>.

Mascagni P, Vardazaryan A, Alapatt D, Urade T, Emre T, Fiorillo C, et al. Artificial intelligence for surgical safety: automatic assessment of the critical view of safety in laparoscopic cholecystectomy using deep learning. *Ann Surg.* 2022;275(5):955-61.

<https://doi.org/10.1097/SLA.0000000000004700>.

Marcus HJ, Hughes-Hallett A, Payne CJ, Cundy TP, Nandi D, Yang GZ, et al. Trends in the diffusion of robotic surgery: a retrospective observational study. *Int J Med Robot.* 2017;13(4):e1870.

<https://doi.org/10.1002/rcs.1870>.

Maier-Hein L, Eisenmann M, Reinke A, Onogur S, Stankovic M, Scholz P, et al. Why rankings of biomedical image analysis competitions should be interpreted with care. *Nat Commun.* 2018;9(1):5217.

<https://doi.org/10.1038/s41467-018-07619-7>.