

# Explainable Machine Learning for Prediction of Early Postoperative Nausea and Vomiting After General Anesthesia

Hsiao-Cheng Chang <sup>1,2</sup>, Li-Yun Chen <sup>1,3</sup>, Yu-Shiang Lin <sup>1</sup>

<sup>1</sup>In-Service Master Program in Artificial Intelligence in Medicine, College of Medicine, Taipei Medical University, Taipei, Taiwan; <sup>2</sup>Department of Anesthesiology, Cathay General Hospital, Taipei, Taiwan; <sup>3</sup>Division of Respiratory Therapy, Department of Chest Medicine, Taipei Veterans General Hospital, Taipei, Taiwan

Correspondence: Yu-Shiang Lin, In-Service Master Program in Artificial Intelligence in Medicine, College of Medicine, Taipei Medical University, No. 250, Wuxing Street, Xinyi District, Taipei City, 110, Taiwan, Email eriklin@tmu.edu.tw

**Purpose:** Postoperative nausea and vomiting (PONV) remains one of the most common adverse effects associated with anesthesia care. This study aimed to explore the feasibility of applying machine learning models trained exclusively on routinely available non-invasive clinical indicators to predict early PONV risk. Explainable artificial intelligence techniques were also employed to identify the most influential predictors of early PONV.

**Patients and Methods:** A retrospective dataset from Cathay General Hospital, including 927 patient cases and 16 non-invasive clinical indicators, was used to investigate early PONV risk prediction. This study evaluated the predictive performance of several traditional machine learning models, deep learning architectures, and ensemble learning methods to compare their classification capabilities.

**Results:** Overall, the models demonstrated moderate discriminative performance. The random forest model achieved an accuracy of 83.5% with balanced precision (80.81%) and recall (83.5%), while the logistic regression model attained an AUC of 0.6905. Analysis of positive SHAP values identified the top 7 most influential predictors of early PONV. These included pharmacologic interventions (eg, neostigmine), pre-existing comorbidities (eg, history of nausea and vomiting, history of cardiovascular disease), demographic characteristics (eg, gender), postoperative pain, and anesthetic and surgical factors (eg, type of surgery and duration of anesthesia). Moreover, SHAP analysis revealed that the use of dexamethasone was negatively associated with the predicted risk in the model, suggesting its potential protective role in the prevention of early PONV.

**Conclusion:** By generating explainable outputs, this study bridges the gap between algorithmic prediction and clinical decision-making, allowing anesthesiologists to better recognize underlying risk factors and make informed, evidence-based decisions in perioperative management.

**Keywords:** postoperative nausea and vomiting, machine learning, explainable artificial intelligence, risk prediction, anesthesiology

## Introduction

Postoperative nausea and vomiting (PONV) represent one of the most common and distressing complications in the perioperative setting. In the absence of prophylactic intervention, its incidence has been reported to range from 20% to 30% in the general surgical population, and to reach as high as 70–80% in patients with multiple risk factors.<sup>1,2</sup> The etiology of PONV is multifactorial, encompassing patient-related characteristics, surgical procedures, and anesthetic techniques. Among these, pharmacological agents administered during the perioperative period play a pivotal role in modulating PONV risk.

Postoperative nausea and vomiting (PONV) is a multifactorial complication influenced by anesthetic, surgical, and patient-related factors. Based on previous studies, several anesthetic and adjunctive agents modulate central and peripheral emetogenic pathways; neostigmine, commonly used for reversal of non-depolarizing neuromuscular blockade, increases acetylcholine availability and has been associated with an elevated risk of PONV.<sup>3</sup> Opioids, despite their

analgesic efficacy, contribute to PONV through  $\mu$ -opioid receptor activation in the vestibular system and chemoreceptor trigger zone, as well as through suppression of gastrointestinal motility.<sup>4–6</sup> Previous studies have reported that prolonged anesthesia duration and the use of volatile anesthetics are associated with an increased risk of PONV, potentially reflecting cumulative emetogenic exposure.<sup>7</sup> Surgical characteristics, including longer procedure duration and specific operative types, such as gynecological or intra-abdominal surgery, independently contribute to PONV.<sup>8,9</sup> Among patient-related factors, female sex, younger age, and a prior history of PONV remain the most consistent predictors, underscoring the importance of integrated perioperative risk stratification.<sup>2</sup> Based on previous evidence, the preoperative identification of high-risk factors for postoperative nausea and vomiting (PONV) may facilitate early recognition of high-risk patients and contribute to an effective reduction in postoperative PONV risk.

Current risk assessment tools for postoperative nausea and vomiting (PONV) have notable structural and practical limitations. Most existing models rely on a limited set of variables, such as sex, smoking status, and postoperative opioid use, and therefore fail to adequately capture individual physiological differences, surgical characteristics, and perioperative analgesic strategies, resulting in limited predictive accuracy and clinical interpretability. To address these limitations, the present study applies interpretable machine learning techniques that integrate perioperative variables to predict early PONV risk, identify key contributing factors, and enhance model transparency and clinical relevance.

Machine learning has been increasingly adopted in healthcare, as it enables the identification of complex patterns and relationships from large datasets beyond the capabilities of traditional rule-based approaches. In recent years, machine learning has been increasingly applied in clinical disciplines, including internal medicine and surgery, to support disease risk assessment and prognosis prediction. These approaches have demonstrated robust performance and considerable clinical potential. For example, Hagan et al evaluated multiple machine learning models across two cardiovascular disease datasets and reported that the Random Forest algorithm achieved an accuracy of 74%, demonstrating robustness and reliability in heterogeneous data environments.<sup>10</sup> Similarly, Luo et al applied machine learning techniques to predict recovery outcomes in patients with Bell's palsy, showing that Logistic Regression achieved the best predictive accuracy at 3 and 9 months (AUCs of 0.751 and 0.720, respectively), with age and prednisolone use identified as significant predictors.<sup>11</sup>

In the surgical domain, Mai et al analyzed 353 patients undergoing hemihepatectomy for hepatocellular carcinoma, where an artificial neural network achieved AUCs of 0.880 and 0.876 in the training and test sets, respectively, for predicting severe posthepatectomy liver failure. This model may aid in identifying intermediate- and high-risk patients, thereby facilitating timely interventions.<sup>12</sup> Likewise, Salat et al developed an extreme gradient boosting model to predict cardiopulmonary complications after lung resection in 1360 patients, achieving an AUC of 0.75 and an accuracy of 70%. These findings suggest that machine learning enables individualized risk prediction and may support surgical decision-making.<sup>13</sup> In neurosurgery, Farrokhi et al demonstrated that supervised machine learning models achieved high discriminatory performance in predicting complications following deep brain stimulation surgery, with AUCs of 0.86 for any complication and 0.97 for infection, indicating potential utility in perioperative risk assessment and treatment planning.<sup>14</sup>

Applications of machine learning have also extended to predicting PONV. Zhou et al reported that among several algorithms predicting early PONV, CNN-RNN achieved the highest accuracy (0.872), while Logistic Regression, SVC, and AdaBoost achieved the best AUCs (0.732, 0.731, and 0.722, respectively), with Logistic Regression and SVC showing the most consistent overall performance.<sup>15</sup> Kim et al analyzed 106,860 adult patients and demonstrated that models incorporating known risk and mitigating factors yielded AUROCs of 0.54–0.69, with opioid use via patient-controlled analgesia identified as a dominant predictor.<sup>16</sup> Furthermore, Zheng et al applied machine learning to 1154 patients to predict delayed clinically important PONV (CIPONV), with the Random Forest model achieving an AUC of 0.737 in the test cohort. This interpretable model facilitates individualized risk prediction and may assist in early identification and prevention of CIPONV in high-risk patients.<sup>17</sup> Although these studies have demonstrated promising performance in PONV prediction using machine learning, limited research has been conducted among Taiwanese and other Asian populations, which restricts the generalizability of these predictive models across different ethnic groups.

The present study aims to develop and validate an early prediction model for postoperative nausea and vomiting (PONV) using routinely available, noninvasive clinical variables and patient medical history to facilitate individualized risk stratification. The predictive performance of multiple machine learning models will be systematically evaluated and compared to identify approaches with superior discriminative capability. Furthermore, by incorporating interpretable

machine learning methodologies, this study seeks to assess the concordance between model-identified key risk factors and established clinical evidence, thereby enhancing model transparency and strengthening clinical confidence in the applicability of the proposed prediction framework.

## Materials and Methods

### Dataset

This study complied with the Helsinki Declaration and was approved by the Institutional Review Board of Cathay General Hospital (CGH-P114040). The informed consent was exempted, as it is a retrospective study utilizing a database, and the data have been anonymized, preventing any identification of individual cases.

In this study, data were retrospectively collected from medical records between January 1, 2019, and November 1, 2024, at Cathay General Hospital. A total of 927 patients and 16 non-invasive clinical indicators were included as input features to develop predictive models for early postoperative nausea and vomiting (Table 1). The dataset included adult inpatients who were admitted to the post-anesthesia care unit (PACU) following surgical procedures performed under general anesthesia with endotracheal intubation.

### Data Preprocessing

Patients were excluded if they met any of the following criteria: received regional anesthesia or underwent surgery under light sedation, had an American Society of Anesthesiologists (ASA) physical status classification of IV or V, or required mechanical ventilation support postoperatively. Early postoperative nausea and vomiting was defined as events occurring during the post-anesthesia care unit (PACU) stay, from extubation to discharge, which typically occurred within approximately 1 hour postoperatively at our institution. Pharmacologic interventions considered PONV events included the administration of rescue antiemetic medications, specifically metoclopramide or prochlorperazine injections, during the PACU stay.

**Table 1** Feature Description of the Early Postoperative Nausea and Vomiting Dataset

No	Feature Name	Attribute	Data Type
1	Age (Years)	Age	Int (Years)
2	Gender	Gender	0: Female; 1: Male
3	Weight (kg)	Weight	Int (Kg)
4	ASA	Asa	0: ASA = 2; 1: ASA = 3
5	Type of Surgery	Type of surgery	1: General Surgery 2: Orthopedics 3: Neurosurgery 4: Thoracic Surgery 5: Otolaryngology 6: Obstetrics and Gynecology 7: Urology 8: Plastic Surgery
6	History of Nausea and Vomiting	History of nausea and vomiting	0: No; 1: Yes
7	History of Cardiovascular Disease	History of cardiovascular disease	0: No; 1: Yes
8	History of Gastrointestinal Disease	History of gastrointestinal disease	0: No; 1: Yes
9	Surgery Time $\geq 100$ (min)	Surgery time	0: No; 1: Yes
10	Anesthesia Time $\geq 120$ (min)	Anesthesia time	0: No; 1: Yes
11	Inhalation Anesthesia	Inhalation anesthesia	0: No; 1: Yes
12	Neostigmine	Neostigmine	0: No; 1: Yes
13	Opioid	Opioid	0: No; 1: Yes
14	Dexamethason	Dexamethason	0: No; 1: Yes
15	Multimodal Pain Management	Multimodal pain management	0: No; 1: Yes
16	Pain	Pain	0: No; 1: Yes

In this study, a total of sixteen clinical and perioperative features were included as predictor variables. Clinical data were obtained from the institutional electronic patient record system, which included the following features: demographic variables consisted of age, gender, and body weight. Surgical characteristics included the type of surgery, categorized into eight subgroups: general surgery, orthopedic surgery, neurosurgery, thoracic surgery, otolaryngology, obstetrics and gynecology, urology, and plastic surgery. Furthermore, the investigation encompassed an evaluation of surgical duration as well as anesthesia duration. As this study is based on retrospective clinical data, and surgery duration and anesthesia duration represent different aspects of the clinical workflow (surgical procedure versus overall anesthesia management), these variables were considered clinically meaningful but not fully overlapping and were retained as predictor variables. Medical history variables included prior history of nausea and vomiting, cardiovascular disease, and gastrointestinal disease. Anesthetic and pharmacological variables included the use of inhalation anesthesia, neostigmine, opioids, and dexamethasone. Postoperative management factors included multimodal pain management strategies and the presence of postoperative pain. A detailed overview of these features is provided in [Table 1](#). All variables were selected based on clinical relevance and prior evidence in the literature regarding their potential association with postoperative nausea and vomiting.

In this study, all 927 patients had complete data for all 16 features; therefore, no missing data handling or imputation procedures were required. The continuous variables were standardized prior to model training using the StandardScaler (z-score normalization) to eliminate the influence of differing feature scales. Categorical variables were converted into numerical representations according to predefined categories before being included in the analysis. One-hot encoding was not applied in order to avoid excessive feature dimensionality, given the limited sample size.

## Machine Learning Methods

A retrospective dataset from a single medical center, comprising 927 patient cases and 16 non-invasive clinical features, was used to develop the early PONV risk prediction models. All analyses were performed using Python within the Visual Studio Code (VS Code) environment, and machine learning models were implemented using the scikit-learn library. This study assessed the predictive performance of seven traditional machine learning models—Logistic Regression, Stochastic Gradient Descent (SGD), Support Vector Machine (SVM), K-Nearest Neighbors (KNN), Decision Tree, Multilayer Perceptron (MLP), and Artificial Neural Network (ANN); two deep learning models—Long Short-Term Memory (LSTM) and Convolutional Neural Network (CNN); and three ensemble learning models—Random Forest, Extreme Gradient Boosting (XGBoost), and Light Gradient Boosting Machine (LightGBM). The dataset was randomly split into training (80%) and testing (20%) subsets, following the Pareto principle.<sup>18</sup> To ensure robustness and mitigate variance in performance estimation, five-fold cross-validation was performed on the entire dataset due to the limited sample size. Model performance metrics were calculated for each fold and reported as the mean across the five folds. Considering the inherent class imbalance in the original dataset, the Synthetic Minority Oversampling Technique (SMOTE) was applied only to the training set after the initial data split. No oversampling was performed on the test set. The SMOTE sampling ratio was set to 1:1 to balance the minority and majority classes within the training set.

In this study, most traditional machine learning models adopted default or near-default parameter settings provided by the scikit-learn library to maintain consistency across model comparisons and to avoid optimistic bias resulting from excessive hyperparameter tuning in the context of a limited sample size. For neural network-based models, conservative and fixed architectures and training settings were used to control model complexity. The complete hyperparameter settings are provided in [Table S1](#).

To mitigate overfitting, several strategies were adopted. For the logistic regression model, default regularization settings were used to constrain model complexity. For neural network-based models, early stopping was applied with validation loss monitored to prevent overtraining.

## Model Evaluation

The predictive performance of each model was evaluated using several metrics, including the area under the receiver operating characteristic curve (AUC), accuracy, precision, recall, and F1 score. All metrics were derived from the independent test dataset. The receiver operating characteristic (ROC) curves were generated using the predicted probabilities produced by each model and plotted as the true positive rate (TPR; sensitivity) against the false positive

rate (FPR;  $1 - \text{specificity}$ ). The AUC represents the area under the ROC curve and quantifies the overall ability of the model to discriminate between positive and negative classes across all possible classification thresholds. To further assess classification outcomes, a confusion matrix was employed, providing a comprehensive overview of the relationship between actual and predicted classifications, consisting of true positives ( $TP$ ), true negatives ( $TN$ ), false positives ( $FP$ ), and false negatives ( $FN$ ). The evaluation metrics were defined as follows:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

$$\text{Precision} = \frac{TP}{TP + FP}$$

$$\text{Recall} = \frac{TP}{TP + FN}$$

$$\text{F1 score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

Where  $TP$  represents cases in which the model correctly predicted the presence of early PONV, while  $TN$  denotes cases in which the absence of early PONV was accurately identified. Higher values of  $TP$  and  $TN$  indicate greater discriminative ability of the model in distinguishing between individuals with and without early PONV. Analysis of the confusion matrix provides valuable insights into both the strengths and limitations of the model, enabling a more comprehensive evaluation of its capacity to detect early PONV risk in clinical applications.

## Results

### Baseline Characteristics

The dataset utilized in this study provides detailed records of risk factors associated with PONV. The baseline clinical characteristics of the patients are summarized in Table 2. A total of 927 individuals were included, comprising 511 females (55.1%) and 416 males (44.9%). Overall, 140 patients (15.1%) were diagnosed with PONV.

**Table 2** Baseline Characteristics of PONV Patients

Characteristics	Absence of PONV	Presence of PONV
<b>Numbers</b>	787	140
<b>Age (years)</b>	53.4 ± 17.2	51.2 ± 16.7
<b>Weight (kg)</b>	64.8 ± 14.5	61.4 ± 11.0
<b>Types of Surgery</b>		
1 General Surgery	270 (34.3%)	51 (36.4%)
2 Orthopedics	86 (10.9%)	15 (10.7%)
3 Neurosurgery	80 (10.2%)	20 (14.3%)
4 Thoracic Surgery	53 (6.7%)	4 (2.9%)
5 Otolaryngology	128 (16.3%)	27 (19.3%)
6 Obstetrics and Gynecology	90 (11.4%)	15 (10.7%)
7 Urology	44 (5.6%)	3 (2.1%)
8 Plastic Surgery	36 (4.6%)	5 (3.6%)
<b>Gender</b>		
Female (0)	415 (52.7%)	96 (68.6%)
Male (1)	372 (47.3%)	44 (31.4%)

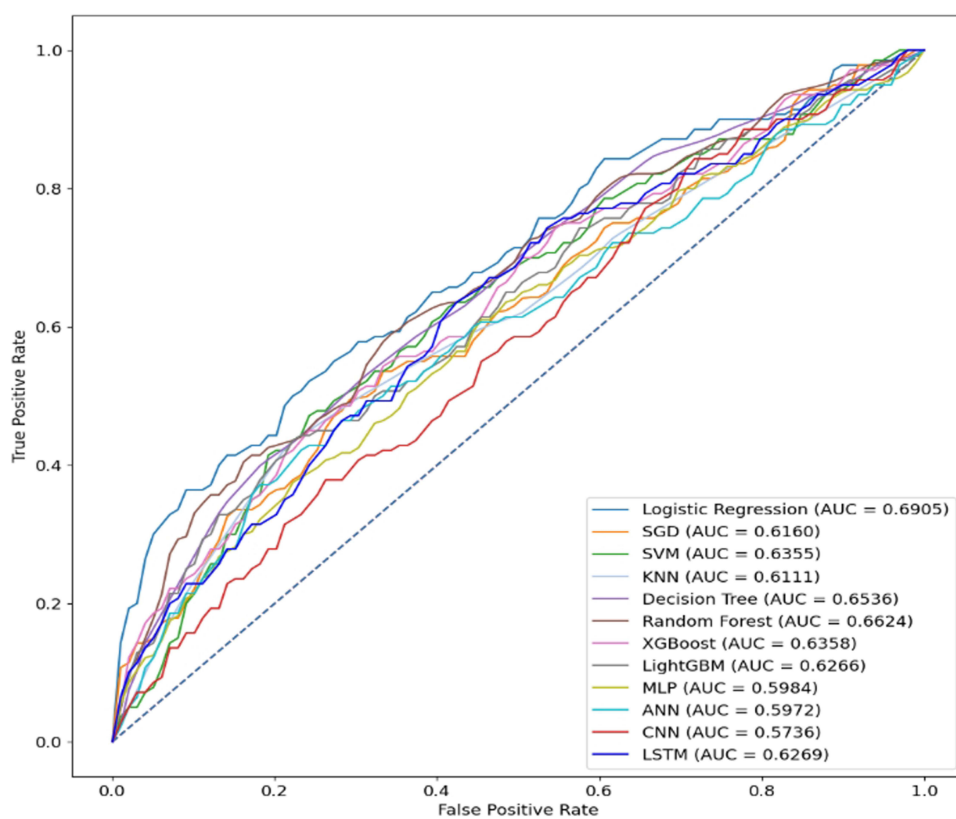
(Continued)

Table 2 (Continued).

Characteristics	Absence of PONV	Presence of PONV
<b>ASA</b>		
II (2)	519 (65.9%)	105 (75.0%)
III (3)	268 (34.1%)	35 (25.0%)
<b>History of Cardiovascular</b>		
No (0)	520 (66.1%)	109 (77.9%)
Yes (1)	267 (33.9%)	31 (22.1%)
<b>History of Gastroenterology</b>		
No (0)	571 (72.6%)	102 (72.9%)
Yes (1)	216 (27.4%)	38 (27.1%)
<b>History of Nausea and Vomiting</b>		
No (0)	769 (97.7%)	113 (80.7%)
Yes (1)	18 (2.3%)	27 (19.3%)
<b>Surgery Time <math>\geq</math>100 min</b>		
No (0)	322 (40.9%)	48 (34.3%)
Yes (1)	465 (59.1%)	92 (65.7%)
<b>Anesthesia Time <math>\geq</math>120 min</b>		
No (0)	322 (40.9%)	48 (34.3%)
Yes (1)	465 (59.1%)	92 (65.7%)
<b>Inhalation Anesthesia</b>		
No (0)	94 (11.9%)	9 (6.4%)
Yes (1)	693 (88.1%)	131 (93.6%)
<b>Dexamethasone</b>		
No (0)	648 (82.3%)	122 (87.1%)
Yes (1)	139 (17.7%)	18 (12.9%)
<b>Neostigmine</b>		
No (0)	405 (51.5%)	36 (25.7%)
Yes (1)	382 (48.5%)	104 (74.3%)
<b>Pain</b>		
No (0)	556 (70.6%)	75 (53.6%)
Yes (1)	231 (29.4%)	65 (46.4%)
<b>Multimodal Pain Management</b>		
No (0)	620 (78.8%)	108 (77.1%)
Yes (1)	167 (21.2%)	32 (22.9%)
<b>Opioid</b>		
No (0)	168 (21.3%)	32 (22.9%)
Yes (1)	619 (78.7%)	108 (77.1%)

## Model Performance

The Model's performance was evaluated using AUC along with several quantitative metrics, including accuracy, precision, recall, and F1 score. The analysis was conducted using all 16 clinical and demographic features. As illustrated in Figure 1, the ROC curves provide a graphical representation of the classification performance of each model. The detailed quantitative evaluation metrics are presented in Table 3.



**Figure 1** Comparison of receiver operating characteristic (ROC) curves among machine learning models for early PONV risk prediction.

The findings of this study demonstrate that machine learning models can effectively predict early PONV using non-invasive clinical indicators. Among all evaluated models, the Logistic Regression model achieved the highest AUC (0.6905), reflecting its moderate discriminatory capacity in distinguishing patients at risk of early PONV from those without risk. Furthermore, its performance across precision (80.33%) and recall (66.67%) remained well balanced,

**Table 3** Performance Comparison of Machine Learning Models for Early PONV Risk Prediction

Model	Accuracy	Precision	Recall	F1 Score
LR	0.6667	0.8033	0.6667	0.7114
SGD	0.6063	0.7792	0.6063	0.6615
SVM	0.7389	0.7922	0.7389	0.7610
KNN	0.6429	0.7853	0.6429	0.6911
Decision Tree	0.6257	0.7924	0.6257	0.6780
Random Forest	0.8350	0.8081	0.8350	0.8176
XGBoost	0.8026	0.7796	0.8026	0.7898
LightGBM	0.8047	0.7848	0.8047	0.7937
MLP	0.7184	0.7621	0.7184	0.7380
ANN	0.7735	0.7825	0.7735	0.7778
CNN	0.7799	0.7664	0.7799	0.7729
LSTM	0.7422	0.7756	0.7422	0.7572

**Abbreviations:** LR, Logistic Regression; SGD, Stochastic Gradient Descent; SVM, Support Vector Machine; KNN, K-Nearest Neighbor; XGBoost, eXtreme Gradient Boosting; LightGBM, Light Gradient Boosting Machine; MLP, Multilayer Perceptron; ANN, Artificial Neural Network; CNN, Convolutional Neural Network; LSTM, Long Short-Term Memory.

indicating that the model not only reduced false-positive predictions but also retained an adequate ability to correctly identify true early PONV cases. Collectively, these findings underscore Logistic Regression as a robust and interpretable algorithm, which offers both clinical applicability and reliability in the context of early PONV risk prediction. Within the traditional models, MLP (AUC = 0.5984) and ANN (AUC = 0.5972) showed marginally inferior performance compared to the best-performing ensemble model.

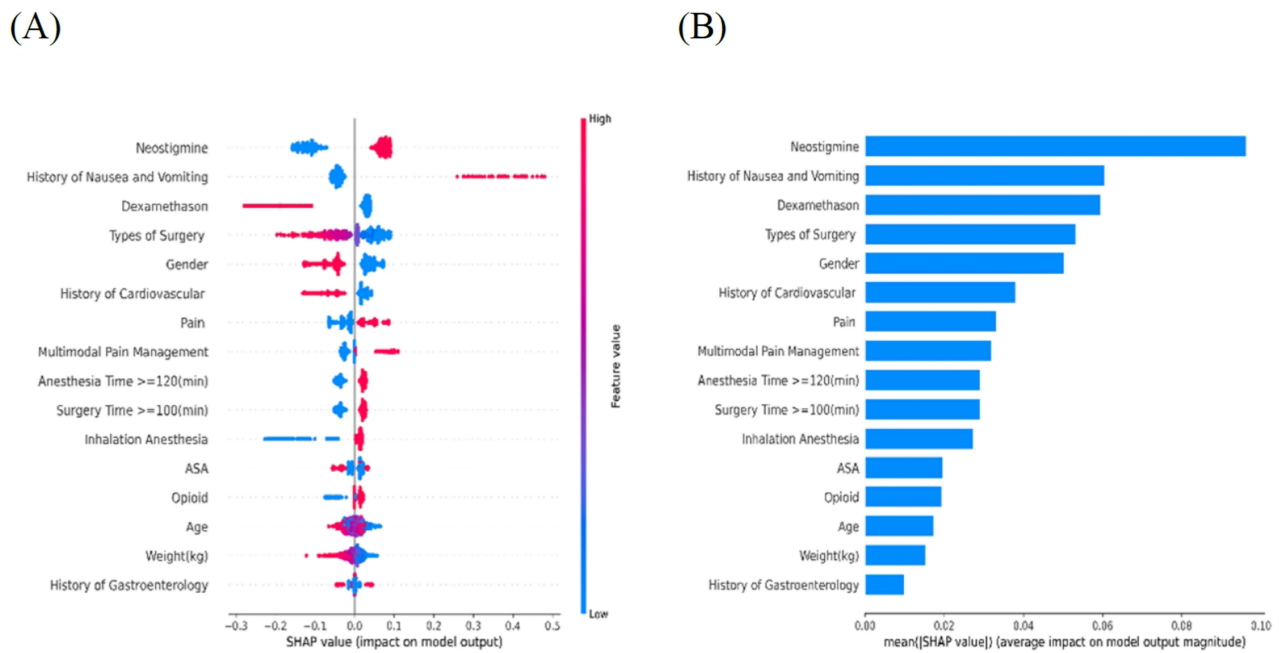
Among the ensemble learning models, the Random Forest algorithm exhibited the strongest overall performance, achieving an accuracy of 83.5%, which indicates a high level of agreement between predicted and observed outcomes. Its classification metrics were consistently well balanced, with a precision of 80.81%, a recall of 83.5%, and an F1 score of 81.76%. The accuracy of LightGBM (80.47%) and XGBoost (80.26%) was slightly lower than that of the Random Forest model. Regarding the deep learning models, the LSTM model achieved an AUC of 0.6269 and an accuracy of 74.22%. The CNN model recorded the lowest AUC (0.5736) among all models.

In summary, among all machine learning models, the Random Forest model demonstrated high accuracy and balanced performance, while the Logistic Regression model achieved the highest AUC.

## Feature Importance Analysis

The SHAP analysis was employed to evaluate the importance of each variable in shaping the model output. SHAP provides a unified framework for interpreting the contribution of individual features to the model's predictions, quantifying both the direction and magnitude of impact. Features were ranked according to their mean absolute SHAP values, with higher ranks indicating greater contributions to early PONV risk prediction. Among the evaluated models, Logistic Regression achieved the highest AUC. Given its relatively better performance within this context, together with its transparent model structure and the clear directional interpretation of feature effects, Logistic Regression was selected as the primary reference model for SHAP-based interpretability analysis.

The SHAP summary plot (Figure 2A) revealed that Neostigmine was the most influential predictor. High SHAP values associated with Neostigmine use emerged as the strongest positive predictor, consistent with prior studies linking acetylcholinesterase inhibitors to an increased risk of early PONV. Similarly, a prior history of nausea and vomiting markedly increased the model's predicted risk. Other relevant features included the type of surgery, gender, and history of



**Figure 2** SHAP-based feature importance analysis for early PONV risk prediction using the logistic regression model. (A) SHAP summary plot illustrating the direction and magnitude of each feature's impact on model output. (B) Mean absolute SHAP value plot ranking the average contribution of each feature to model prediction.

cardiovascular disease, which demonstrated variable but notable effects. Longer anesthesia and surgery durations generally shifted predictions toward a higher risk of early PONV. In contrast, features such as inhalation anesthesia and opioid administration had an influence, but their effects were less pronounced.

Conversely, the administration of dexamethasone exhibited negative SHAP values, particularly when its feature values were high, implying a protective effect against early PONV. These findings align with existing clinical guidelines advocating the prophylactic use of antiemetic agents to prevent PONV. Demographic and physiological factors, such as age, body weight, and ASA classification, exerted a relatively minor influence on the model outputs (Figure 2B).

## Discussion

In the present study, high SHAP values associated with elevated Neostigmine levels indicate a positive contribution to the predicted risk of early PONV. Neostigmine, a commonly used acetylcholinesterase inhibitor in clinical anesthesia practice, is primarily administered at the end of surgery to reverse residual neuromuscular blockade. Its primary mechanism of action involves the inhibition of acetylcholinesterase, thereby increasing acetylcholine concentrations at cholinergic synapses. This pharmacological effect facilitates the restoration of neuromuscular transmission; however, it concurrently activates central cholinergic pathways and augments gastrointestinal motility.<sup>19</sup> Its mechanism of action involves increasing acetylcholine concentrations at cholinergic synapses. This leads to activation of central cholinergic pathways and enhanced gastrointestinal activity, both of which may contribute to the risk of early PONV. There are several types of receptors associated with emetogenic neurotransmitters, including dopamine (D2) receptors, histaminic (H1) receptors, 5-hydroxytryptamine<sub>3</sub> (5-HT<sub>3</sub>) receptors, and muscarinic cholinergic receptors. Cholinesterase inhibitors, such as neostigmine, in particular, have been linked to an increased incidence of PONV.<sup>19</sup> The underlying mechanism may involve stimulation of the vomiting center in the brainstem, which receives emetic signals from various parts of the body.<sup>20</sup> Furthermore, neostigmine promotes increased gastrointestinal motility and secretions. These effects may cause gastrointestinal discomfort and delayed gastric emptying, both of which are recognized contributors to PONV.<sup>21</sup> Previous studies have recommended avoiding the use of acetylcholinesterase inhibitors to reduce the incidence of postoperative vomiting.<sup>22</sup> As an alternative, sugammadex, a selective binding agent for steroidal neuromuscular blockers, offers a non-cholinergic option for reversal.<sup>23</sup> Several studies have reported a lower incidence of PONV in patients receiving sugammadex compared to those who received neostigmine. Therefore, the use of sugammadex as a reversal agent may be considered, particularly in patients at high risk for PONV.<sup>24</sup> In this research, the use of opioids also contributed positively to the model's prediction, suggesting that this variable substantially increases the likelihood of PONV.

In this research, the history of nausea and vomiting emerged as the most influential variable, with high feature values consistently associated with strongly positive SHAP contributions. This indicates that patients with a prior history are at substantially increased predicted risk of early PONV. Prior history of nausea and vomiting, female sex, and younger age are independently and strongly associated with increased risk of PONV.<sup>20</sup> Demographic factors like age and gender demonstrated a relatively significant influence on model outputs. Most research shows that women and young patients are susceptible to PONV. In the previous meta-analysis study of 22 prospective studies involving 95,154 patients, female gender was identified as the most significant patient-related risk factor for PONV (OR = 2.57, 95% CI: 2.32–2.84). This was followed by a history of PONV, which was also strongly predictive (OR = 2.09; 95% CI: 1.90–2.29).<sup>7</sup> Younger age has been associated inversely with PONV risk. In the meta-analysis, each additional decade of age decreased the risk by approximately 12% (OR per decade = 0.88; 95% CI: 0.84–0.9).<sup>7</sup> These patient-related variables are simple to assess preoperatively and form the foundation of risk stratification models that inform prophylactic antiemetic strategies.

In this study, perioperative pain and opioid use were positively associated with the model's predictions, indicating that these variables substantially increased the likelihood of PONV. PONV has emerged as a clinical concern of comparable importance to postoperative pain and warrants particular attention from anesthesiologists.<sup>20</sup> Opioids are widely employed in perioperative care owing to their potent analgesic properties and efficacy in controlling acute postoperative pain. Despite these benefits, opioid administration has been consistently associated with an increased risk of PONV.<sup>25</sup> The emetogenic effect of opioids is primarily mediated through activation of  $\mu$ -opioid receptors located in the chemoreceptor trigger zone (CTZ) of the area postrema and within the vestibular system.<sup>26</sup> For example, opioids act on chemoreceptors in the area postrema, the vestibular system, and the gastrointestinal tract, where the lipophilicity of individual substances

(eg, morphine vs fentanyl) significantly influences their local concentration within components of the vomiting center.<sup>27</sup> Additionally, opioids exert peripheral actions on the gastrointestinal tract, leading to delayed gastric emptying and enhanced visceral afferent stimulation, both of which contribute to nausea and vomiting. The pharmacokinetic characteristics of opioids, particularly their lipophilicity, further modulate their emetogenic potential. More lipophilic opioids rapidly cross the blood-brain barrier, achieving higher local concentrations within central components of the vomiting center, thereby intensifying their propensity to induce PONV.<sup>28</sup> The development and availability of novel pharmacological agents and anesthetic adjuncts have provided opportunities to reduce the incidence of PONV.<sup>20</sup> This study shows that dexamethasone administration was associated with negative SHAP values, indicating a potential protective effect against early PONV. Dexamethasone administration represents an effective approach for minimizing early PONV in surgical patients.

From the study, the type of surgery and surgery duration also showed modest positive contributions to early PONV. Surgery-related risk factors include both the duration and type of surgery. Surgical duration exceeding 60 minutes, as well as certain types of surgical procedures, are independent risk factors associated with an increased incidence of PONV.<sup>2</sup> Additionally, direct and indirect mechanical effects on the gastrointestinal (GI) tract may further contribute to the development of PONV. In adult patients, the incidence of PONV increases markedly with surgical duration, rising from 2.8% for procedures lasting less than 30 minutes to 27.7% for those exceeding 3 hours.<sup>8</sup> Most gastrointestinal surgeries affect the vagus nerve (tenth cranial nerve), which serves as a major peripheral afferent pathway for triggering PONV.<sup>29</sup> Central nervous system (CNS) and ear, nose, and throat (ENT) surgeries may activate both central and peripheral receptors involved in the emetic reflex. Stimulation of the vagus and glossopharyngeal nerves during these procedures may contribute to the development of PONV.<sup>30</sup> Recognizing high-risk procedures and lengthy operations enables clinicians to implement tailored early PONV prophylaxis, including multimodal antiemetic regimens and opioid-sparing analgesia.

The current work investigated longer anesthesia time and inhalation anesthesia, which were also associated with predicted probabilities of early PONV. Duration of anesthesia and the use of volatile anesthetic agents are widely recognized as significant, quantifiable risk factors for PONV.<sup>7</sup> Longer anesthesia duration is associated with early PONV.<sup>31</sup> In patients who were under anesthesia for more than four hours, the incidence of PONV was six times higher than in those under anesthesia for less than two hours (OR = 6.46, 95% CI = 2.08–20.01,  $p = 0.01$ ).<sup>32</sup> Volatile anesthetics are commonly associated with postoperative nausea and vomiting. The use of volatile anesthetic agents was the strongest anesthesia-related predictor of PONV risk (OR = 1.82, 95% CI = 1.56–2.13).<sup>7</sup> Volatile agents cause PONV by stimulating the release of serotonin in the gastrointestinal tract, increasing vestibular sensitivity, and activating the chemoreceptor trigger zone (CTZ) in the brain.<sup>4</sup> In a study of 1180 patients, volatile anesthetics were identified as the factor most affecting the incidence of emesis within the first two hours after surgery. The use of volatile anesthetics increased PONV in a dose-dependent manner, depending on the specific agent chosen.<sup>33</sup> A reduction in baseline risk factors, particularly through strategies such as the avoidance of volatile anesthetics, has been shown to lower the incidence of PONV.<sup>34</sup> Inhalational anesthesia significantly increases PONV (OR = 2.09; 95% CI = 1.21–3.60;  $p = 0.01$ ) compared to total intravenous anesthesia (TIVA).<sup>34</sup> When general anesthesia is required, the use of propofol for both induction and maintenance reduces the incidence of early PONV.<sup>35</sup> Clinical strategies, such as shortening operative time and employing TIVA, are recommended to mitigate this risk.

In this study, convolutional neural networks (CNNs) and long short-term memory networks (LSTMs) were included primarily as comparative models rather than as approaches expected to provide optimal performance. Given that these architectures are designed to capture spatial or sequential patterns, their applicability to tabular clinical data is limited. Consistent with this consideration, CNNs and LSTMs showed relatively weaker performance compared with models that are more commonly used for structured clinical variables. Therefore, the main findings and conclusions of this study are based on models that are more suitable for tabular data, such as logistic regression and tree-based methods.

The contributions of this study are threefold. First, it represents the first investigation in Taiwan, and among Asian populations, to develop a machine learning based predictive model for early postoperative nausea and vomiting (PONV) risk using routinely available non-invasive clinical and anesthetic data. Second, by incorporating SHAP analysis, the study not only quantified the relative importance of individual features but also enhanced the interpretability of machine

learning models for clinical application. Third, by comparing our results with established literature, this study developed an accurate and interpretable predictive model for early PONV, thereby reinforcing the credibility of machine learning applications in clinical practice. The findings also suggest potential practical implications for anesthesiologists by enabling the early identification of high-risk patients, supporting timely prophylactic antiemetic administration, informing perioperative care planning, and potentially contributing to post-anesthesia care unit (PACU) resource planning.

With the increasing maturity of artificial intelligence applications in the medical field, this study demonstrates the potential of machine learning and SHAP interpretability methods in predicting the risk of early PONV. Potential future clinical integration scenarios include embedding the model into existing hospital anesthesia electronic medical record systems to provide early risk alerts during the postoperative period, thereby facilitating timely clinical decision making by anesthesiologists.

Although the proposed models demonstrated moderate discriminative performance, their current performance does not support immediate clinical application. The contribution of this study should therefore be interpreted as incremental, with its primary contribution lying in the application of explainable machine learning methods to enhance interpretability within a specific clinical context. Future research could involve prospective clinical trials, multicenter data integration, and evaluations of the actual impact of incorporating the model into clinical decision support systems on patient outcomes and healthcare quality. In addition, the surgical type was treated as a single multi-level categorical feature in the present study. Future studies may consider more fine-grained, surgery-specific SHAP analyses. Moreover, the assessment of PONV in this study may be subject to subjective variability, particularly given the retrospective study design. Future studies employing prospective designs and standardized assessment tools may help reduce such variability and improve the reliability of the findings.

## Conclusion

This study represents the first investigation in Taiwan and among Asian populations to develop a machine learning-based predictive model for early postoperative nausea and vomiting risk using routinely available non-invasive clinical and anesthetic data. Multiple modeling approaches, including logistic regression, random forest, and other machine learning architectures, were evaluated, demonstrating that meaningful risk patterns for early PONV can be identified from retrospective clinical data. Explainable machine learning analysis further highlighted clinically relevant factors associated with early PONV risk, including neostigmine use, a history of nausea and vomiting, and procedure-related characteristics, supporting existing clinical understanding of emetogenic mechanisms. From a clinical perspective, this study illustrates the potential value of transparent machine learning models in perioperative risk assessment and individualized anesthetic planning. However, although the proposed models demonstrated moderate discriminative performance, their current performance does not support immediate clinical application. Future work should focus on prospective external validation to evaluate robustness and real-world clinical impact.

## Funding

This research was funded by the National Science and Technology Council, Taiwan (grant number NSTC 113-2222-E-038-001-MY3) and the Taipei Medical University (grant number TMU111-AE1-B30).

## Disclosure

The authors report no conflicts of interest in this work.

## References

1. Shim JH. Is it necessary to use prophylactics for preventing PONV? *Korean J Anesthesiol.* 2011;61(2):105–106. doi:10.4097/kjae.2011.61.2.105
2. Apfel CC, Laara E, Koivuranta M, Greim CA, Roewer N. A simplified risk score for predicting postoperative nausea and vomiting: conclusions from cross-tests between two centers. *Anesthesiology.* 1999;91(3):693–700. doi:10.1097/00000542-199909000-00022
3. Tramèr MR, Fuchs-Buder T. Omitting antagonism of neuromuscular block: effect on postoperative nausea and vomiting and risk of residual paralysis. A systematic review. *Br J Anaesth.* 1999;82:379–386. doi:10.1093/bja/82.3.379
4. Kovac AL. Pathophysiology and risk factors for postoperative nausea and vomiting in adults and children. *BJA Educ.* 2025;25(6):234–239. doi:10.1016/j.bjae.2025.02.003

5. Schlesinger T, Meybohm P, Kranke P. Postoperative nausea and vomiting: risk factors, prediction tools, and algorithms. *Curr Opin Anaesthesiol.* 2023;36(1):117–123. doi:10.1097/ACO.0000000000001220
6. Elvir-Lazo OL, White PF, Yumul R, Cruz Eng H. Management strategies for the treatment and prevention of postoperative/postdischarge nausea and vomiting: an updated review. *F1000Res.* 2020;9:983.
7. Apfel CC, Heidrich FM, Jukar-Rao S, et al. Evidence-based analysis of risk factors for postoperative nausea and vomiting. *Br J Anaesth.* 2012;109(5):742–753. doi:10.1093/bja/aes276
8. Sinclair DR, Chung F, Mezei G. Can postoperative nausea and vomiting be predicted? *Anesthesiology.* 1999;91(1):109–118. doi:10.1097/0000542-199907000-00018
9. Kovac AL. Prevention and treatment of postoperative nausea and vomiting. *Drugs.* 2000;59(2):213–243. doi:10.2165/00003495-200059020-00005
10. Hagan R, Gillan CJ, Mallett F. Comparison of machine learning methods for the classification of cardiovascular disease. *Inf Med Unlocked.* 2021;24:100606. doi:10.1016/j.imu.2021.100606
11. Luo JT, Hung YC, Chen GJ, Lin YS. Predicting early treatment effectiveness in Bell's palsy using machine learning: a focus on corticosteroids and antivirals. *Int J Gene Med.* 2024;17:5163–5174. doi:10.2147/IJGM.S488418
12. Mai RY, Lu HZ, Bai T, et al. Artificial neural network model for preoperative prediction of severe liver failure after hemihepatectomy in patients with hepatocellular carcinoma. *Surgery.* 2020;168:643–652. doi:10.1016/j.surg.2020.06.031
13. Salati M, Migliorelli L, Moccia S, et al. A machine learning approach for postoperative outcome prediction: surgical data science application in a thoracic surgery setting. *World J Surg.* 2021;45(5):1585–1594. doi:10.1007/s00268-020-05948-7
14. Farrokhi F, Buchlak QD, Sikora M, et al. Investigating risk factors and predicting complications in deep brain stimulation surgery with machine learning algorithms. *World Neurosurg.* 2020;134:325–338. doi:10.1016/j.wneu.2019.10.063
15. Zhou CM, Wang Y, Xue Q, et al. Predicting early postoperative PONV using multiple machine-learning- and deep-learning-algorithms. *BMC Med Res Meth.* 2023;23:133. doi:10.1186/s12874-023-01955-z
16. Kim JH, Cheon BR, Kim MG, et al. Postoperative nausea and vomiting prediction: machine learning insights from a comprehensive analysis of perioperative data. *Bioengineering.* 2023;10(10):1152. doi:10.3390/bioengineering10101152
17. Zheng Z, Huang Y, Zhao Y, Shi J, Zhang S, Zhao Y. A machine learning-based prediction model for delayed clinically important postoperative nausea and vomiting in high-risk patients undergoing laparoscopic gastrointestinal surgery. *Am J Surg.* 2024;237:115912. doi:10.1016/j.amjsurg.2024.115912
18. Dunford R, Q S, Tamang E. The pareto principle. *Plymouth Stud Sci.* 2014;7:140–148.
19. Lee OH, Choi GJ, Kang H, et al. Effects of sugammadex vs. pyridostigmine-glycopyrrolate on post-operative nausea and vomiting: propensity score matching. *Acta Anaesthesiol Scand.* 2017;61(1):39–45. doi:10.1111/aas.12813
20. Zhang Z, Wang X. The neural mechanism and pathways underlying postoperative nausea and vomiting: a comprehensive review. *Eur J Med Res.* 2025;30:362. doi:10.1186/s40001-025-02632-1
21. Koyuncu O, Turhanoglu S, Akkurt CO, et al. Comparison of sugammadex and conventional reversal on postoperative nausea and vomiting: a randomized, blinded trial. *J Clin Anesth.* 2015;27:51–56. doi:10.1016/j.jclinane.2014.08.010
22. Jokela R, Koivuranta M. Tropisetron or droperidol in the prevention of postoperative nausea and vomiting. A comparative, randomised, double-blind study in women undergoing laparoscopic cholecystectomy. *Acta Anaesthesiol.* 1999;43:645–650. doi:10.1034/j.1399-6576.1999.430609.x
23. Ju JW, Hwang IE, Cho HY, Yang SM, Kim WH, Lee HJ. Effects of sugammadex versus neostigmine on postoperative nausea and vomiting after general anesthesia in adult patients: a single-center retrospective study. *Sci Rep.* 2023;13(1):5422. doi:10.1038/s41598-023-32730-1
24. Mat NISN, Yeoh CN, Maaya M, Zain JM, Ooi JSM. Effects of sugammadex and neostigmine on post-operative nausea and vomiting in ENT surgery. *Front Med.* 2022;9:905131. doi:10.3389/fmed.2022.905131
25. Kovac AL. Update on the management of postoperative nausea and vomiting. *Drugs.* 2013;73(14):1525–1547. doi:10.1007/s40265-013-0110-7
26. Watcha MF, White PF. Postoperative nausea and vomiting. Its etiology, treatment, and prevention. *Anesthesiology.* 1992;77(1):162–184. doi:10.1097/0000542-199207000-00023
27. Chan KS, Chen WH, Gan TJ, et al. Development and test of a composite score based on clinically meaningful events for the opioid related symptom distress scale. *Qual Life Res.* 2009;18:1331–1340. doi:10.1007/s11136-009-9547-2
28. Horn CC. The medical implications of gastrointestinal vagal afferent pathways in nausea and vomiting. *Curr Pharm Des.* 2014;20(16):2703–2712. doi:10.2174/13816128113199990568
29. Halliday TA, Sundqvist J, Hultin M, Walldén J. Post-operative nausea and vomiting in bariatric surgery patients: an observational study. *Acta Anaesthesiol Scand.* 2017;61:471–479. doi:10.1111/aas.12884
30. Ebrahim Soltani A, Mohammadinasab H, Arbabi S, Godarzi M, Mohtaram R. Prophylactic P6 acupressure, ondansetron, metoclopramide and placebo for prevention of vomiting and nausea after strabismus surgery. *Eur J Anaesthesiol.* 2009;26:129.
31. Emma J, Magnus H, Tomi M, Jakob W. Early post-operative nausea and vomiting: a retrospective observational study of 2030 patients. *Acta Anaesthesiologica Scandinavica.* 2021;65:1229–1239. doi:10.1111/aas.13936
32. Apipan B, Rummasak D, Wongsirichat N. Postoperative nausea and vomiting after general anesthesia for oral and maxillofacial surgery. *J Dent Anesth Pain Med.* 2016;16(4):273–281. doi:10.17245/jdapm.2016.16.4.273
33. Apfel CC, Kranke P, Katz MH, et al. Volatile anaesthetics may be the main cause of early but not delayed postoperative vomiting: a randomized controlled trial of factorial design. *Br J Anaesth.* 2002;88:659–668. doi:10.1093/bja/88.5.659
34. Domene SS, Fulginiti D, Thompson A, et al. Inhalation anesthesia and total intravenous anesthesia (TIVA) regimens in patients with obesity: an updated systematic review and meta-analysis of randomized controlled trials. *J Anesth Analg Crit Care.* 2025;5:15. doi:10.1186/s44158-025-00234-1
35. Visser K, Hassink EA, Bonsel GJ, Moen J, Kalkman CJ. Randomized controlled trial of total intravenous anesthesia with propofol versus inhalation anesthesia with isoflurane nitrous oxide: postoperative nausea and vomiting and economic analysis. *Anesthesiology.* 2001;95:616–626. doi:10.1097/0000542-200109000-00012

## Journal of Multidisciplinary Healthcare

### Publish your work in this journal

The Journal of Multidisciplinary Healthcare is an international, peer-reviewed open-access journal that aims to represent and publish research in healthcare areas delivered by practitioners of different disciplines. This includes studies and reviews conducted by multidisciplinary teams as well as research which evaluates the results or conduct of such teams or healthcare processes in general. The journal covers a very wide range of areas and welcomes submissions from practitioners at all levels, from all over the world. The manuscript management system is completely online and includes a very quick and fair peer-review system. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.

Submit your manuscript here: <https://www.dovepress.com/journal-of-multidisciplinary-healthcare-journal>

**Dovepress**  
Taylor & Francis Group