**SUPPLEMENTARY MATERIAL**

**eTable 1: Data Dictionary**

|  |  |  |
| --- | --- | --- |
| **Variable Name** | **Data Type** | **Description** |
| age\_at\_discharge | Numerical | Age of patient at time of discharge in years. Excludes <18 and >130 years excluded. |
| sex | Categorical | Genotypical sex of patient. ‘Transgender’ and ‘intersex’ collapsed into 'Other'. |
| average\_BMI | Numerical | Average BMI (kg/m²) during encounter. Excludes <5, >200. Missing data imputed with mean. |
| asa\_score | Ordinal | American Society of Anesthesiologists (ASA) co-morbidity score. |
| surgical\_encounter\_type | Categorical | Type of surgical encounter. |
| case\_service | Categorical | Main surgical service. |
| scheduled\_procedure | Categorical | Name of scheduled procedure. Selects general surgery procedures. |
| procedure\_approach | Categorical | Approach for the scheduled procedure (e.g., Open, Laparoscopic). |
| diagnosis\_category | Categorical | ICD10 category of the most responsible diagnosis. Based on most responsible diagnosis. |
| anesthetic\_type | Categorical | Type of anesthesia used. Collapsed into General, Regional, Sedation, Local. |
| day\_of\_year | Numerical | Day of the year on which the case is scheduled. Based on case date. |
| day\_of\_week | Ordinal | Day of the week on which the case is scheduled. Based on case date. |
| weekend\_indicator | Binary | Indicates if the case is scheduled on a weekend. Based on case date. |
| surgical\_location\_hospital | Categorical | Hospital where the surgery takes place. Based on surgical location. |
| surgical\_location\_service | Binary | Surgical specialty of the OR. Based on surgical location. |
| primary\_surgeon\_id | Categorical | Unique identifier for the primary surgeon. Categories with <1% frequency collapsed into 'Other'. |
| scheduled\_surgeon | Binary | Indicates if there is an additional scheduled surgeon (as surgical assist). |
| first\_case\_of\_day | Binary | Indicates if the case is the first scheduled of the day in the assigned OR. |
| last\_case\_of\_day | Binary | Indicates if the case is the last scheduled of the day in the assigned OR. |
| morning\_scheduled | Binary | Indicates if the case is scheduled to start in the morning (AM). Based on scheduled start time. |
| actual\_case\_time\_minutes | Numerical | Duration of the case in minutes. Excludes >8000 minutes. |

**eTable 2: Hyperparameter Optimization Ranges**

|  |  |  |
| --- | --- | --- |
| **Model Type** | **Hyperparameter** | **Optimization Range** |
| Linear Regression | N/A | N/A |
| Ridge Regression | alpha | 1e-5 to 1e2 (log scale) |
| Lasso Regression | alpha | 1e-5 to 1e2 (log scale) |
| SVR | C | 1e-3 to 1e2 (log scale) |
| epsilon | 1e-3 to 1.0 (log scale) |
| kernel | {'linear', 'poly', 'rbf'} |
| Random Forest | n\_estimators | 100 to 1000 |
| max\_depth | 3 to 30 |
| min\_samples\_split | 2 to 10 |
| min\_samples\_leaf | 1 to 4 |
| GBM | n\_estimators | 100 to 1000 |
| learning\_rate | 1e-3 to 1.0 (log scale) |
| max\_depth | 3 to 30 |
| subsample | 0.5 to 1.0 |
| XGBoost | n\_estimators | 100 to 1000 |
| learning\_rate | 1e-3 to 1.0 (log scale) |
| max\_depth | 3 to 30 |
| subsample | 0.5 to 1.0 |
| colsample\_bytree | 0.5 to 1.0 |
| Neural Network | units\_l1 | 64 to 512 |
| activation\_l1 | {'relu', 'leaky\_relu', 'elu'} |
| dropout\_l1 | 0.0 to 0.5 |
| l2\_l1 | 1e-6 to 1e-2 |
| n\_layers | 2 to 5 |
| optimizer | {'adam', 'nadam', 'rmsprop'} |
| learning\_rate | 1e-5 to 1e-2 (log scale) |
| batch\_size | 16 to 128 |

**eFigure 2: K-fold cross validation**

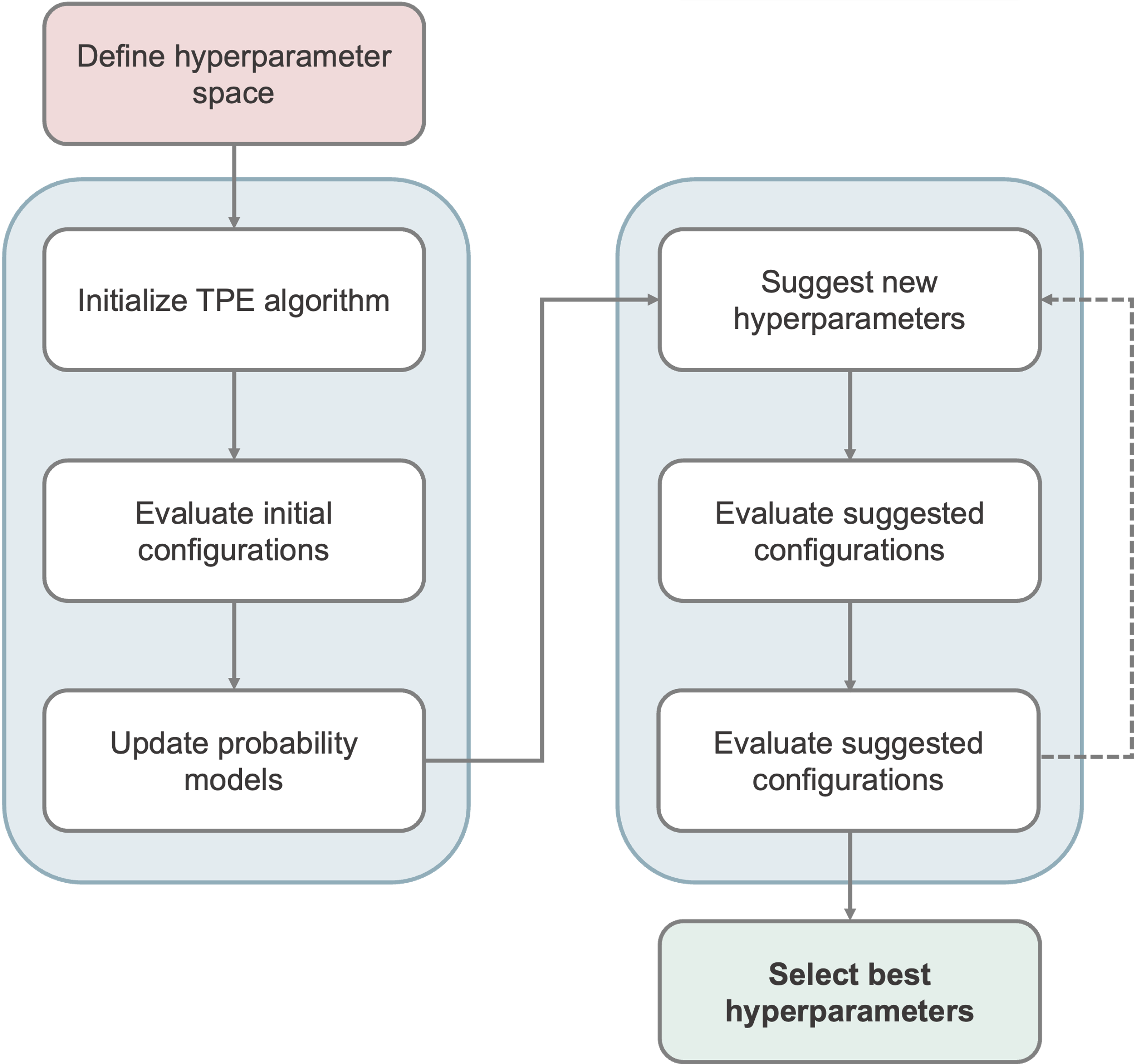
A diagram of a training

AI-generated content may be incorrect.

**eFigure 2.** Schematic of the 10-fold cross validation methodology used in this study for model training and testing. Application of pre-processing steps, hyperparameter tuning and calculation of evaluation metrics were performed indpdently for each fold.

Abbreviations: represents evaluation metrics (i.e., MSE, mean squared error; RMSE, root mean squared error; MAE, mean absolute error; MAPE, mean absolute percentage error; and *R*2, coefficient of determination) calculated across fold

**eFigure 2: Hyperparameter optimization workflow**



**eFigure 2.** Flowchart of the hyperparameter optimization process using the Tree-structured Parzen Estimator (TPE) algorithm, from defining the hyperparameter space to selecting the best configuration.

**Supplementary Material 1: Model Evaluation Metrics**

**Mean Squared Error (MSE)** measures the average of the squares of the errors between the actual and predicted surgery times. MSE is particularly sensitive to larger errors, making it useful for identifying models that minimize significant prediction discrepancies.

Where:

= true case duration

= predicted case duration

*=* number of observations

**Root Mean Squared Error (RMSE):** This metric is the square root of the MSE, providing an error measure in the same units as the case time of surgery (minutes). RMSE is useful for understanding the typical magnitude of errors in the predictions.

**Mean Absolute Error (MAE):** This metric measures the average magnitude of the errors in predicting surgery time, without considering their direction. It is less sensitive to outliers compared to MSE and RMSE, making it a robust measure of model accuracy.

**Mean Absolute Percentage Error (MAPE):** This metric expresses the prediction accuracy as a percentage by measuring the average of the absolute percentage errors between the actual and predicted surgery times. It provides a relative measure of error, which is useful for comparing the accuracy across different scales.

**Coefficient of Determination (***R*2**):** This metric indicates the proportion of variance in the actual surgery times that can be explained by the predictive model. An *R*2 value closer to 1 suggests that the model accurately captures the variability in surgery times.

Where:

= mean of true case duration