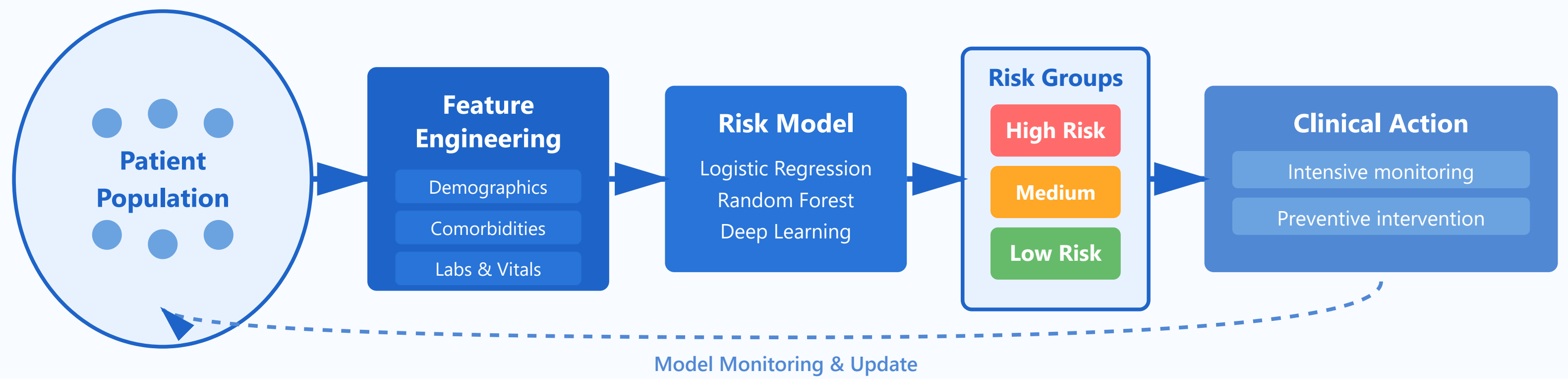
Risk Stratification



📊 Clinical Risk Scores

- CHADS-VASc (stroke risk)
- MELD (liver disease)
- GRACE (cardiac events)
- Point-based scoring systems

🤖 Model Development

- Logistic regression
- Cox proportional hazards
- Gradient boosting machines
- Neural networks

🔧 Feature Engineering

📈 Calibration & Implementation

- Aggregating encounter data
- Temporal patterns
- Medication burden scores
- Comorbidity indices (Charlson, Elixhauser)

- Calibration plots
- Decision curve analysis
- Integration into EHR alerts
- Continuous model monitoring

Detailed Guide to Risk Stratification Components

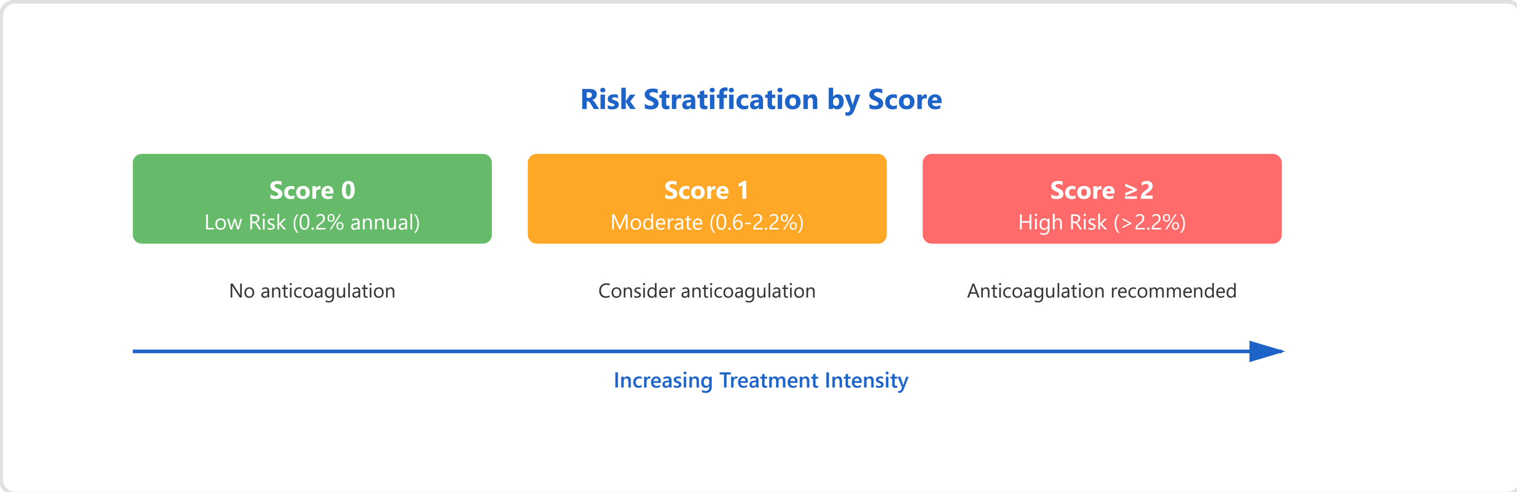
📊 Clinical Risk Scores

Clinical risk scores are validated, point-based systems that quantify patient risk using easily obtainable clinical variables. These scores enable standardized risk assessment across different healthcare settings and support clinical decision-making.

Example: CHA_2DS_2-VASc Score for Stroke Risk in Atrial Fibrillation

Risk Factor	Points
Congestive heart failure	1
Hypertension	1
Age ≥75 years	2
Diabetes mellitus	1

Risk Factor	Points
Stroke/TIA/thromboembolism	2
Vascular disease	1
Age 65-74 years	1
Scex category (female)	1



Example: MELD Score for Liver Disease Severity

$$MELD = 3.78 \times \ln(\text{bilirubin}) + 11.2 \times \ln(\text{INR}) + 9.57 \times \ln(\text{creatinine}) + 6.43$$

The MELD score ranges from 6 to 40 and predicts 3-month mortality in patients with end-stage liver disease. It is used for organ allocation in liver transplantation.

Key Advantages of Clinical Risk Scores:

- Simple and interpretable for clinicians
- Validated across multiple populations
- Easy to calculate at bedside
- Support guideline-based care
- Facilitate risk communication with patients

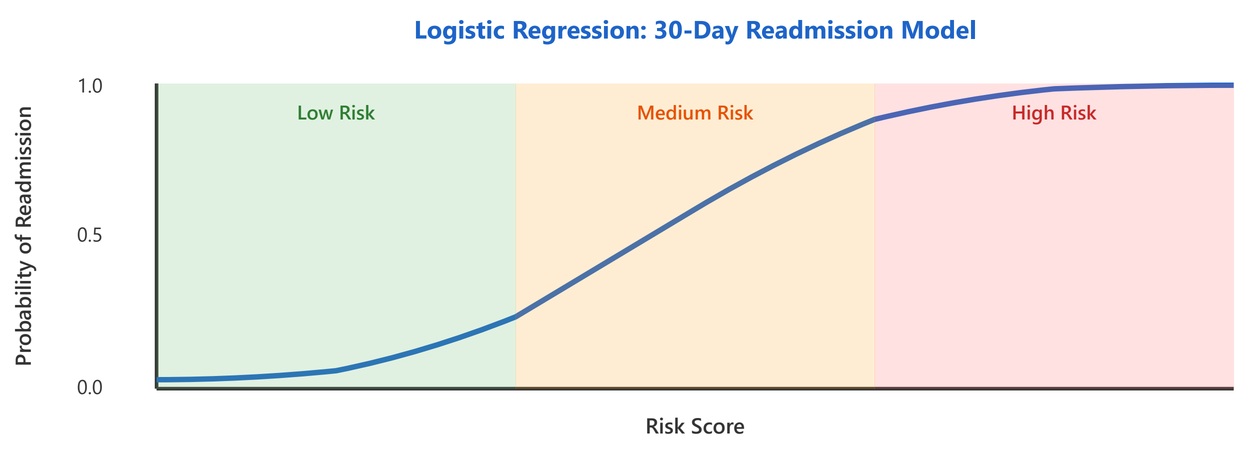
🤖 Model Development

Modern risk stratification extends beyond simple scoring systems by leveraging advanced statistical and machine learning techniques. These models can capture complex interactions between variables and provide more accurate risk predictions.

1. Logistic Regression

The foundation of risk prediction modeling, logistic regression estimates the probability of a binary outcome (e.g., readmission vs. no readmission) based on predictor variables.

$$P(Y=1) = 1 / (1 + e^{-(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + ... + \beta_n X_n)})$$

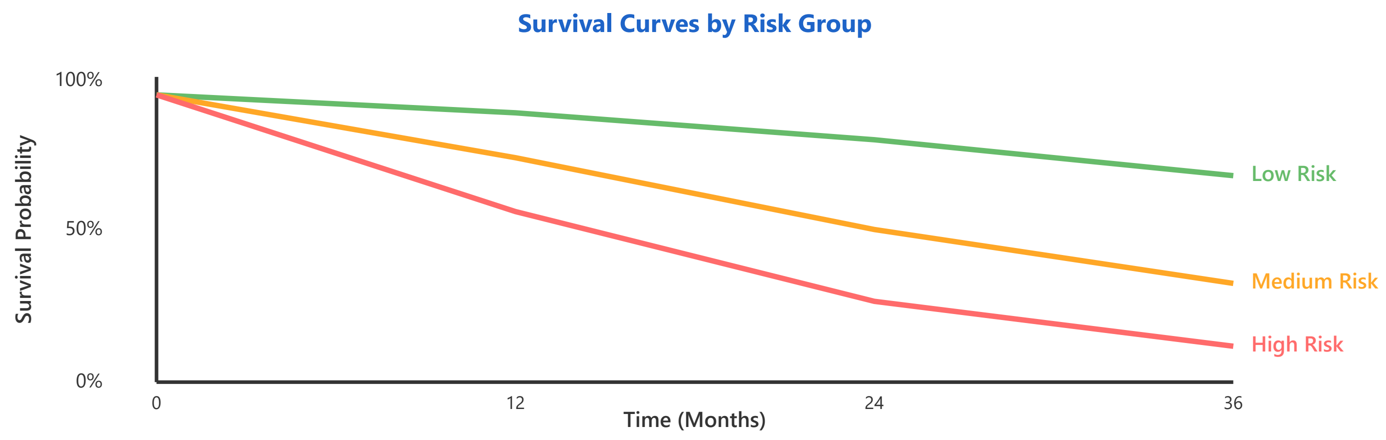


Example Features: Age, number of prior admissions, comorbidity count, length of stay, discharge disposition, lab abnormalities

2. Cox Proportional Hazards Model

Used for time-to-event analysis, Cox models estimate the hazard (instantaneous risk) of an event occurring at any given time, accounting for censoring.

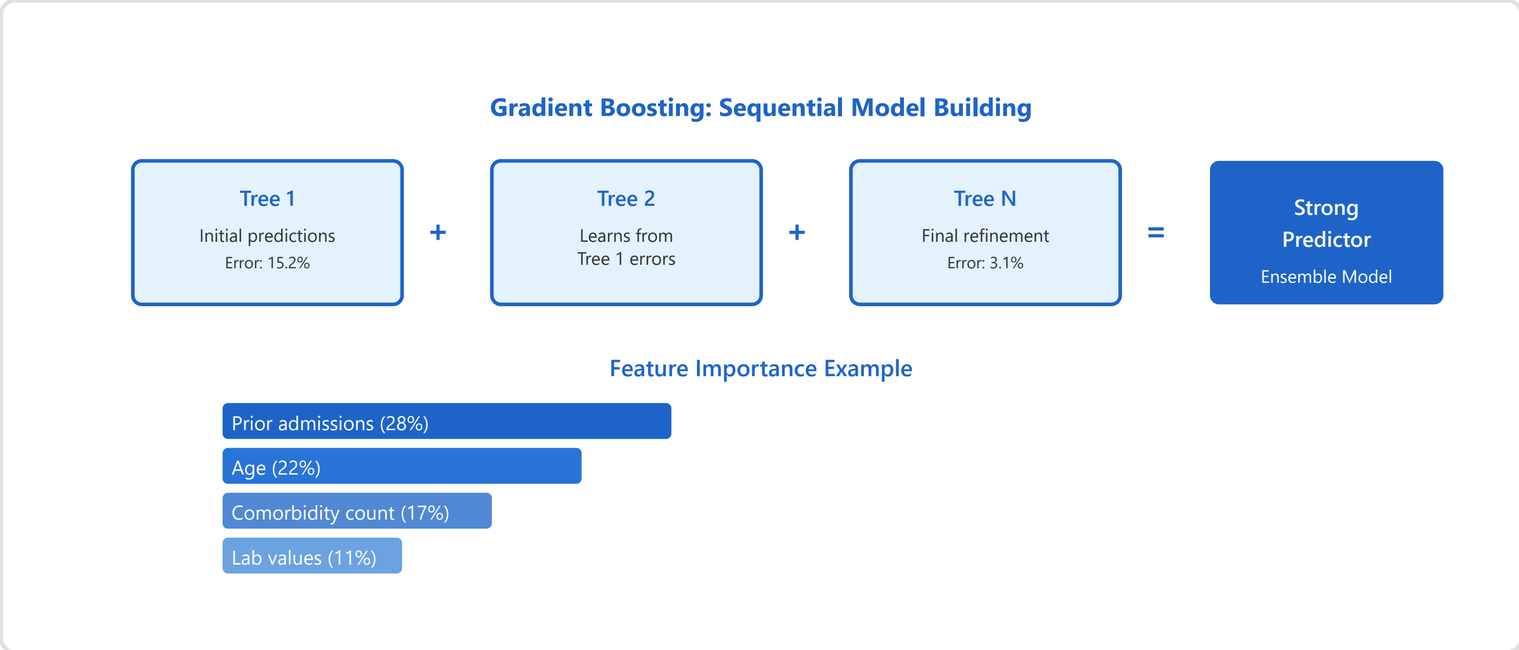
$$h(t) = h_0(t) \times e^{(\beta_1 X_1 + \beta_2 X_2 + ... + \beta_n X_n)}$$



Applications: Mortality prediction, time to disease progression, recurrence-free survival

3. Gradient Boosting Machines (GBM)

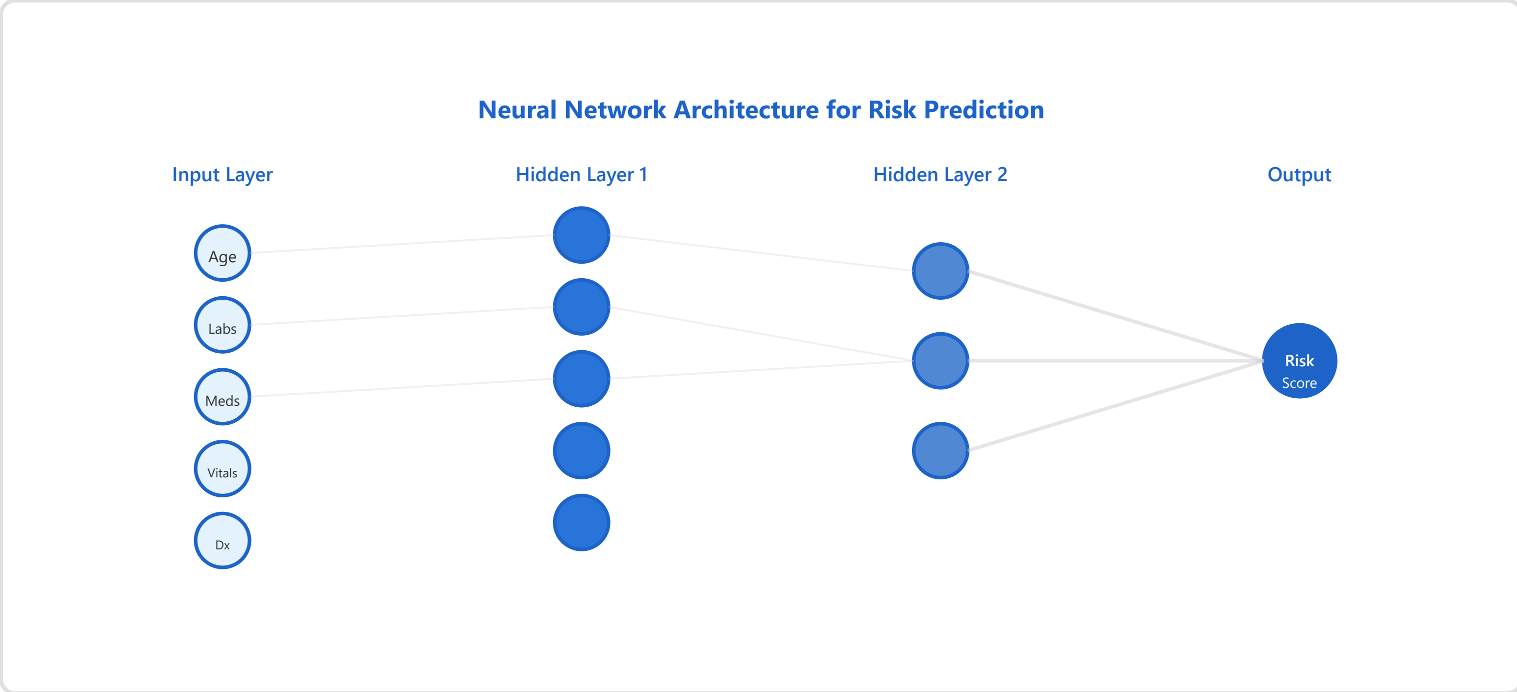
Ensemble learning methods that build multiple weak prediction models (typically decision trees) sequentially, with each model correcting errors from previous models.



Popular Algorithms: XGBoost, LightGBM, CatBoost - widely used in healthcare competitions and production systems

4. Neural Networks (Deep Learning)

Deep learning models can automatically learn hierarchical representations from raw data, particularly effective for unstructured data like medical imaging, clinical notes, and time-series data.



Advanced Architectures: Recurrent Neural Networks (RNNs) for temporal data, Convolutional Neural Networks (CNNs) for imaging, Transformers for clinical notes

Model Selection Considerations:

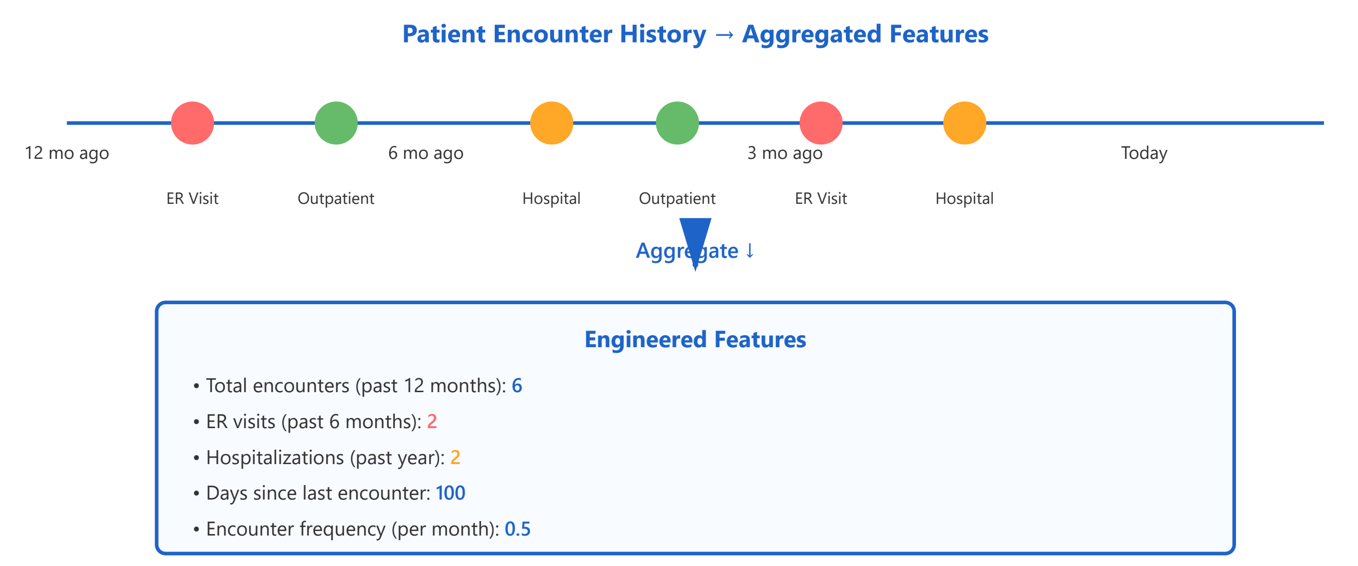
- Interpretability requirements: Logistic regression and decision trees are more interpretable
- Data volume: Deep learning requires large datasets (typically >10,000 samples)
- Feature complexity: Nonlinear relationships favor tree-based or neural methods
- Computational resources: Complex models require more training and inference time
- Regulatory constraints: Some settings require explainable models

🔧 Feature Engineering

Feature engineering transforms raw clinical data into meaningful predictive variables. Well-engineered features can dramatically improve model performance and clinical utility.

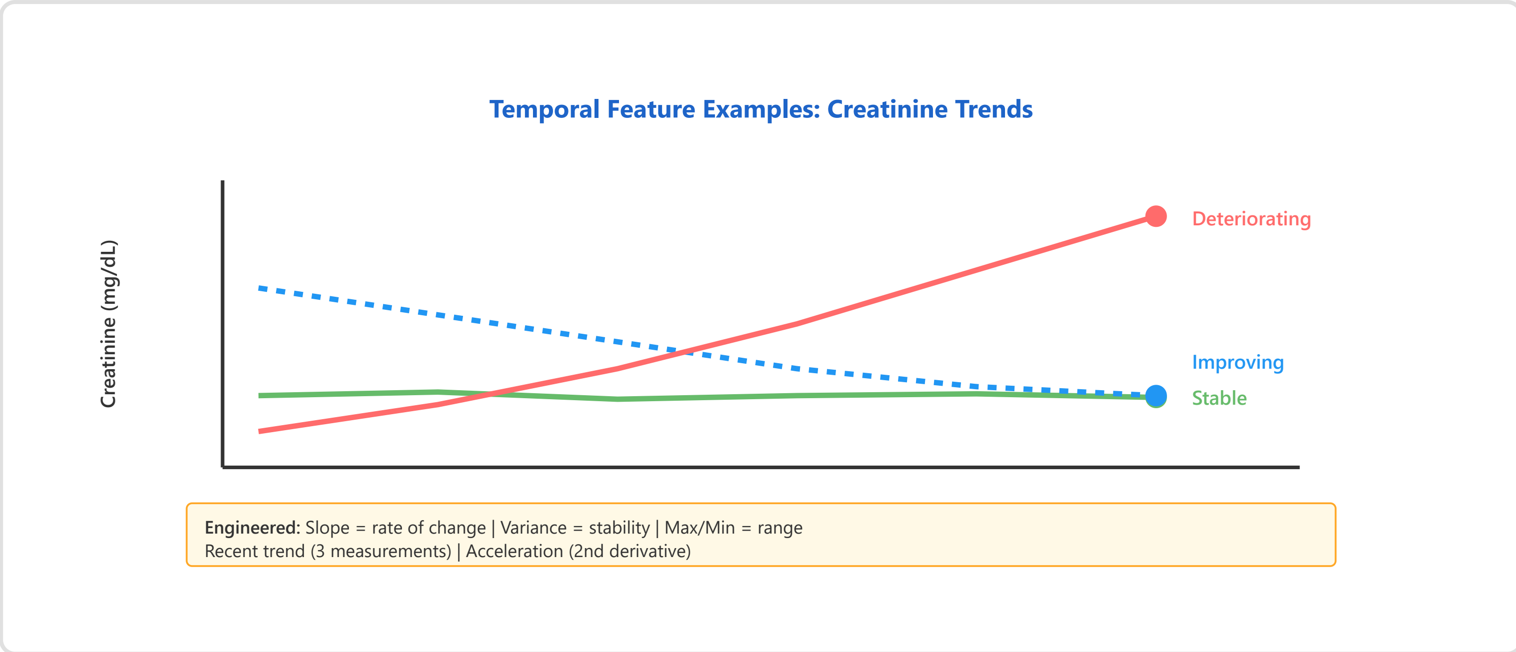
1. Aggregating Encounter Data

Electronic health records contain multiple encounters per patient. Aggregation creates summary features that capture patterns over time.



2. Temporal Patterns

Capturing trends and changes in clinical variables over time often provides stronger predictive signals than single point measurements.



Common Temporal Features: Slope, variance, rate of change, time above/below threshold, crossing frequency

3. Medication Burden Scores

Polypharmacy (multiple medications) increases complexity and risk. Medication burden scores quantify this complexity.

Medication Category	Count	Weight	Score
High-risk medications (anticholinergics, sedatives)	2	×3	6

Medication Category	Count	Weight	Score
Chronic disease medications	5	×1	5
As-needed medications	3	×0.5	1.5
Total Medication Burden Score			**12.5**

Additional Features: Number of medication changes, high-risk drug-drug interactions, medication adherence patterns

4. Comorbidity Indices

Standardized methods to quantify disease burden from diagnosis codes.

Charlson Comorbidity Index Example

Patient Diagnoses

Diagnosis	Points
Diabetes with complications	+2
Chronic heart failure	+1
COPD	+1
Mild liver disease	+1
Peripheral vascular disease	+1
Age adjustment (65-74 years)	+1

Charlson Score: 7

Score Interpretation

- Score 0-1: Low risk
  1-year mortality: ~12%
- Score 2-4: Moderate risk
  1-year mortality: ~26%
- Score ≥5: High risk
  1-year mortality: ~52%

▲ This Patient (Score 7)

Elixhauser Index is an alternative with 31 conditions, often used for inpatient outcomes

Feature Engineering Best Practices:

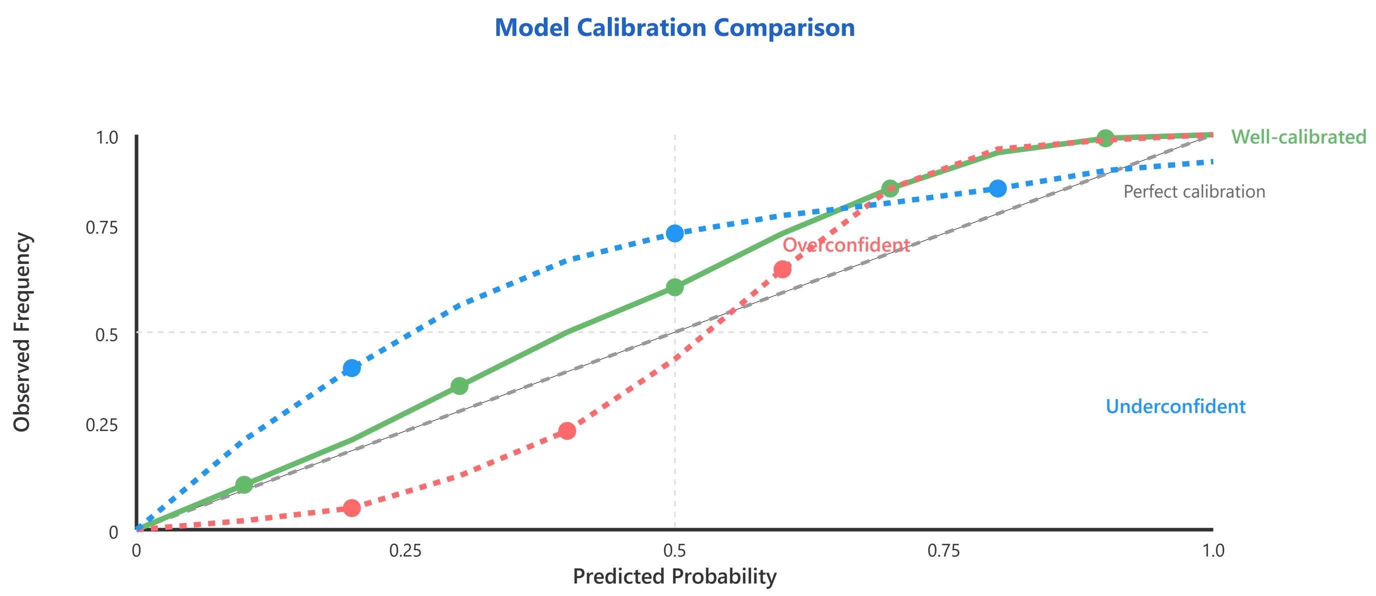
- Domain knowledge integration: Collaborate with clinicians to identify meaningful features
- Handle missing data appropriately: Imputation, missingness indicators, or missing as a category
- Time windows: Choose lookback periods that balance information and clinical relevance
- Feature scaling: Normalize or standardize features for distance-based algorithms
- Avoid data leakage: Only use information available at prediction time

📈 Calibration & Implementation

Developing an accurate model is only the first step. Proper calibration ensures predicted probabilities match observed outcomes, and thoughtful implementation determines real-world clinical impact.

1. Calibration Plots

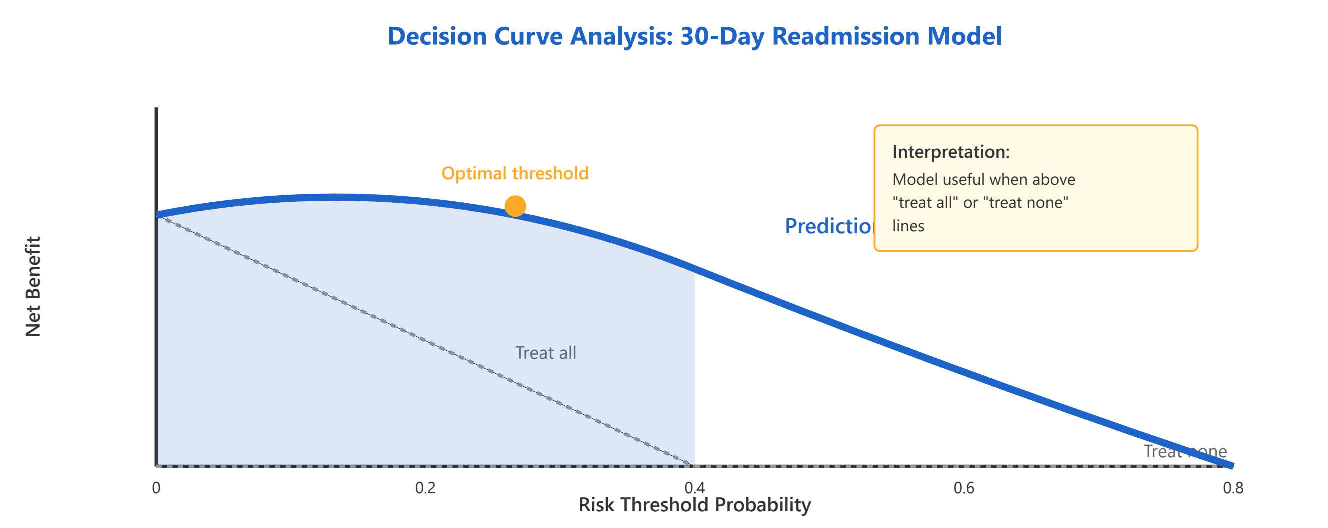
Calibration plots compare predicted probabilities to observed frequencies. A perfectly calibrated model has predictions that match reality.



Interpretation: Overconfident models predict extreme probabilities that don't match reality. Underconfident models cluster predictions around 0.5. Well-calibrated models align with the diagonal.

2. Decision Curve Analysis

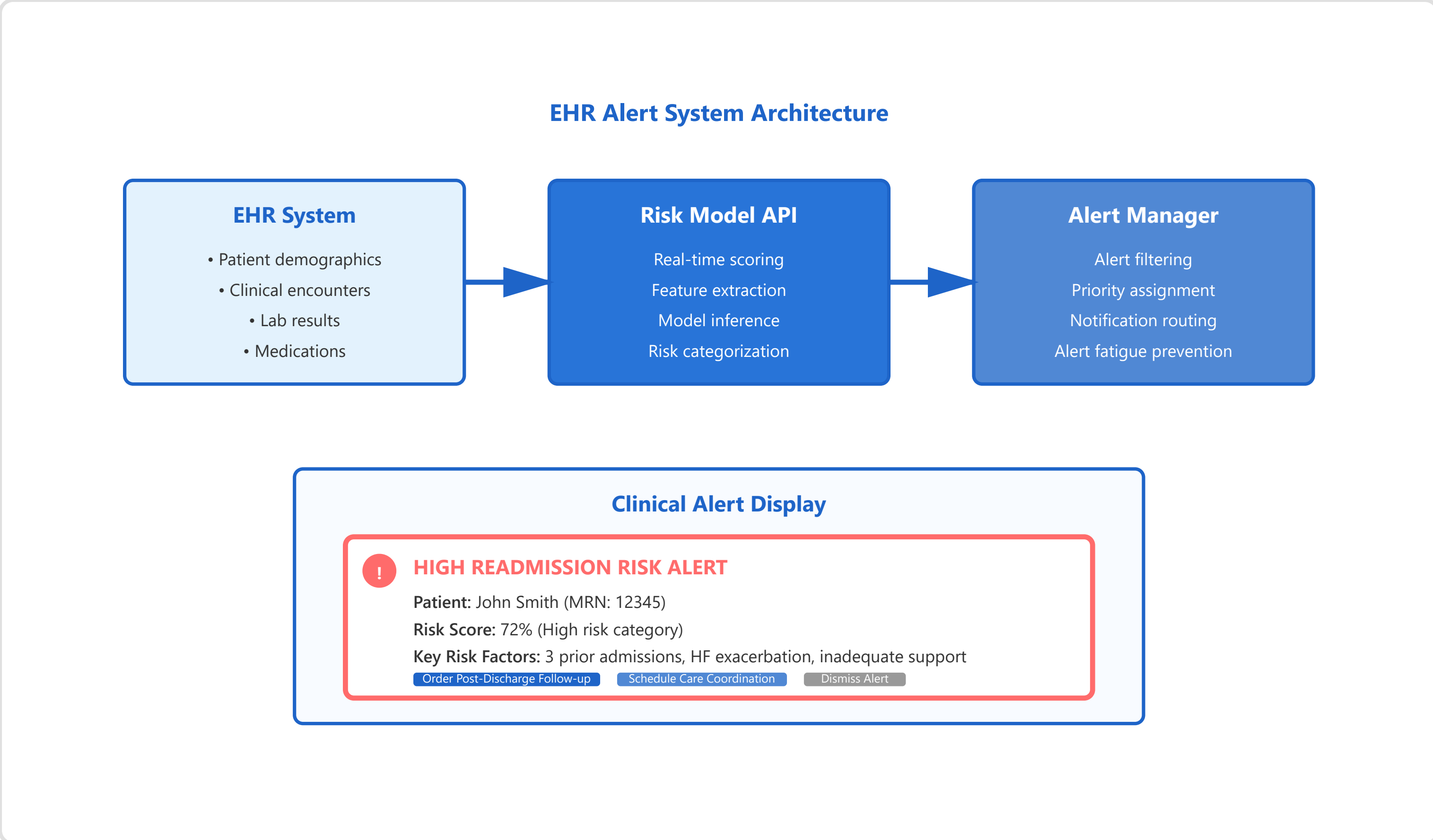
Decision curve analysis evaluates the clinical utility of a prediction model by quantifying the net benefit across different decision thresholds.



Clinical Insight: The model provides maximum net benefit at threshold ~0.25, suggesting interventions should target patients with predicted risk above 25%.

3. Integration into EHR Alerts

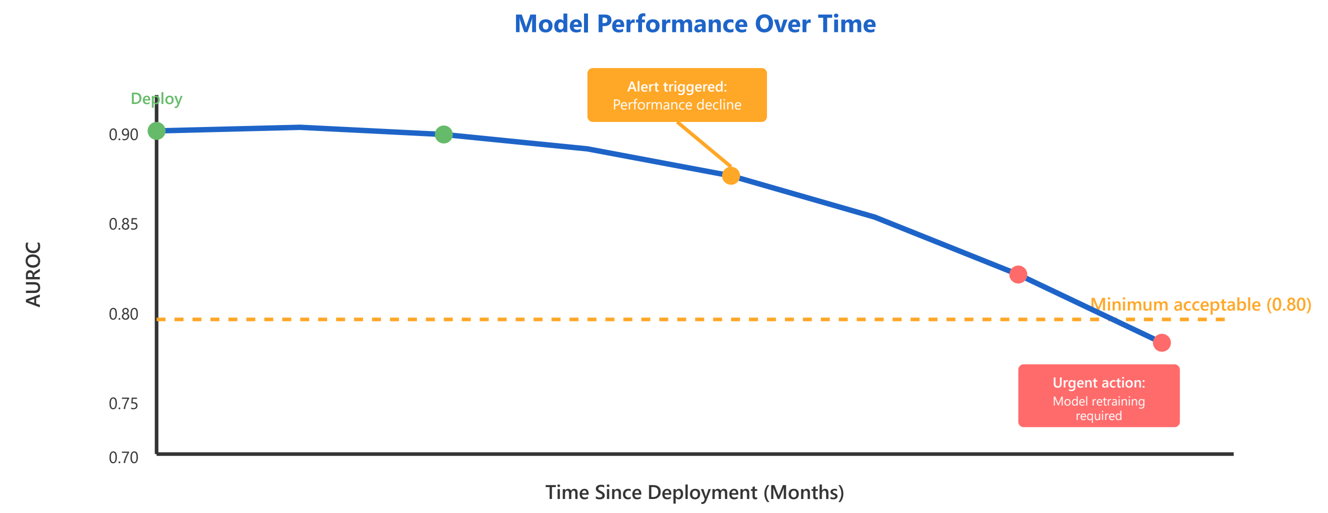
Effective implementation requires seamless integration into clinical workflows through the electronic health record.



Implementation Considerations: Alert timing (admission, discharge, outpatient), actionable recommendations, minimal workflow disruption, override capability

4. Continuous Model Monitoring

Models degrade over time due to population drift, changing clinical practices, and evolving healthcare systems. Continuous monitoring is essential.



Monitoring Metric	Purpose	Action Threshold
Discrimination (AUROC)	Model's ability to separate risk groups	Drop >5% from baseline
Calibration	Predicted vs. observed agreement	Hosmer-Lemeshow p<0.05
Alert volume	Operational sustainability	>15% high-risk alerts
Alert override rate	Clinical utility/acceptance	>40% dismissal rate
Feature drift	Population/practice changes	Distribution shift >20%

Implementation Success Factors:

- Clinician engagement: Involve end-users from design through deployment

- Alert fatigue mitigation: Limit alert frequency, provide actionable recommendations
- Workflow integration: Minimize clicks, align with existing processes
- Transparency: Explain risk factors and model logic to build trust
- Evaluation plan: Define success metrics before deployment (outcomes, usage, satisfaction)
- Governance structure: Establish oversight for model updates and deactivation criteria