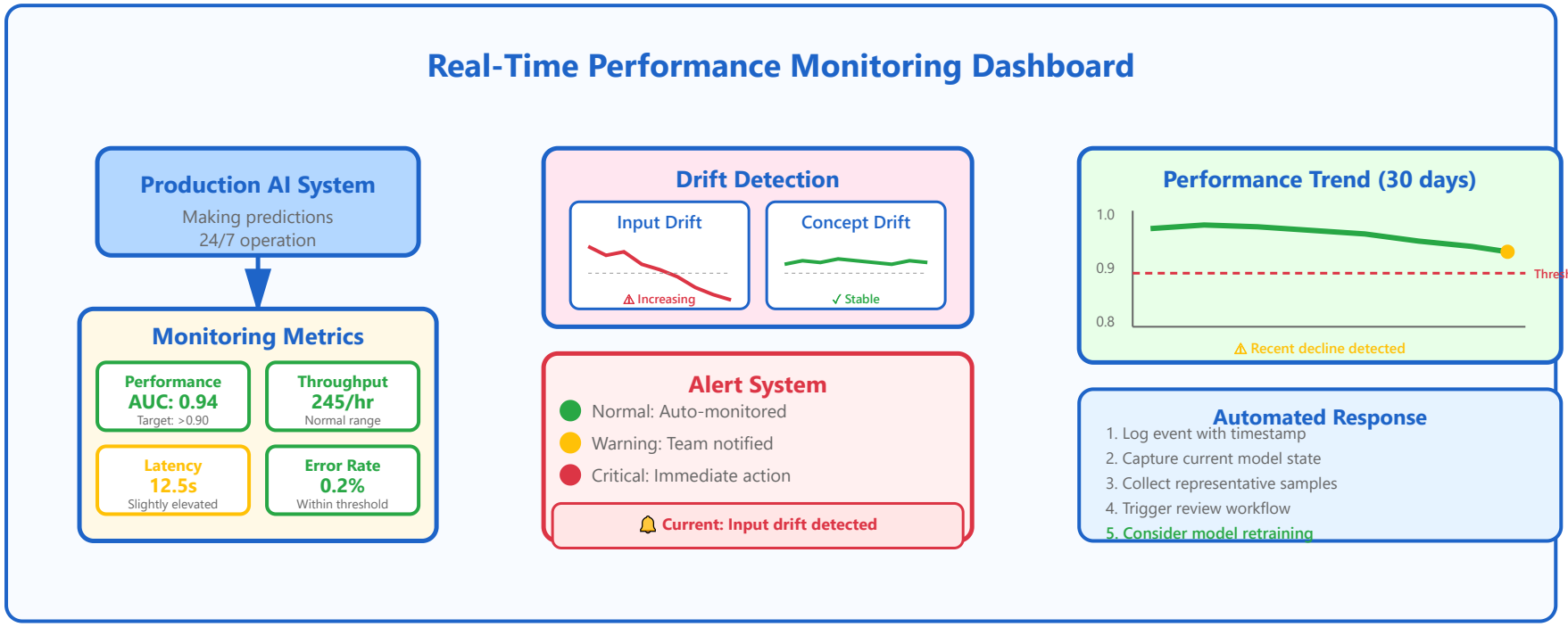
Continuous Monitoring



Real-World Metrics

Sensitivity, specificity in production. Compare to validation performance

Alert Systems

Automated alerts for anomalies. Performance degradation, unusual predictions

Performance Degradation

Detection of model staleness. Dataset shift from new equipment or protocols

Update Strategies

When and how to retrain. Regulatory considerations for algorithm changes

Regulatory Compliance

Documentation for audits. Adverse event reporting to FDA

Detailed Category Explanations

1 Real-World Metrics

Real-world metrics track how your AI model performs with actual clinical data in production environments. These metrics are critical for ensuring that your model maintains its expected performance when deployed in real healthcare settings.

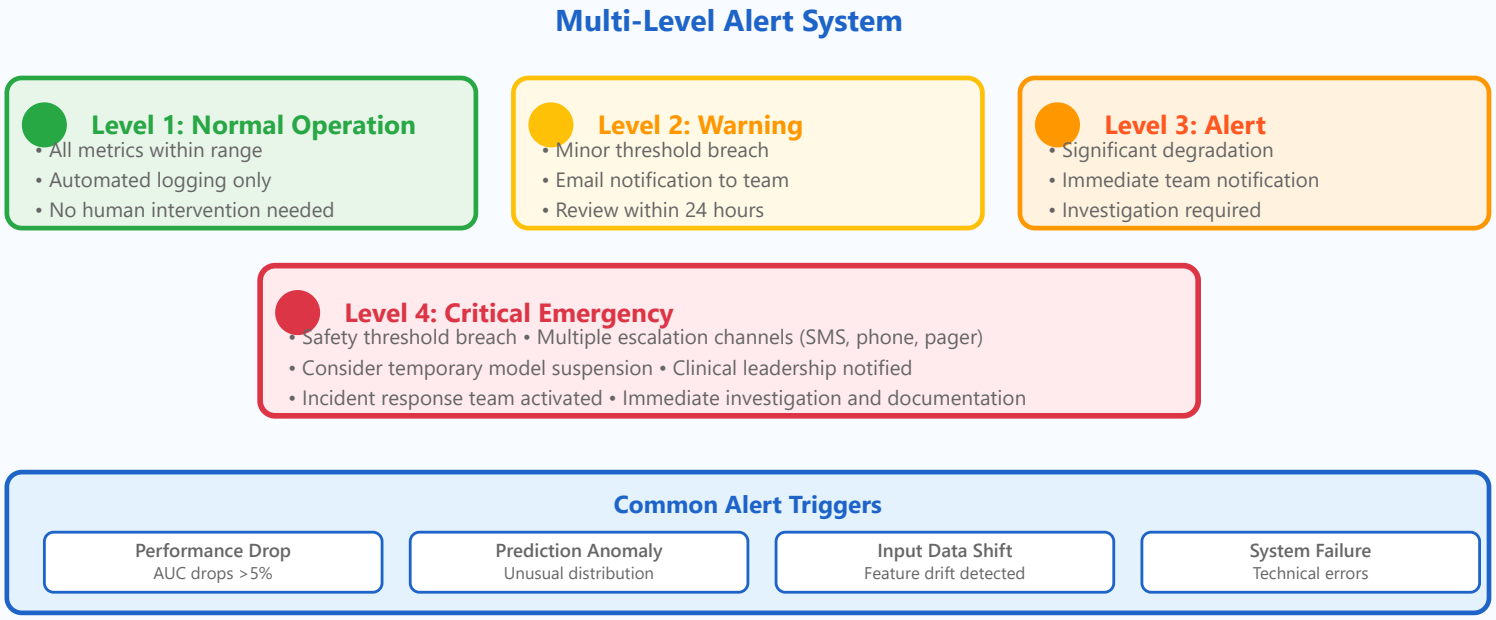
💡 Key Insight: Small performance degradation (3-4%) in production is common due to real-world data variability, but consistent monitoring ensures it stays within acceptable ranges.

Critical Monitoring Points:

- **Ground Truth Collection:** Establish systematic processes for obtaining confirmed diagnoses to validate predictions
- **Temporal Trends:** Monitor daily, weekly, and monthly performance to identify gradual degradation
- **Demographic Subgroups:** Track performance across age, gender, ethnicity to ensure equitable performance
- **Clinical Context:** Evaluate performance in different clinical scenarios (emergency vs. routine screening)
- **False Positive/Negative Analysis:** Regularly review misclassified cases to identify systematic errors

2 Alert Systems

Alert systems provide automated, real-time notifications when your AI model exhibits anomalous behavior or performance issues. These systems act as an early warning mechanism to prevent potential clinical errors.



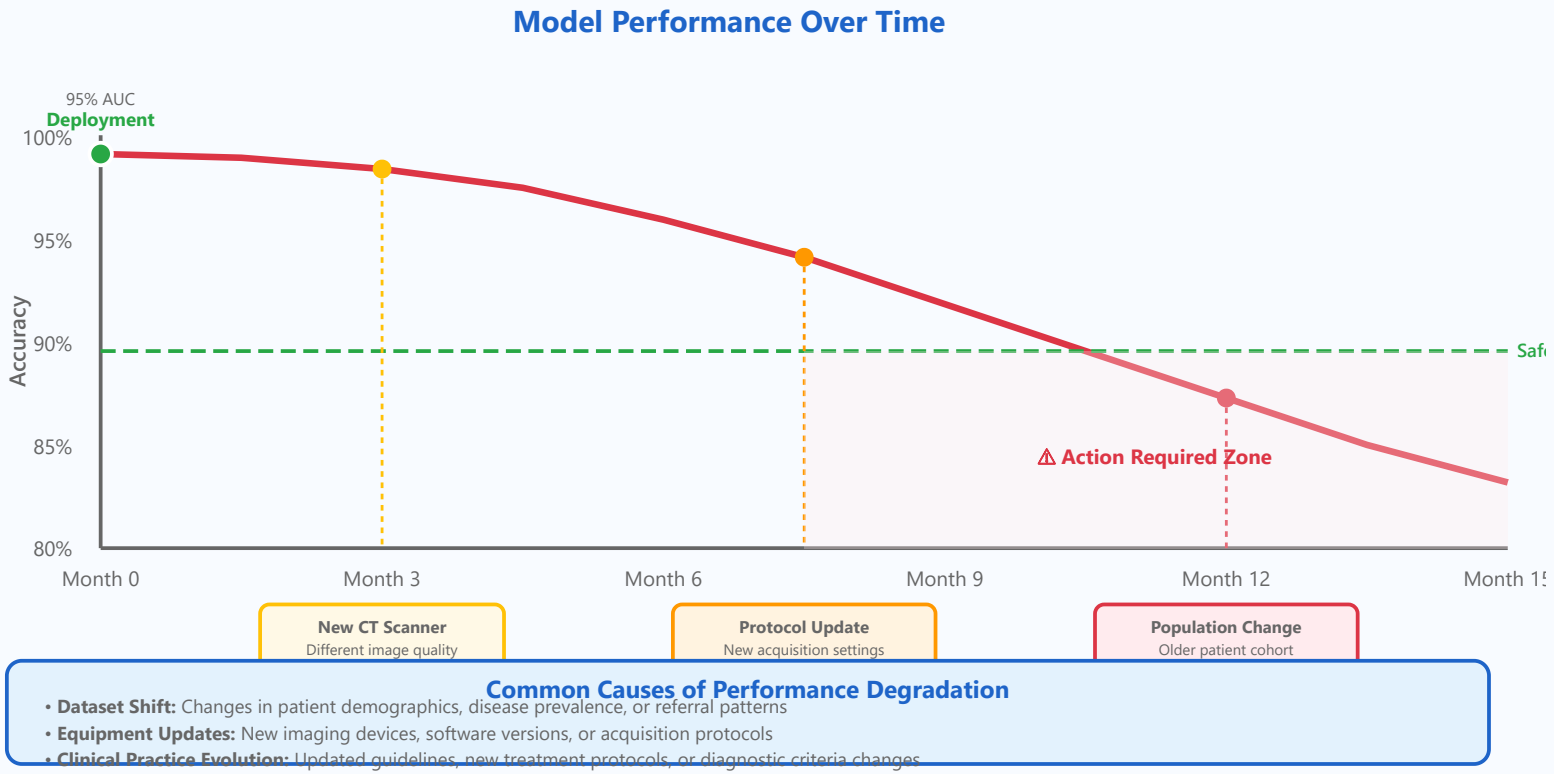
💡 Key Insight: Alert fatigue is real. Set thresholds carefully to minimize false alarms while ensuring genuine issues are caught. Aim for <5 false alerts per month.

Alert System Best Practices:

- **Threshold Calibration:** Set alert thresholds based on clinical significance, not just statistical deviation
- **Multi-Channel Notification:** Use appropriate channels (email for minor, SMS for urgent) based on severity
- **Escalation Protocols:** Define clear escalation paths if alerts are not acknowledged within specified timeframes
- **Alert Suppression:** Implement smart suppression to prevent duplicate alerts for the same underlying issue
- **Documentation:** Automatically log all alerts with timestamps, context, and resolution actions for audit trails
- **Regular Review:** Periodically review alert patterns to refine thresholds and reduce false positives

3 Performance Degradation

Performance degradation occurs when AI models gradually lose accuracy over time due to changes in the real-world data distribution. This phenomenon, known as "model staleness," requires systematic detection and response strategies.



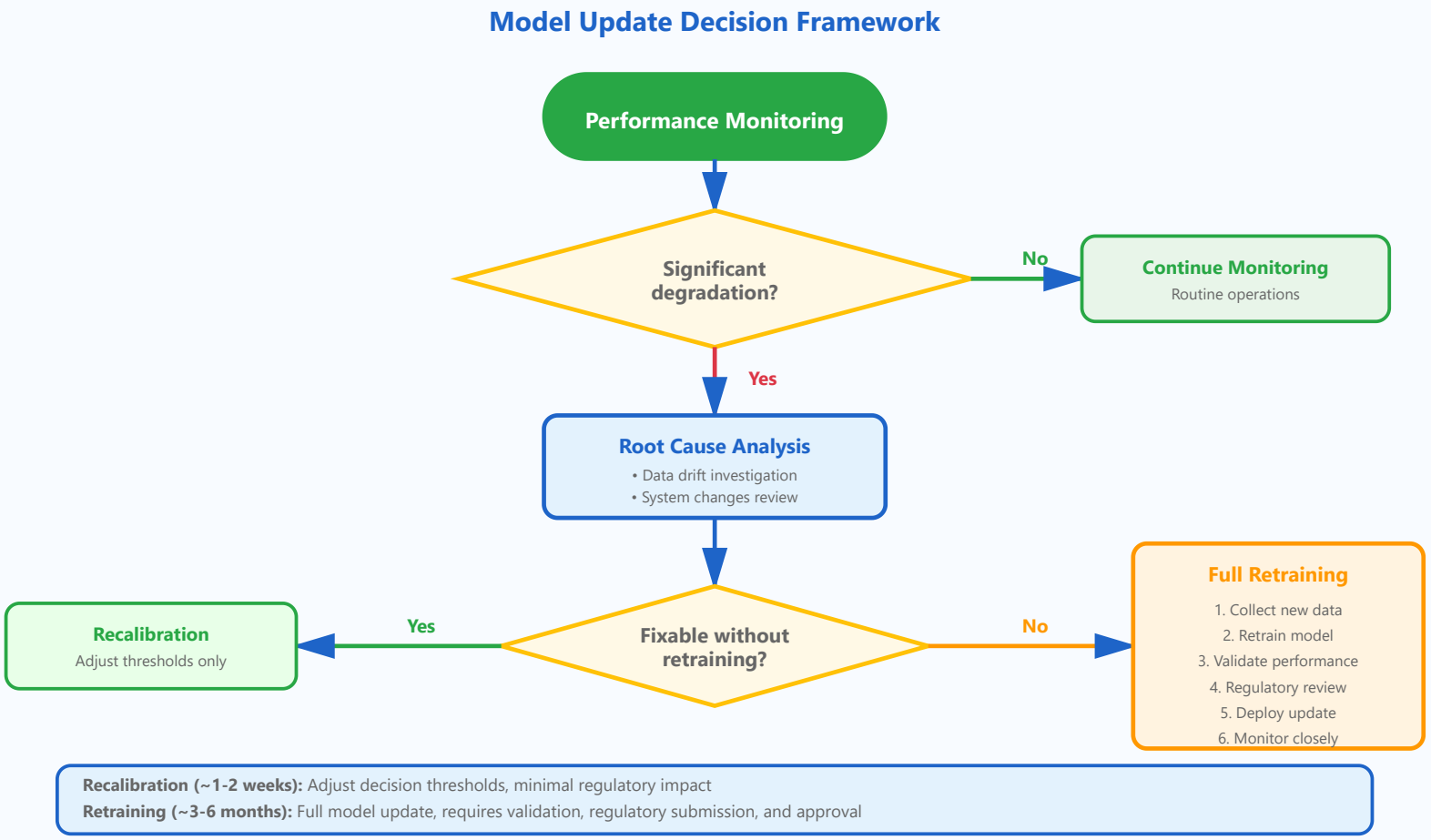
💡 Key Insight: Performance degradation is often gradual and subtle. Regular monitoring with statistical process control charts can detect trends before they become critical issues.

Detection and Response Strategies:

- **Baseline Comparison:** Continuously compare current performance against initial validation metrics
- **Rolling Window Analysis:** Calculate performance over sliding time windows (7-day, 30-day) to detect trends
- **Statistical Tests:** Apply control charts and statistical tests to identify significant changes
- **Root Cause Analysis:** When degradation is detected, investigate potential causes (data shift, technical changes)
- **Mitigation Strategies:** Implement temporary measures (increase review, adjust thresholds) while planning retraining
- **Retraining Decision:** Establish clear criteria for when retraining is necessary vs. when recalibration suffices

4 Update Strategies

Deciding when and how to update AI models in production requires careful consideration of performance metrics, regulatory requirements, and clinical impact. A systematic approach ensures updates improve rather than disrupt care delivery.



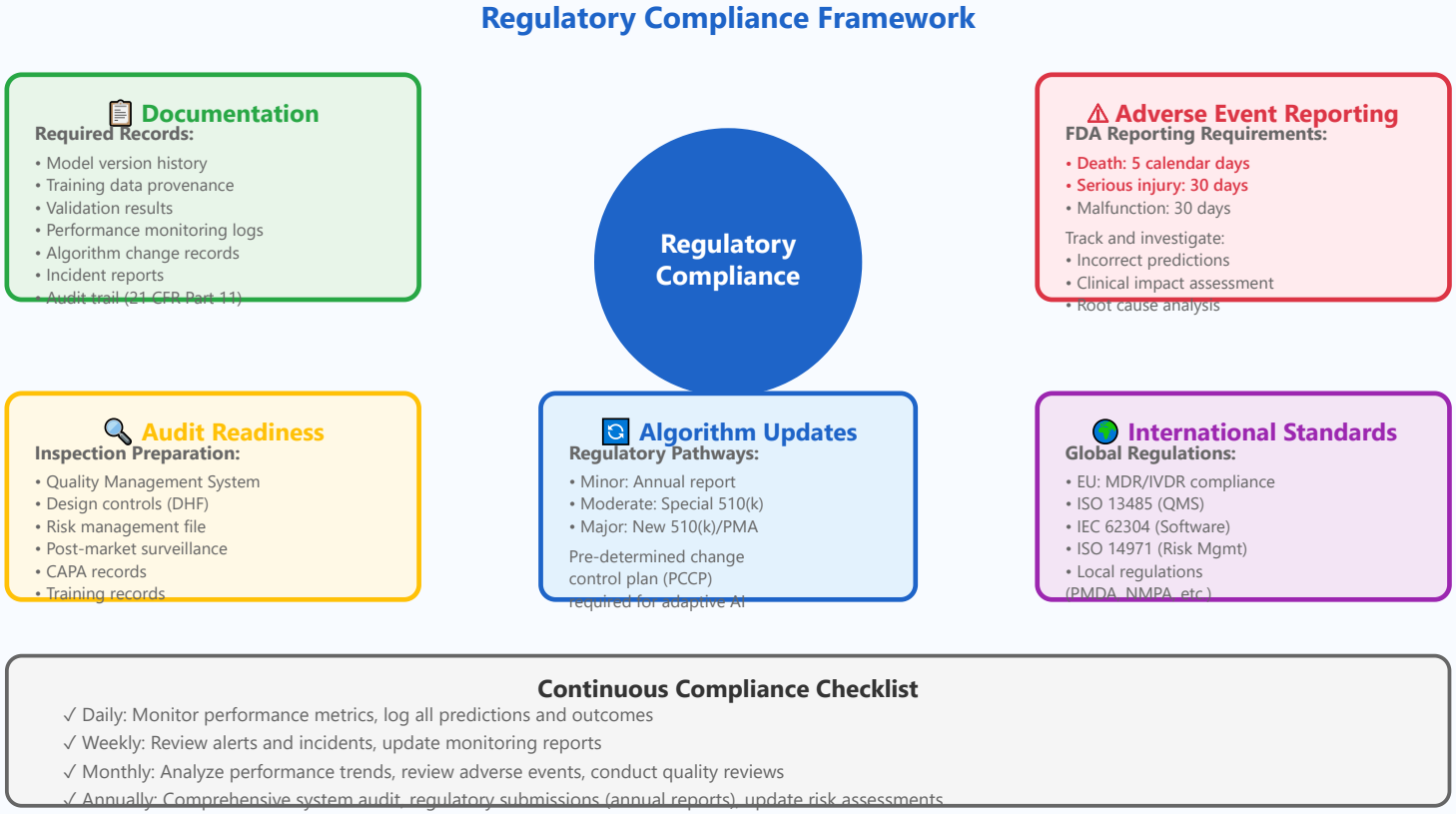
💡 Key Insight: Not all performance issues require full retraining. Simple threshold adjustments (recalibration) can often restore performance while avoiding lengthy regulatory processes.

Update Strategy Considerations:

- **Trigger Criteria:** Define clear performance thresholds that mandate investigation (e.g., AUC drops below 0.90)
- **Data Collection:** Maintain systematic processes for collecting representative training data from production
- **Version Control:** Implement robust model versioning and rollback capabilities in case updates fail
- **A/B Testing:** When possible, deploy updates to subset of users first to validate improvements
- **Regulatory Pathway:** Understand FDA requirements for algorithm changes (510(k) vs. annual report)
- **Stakeholder Communication:** Keep clinical users informed about updates and potential workflow impacts
- **Validation Protocol:** Revalidate updated models using current data before deployment
- **Monitoring Intensification:** Increase monitoring frequency immediately after updates to catch issues early

5 Regulatory Compliance

Maintaining regulatory compliance for AI/ML medical devices requires comprehensive documentation, systematic monitoring, and timely reporting of adverse events. These practices ensure patient safety and satisfy regulatory obligations.



💡 Key Insight: Proactive compliance is easier than reactive compliance. Automated logging and documentation systems should be built into your AI infrastructure from day one.

Compliance Best Practices:

- **Automated Logging:** Implement systems that automatically capture all required documentation without manual intervention
- **Change Control:** Establish formal processes for evaluating and documenting all system changes
- **Traceability:** Maintain complete traceability from requirements through deployment and monitoring
- **Incident Management:** Create clear workflows for detecting, investigating, and reporting adverse events
- **Periodic Review:** Schedule regular compliance reviews to identify and address gaps proactively
- **Training Programs:** Ensure all team members understand their regulatory responsibilities
- **Vendor Management:** If using third-party components, maintain documentation of their regulatory status

- **International Considerations:** Plan for compliance with regulations in all markets where device will be sold