AI Post-Market Surveillance: Locked vs. Continuous Learning

2/25/2026 • 60 min read

post-market surveillance | ai medical devices | continuous learning | locked algorithms | algorithmic drift | fda regulation | samd | medical device validation | machine learning

© 2026 IntuitionLabs.ai - North America's Leading AI Software Development Firm for Pharmaceutical & Biotech. All rights reserved.

Executive Summary

Post-market surveillance (PMS) of medical devices is a critical component of ensuring ongoing safety and effectiveness after regulatory clearance. In the rapidly evolving field of artificial intelligence (AI)–enabled medical devices (AIaMDs), PMS faces unique challenges and opportunities. AIaMDs may be designed as **locked algorithms**, fixed once deployed, or as **continuous learning systems** that adapt and improve with new data over time. This report provides a detailed, evidence-based analysis of PMS for AI medical devices, with a focus on the contrast between continuous learning and locked models. Our executive summary highlights key findings:

- **Definitions and Scope:** We define *post-market surveillance* as the activities by which manufacturers and regulators collect and analyze real-world data on device performance after market entry (www.who.int). *Locked AI models* are algorithms "in a locked state when changes are not permitted" ([1] innolitics.com), while *continuous learning AI* refers to systems that retrain or update with each new data point during operation ([2] innolitics.com).

- **Regulatory Context:** Global regulators (FDA, EU, IMDRF, WHO, etc.) recognize the novelty of adaptive AI and are developing frameworks. The FDA has over 950 AI/ML-enabled devices authorized by mid-2024 ([3] www.healthcaredive.com), and recently issued draft guidance requiring manufacturers to describe plans for postmarket performance monitoring of AI devices ([4] www.fda.gov). In contrast, the EU's Medical Device Regulation (MDR) and In Vitro Diagnostic Regulation (IVDR) currently treat AI as static software, requiring re-certification for any significant updates ([5] goodlifesci.sidley.com) ([6] pmc.ncbi.nlm.nih.gov). Notably, the FDA's proposed *predetermined change control plan* (SaMD Pre-Specifications and Algorithm Change Protocol) would allow pre-specified continuous updates within a defined plan ([7] pmc.ncbi.nlm.nih.gov) ([8] www.ncbi.nlm.nih.gov).

- **Opportunities and Risks of Continuous Learning:** Continuous learning AI can adapt to new patient data, potentially improving accuracy and personalization over time ([9] www.ncbi.nlm.nih.gov). However, this brings new risks: *catastrophic forgetting* (where new data overrides previous knowledge), overfitting to non-representative cases, and generalization failures ([9] www.ncbi.nlm.nih.gov). Surveillance must therefore include ongoing evaluation of effectiveness, safety, and bias as the model evolves ([10] www.ncbi.nlm.nih.gov) ([11] pmc.ncbi.nlm.nih.gov). The unpredictability of continuous updates raises questions of "when and how" to re-evaluate a device's performance ([10] www.ncbi.nlm.nih.gov). Regulators and developers emphasize the need for rigorous *cyclic revalidation*, robust monitoring plans, and transparent reporting.

- **Challenges of Locked Models:** Locked AI devices are simpler to validate premarket but can become **outdated**. Their performance may "drift" if underlying healthcare practices or patient populations change ([12] www.ncbi.nlm.nih.gov). Locked algorithms require periodic re-evaluation and updates (often via new regulatory submissions) to maintain safety ([12] www.ncbi.nlm.nih.gov). Real-world evidence shows limitations: for example, the FDA-cleared IDx-DR retinopathy detector (a locked system) could not analyze 26.1% of real-world patient images, particularly in cases of small pupils ([13] pmc.ncbi.nlm.nih.gov). Continuous learning could have gradually improved in challenging cases like these, highlighting a trade-off between static validation and adaptability.

- **Post-Market Surveillance Methods:** Traditional PMS tools (adverse event reporting, registries, and device tracking) apply to AIaMDs, but must be expanded. Methods such as **active performance monitoring** (collecting key performance metrics during routine use), real-world data analysis, user feedback loops, and specialized registries are needed ([4] www.fda.gov) ([11] pmc.ncbi.nlm.nih.gov). The FDA enforces mandatory tracking and malfunction reporting, and may require post-approval studies or real-world evidence projects ([14] www.fda.gov). WHO guidance stresses systematic data collection and evaluation of real-world use to trigger corrective actions when needed (www.who.int). For AIaMDs, surveillance should include monitoring for *algorithmic drift*, emergent biases, cybersecurity incidents, and degradation of performance over time.

- **Case Studies:** We review illustrative examples. The AI-based skin cancer app *SkinVision* (CE-marked, locked model) significantly over-detected benign lesions as high-risk in a real-world evaluation (41–83% sensitivity, 60–83% specificity), leading to many false alarms and low user trust ([15] pmc.ncbi.nlm.nih.gov). IDx-DR (AI for diabetic retinopathy) was cleared by FDA, then updated with a new "training mode" via a separate 510(k) almost three years later ([16] pmc.ncbi.nlm.nih.gov). The HeartFlow FFR_{CT} AI for coronary artery disease has faced regulatory hurdles in Europe because its constantly-updated algorithm "could not be determined at a single point in time, as requested by MDR" ([6] pmc.ncbi.nlm.nih.gov). These cases underscore that real-world performance can diverge from trial settings, and that locked devices often require formal re-approval for updates. Emerging examples of continuous-learning in practice remain rare, but regulators and industry are actively studying frameworks – for example, FDA's Total Product Lifecycle (TPLC) approach.

© 2026 IntuitionLabs.ai - North America's Leading AI Software Development Firm for Pharmaceutical & Biotech. All rights reserved.

- **Data and Evidence:** The proliferation of AI devices is rapid: in 2015 FDA cleared only 6 AI devices, versus 221 in 2023 ([17] www.healthcaredive.com). Radiology constitutes the largest share of AI/ML medical devices ([18] pmc.ncbi.nlm.nih.gov). However, analyses find persistently limited clinical data on many AI devices. A RAPS study found most cleared AI/ML devices lacked robust prospective postmarket data . Drifts in real-world performance can be subtle, making bias and confounding key issues ([11] pmc.ncbi.nlm.nih.gov). Experts emphasize evidence-based monitoring: e.g., pulling model outputs ("withholding") or advanced causal analysis to detect true performance changes ([19] pmc.ncbi.nlm.nih.gov).
- **Future Directions:** The field is moving towards **proactive lifecycle governance**. The EU AI Act (effective 2026) will classify medical AI as "high-risk," imposing post-market requirements (quality management, incident/data logging, monitoring plans) even if MDR lacks specifics ([20] pmc.ncbi.nlm.nih.gov) ([5] goodlifesci.sidley.com). The FDA and global bodies are pushing for "good machine learning practices" (GMLP) and new guidelines on periodic re-validation. Big emerging themes include use of Real-World Evidence (RWE) to inform updates, advanced surveillance analytics (e.g., machine learning for signal detection ([21] jamanetwork.com)), and harmonized international standards (IMDRF, WHO). Manufacturers are advised to build PMS into device design (e.g. performance logging, human oversight mechanisms) and to engage regulators early on lifecycle plans ([4] www.fda.gov) ([8] www.ncbi.nlm.nih.gov).

In conclusion, effective post-market surveillance of AI medical devices requires balancing innovation with patient safety. Locked and continuous-learning models present different risk profiles: locked models need safeguards against time-dependent degradation ([12] www.ncbi.nlm.nih.gov), while continuous models require rigorous control of unanticipated changes ([10] www.ncbi.nlm.nih.gov). This report explores scientific, regulatory, and real-world perspectives in depth, providing data, expert consensus, and case analysis. It highlights emerging solutions such as algorithm change protocols, real-world performance monitoring plans, and collaborative registries. Robust PMS frameworks will be crucial for the next generation of AI medical devices to ensure they remain safe, effective, and equitable as technology and data evolve.

Introduction and Background

The last decade has seen an **explosive growth in AI-enabled medical devices (AIaMDs)**. According to FDA data, approximately 950 AI/ML-enabled medical devices were cleared by U.S. regulators between 1995 and mid-2024 ([3] www.healthcaredive.com). The approval rate increased sharply – only 6 devices were cleared in 2015 versus over 220 in 2023 ([17] www.healthcaredive.com). Radiology algorithms dominate the market, reflecting the early adoption of AI in medical imaging ([18] pmc.ncbi.nlm.nih.gov). Fields like cardiology and pathology are the next frontier. These devices span a broad range: from clinical decision-support tools and diagnostic engines (e.g. automated image analysis) to therapy guidance (e.g. closed-loop insulin delivery) and patient-facing screening apps.

With rapid innovation comes a critical need for **post-market surveillance (PMS)**. PMS is defined, for medical devices, as the activities through which manufacturers and regulators collect and evaluate experience gained from devices that have been marketed, in order to identify safety or performance issues (www.who.int). The goal of PMS is to ensure that devices continue to meet safety and performance requirements throughout their life cycle, and to trigger corrective actions if risks arise. The 2021 WHO global guidance on medical device surveillance states:

> *"Post-market surveillance is a set of activities conducted by manufacturers to collect and evaluate experience gained from medical devices that have been placed on the market, and to identify the need to take any action"* (www.who.int).

PMS for traditional medical devices involves collecting data from adverse event reports, device tracking systems, registries, observational studies, and user feedback. Regulatory mandates vary: for example, the EU Medical Device Regulation (MDR 2017/745) requires manufacturers to implement PMS plans, monitoring, and periodic safety update reports, while in the U.S. FDA enforces adverse event reporting (MDR), device tracking, and may require post-approval studies or section 522 surveys for certain devices ([14] www.fda.gov). AIaMDs introduce **new complexities** to PMS. Unlike classical devices, AI algorithms (especially those based on machine learning) can "learn" from data and change their behavior over time. This opens the door to concept drift (changes in input data), model drift (changes in performance), and emergent biases, all of which may impact safety and effectiveness.

© 2026 IntuitionLabs.ai - North America's Leading AI Software Development Firm for Pharmaceutical & Biotech. All rights reserved.

A key conceptual distinction arises in how an AIaMD is designed to operate post-deployment:

- A **Locked Algorithm** (or *fixed* model) is deployed in a static form. Once the device is on the market, the AI model's parameters remain unchanged until the manufacturer issues a new version (through another regulatory submission). In IMDRF terminology, an ML-enabled device is *"in a locked state when changes are not permitted"* ([1] innolitics.com). A locked model may still be updated by the sponsor later (batch retraining), but it is not continually adapting in real time. Examples include most current FDA-cleared AI devices: they operate on fixed weights, and any planned update must go through formal review or clearance.
- A **Continuous-Learning AI** system is designed to adapt continuously as it receives new data during normal use. IMDRF defines *"Continuous Learning"* as *"training that leads to change of an MLMD (Machine Learning-enabled Medical Device) with each exposure to data that takes place on an ongoing basis during the operation phase of the MLMD life cycle"* ([2] innolitics.com). Put differently, every new patient case or set of images could be used to retrain or fine-tune the model incrementally (often via online learning algorithms). Continuous learning is also called *"unlocked"* or *"adaptive"* AI. For example, a radiology AI could in principle refine its detection thresholds each time it processes a new scan.

These concepts are distinct from *batch learning*, which lies between them. Batch learning (also termed *"piecemeal"* or *"offline updates"*) involves periodic retraining on a fixed dataset at discrete times (e.g. monthly or annually) separated from normal operation ([22] innolitics.com). In practice, most updates today happen via batch submissions (the device is locked during operation, then updated in batches after regulatory review).

The choice between locked vs continuous has profound implications for PMS. Continuous learning offers the promise of improving over time with new data, potentially catching new disease patterns or adapting to shifting populations ([9] www.ncbi.nlm.nih.gov). However, it also introduces unpredictability: the algorithm's decision boundaries may shift in complex ways, raising concerns about repeatability, validation, and oversight ([9] www.ncbi.nlm.nih.gov) ([10] www.ncbi.nlm.nih.gov). Locked models, by contrast, offer stability (the device behavior is fixed and typically well-characterized at approval) ([1] innolitics.com), but risk degradation if the real-world environment diverges from the training data ([12] www.ncbi.nlm.nih.gov).

This report examines **post-market surveillance for AI medical devices**, focusing on the tension between continuous learning systems and locked models. We survey regulatory frameworks (U.S., EU, global), technical challenges, PMS methodologies, case studies, and emerging solutions. We emphasize evidence-based analysis, citing authoritative sources: regulatory guidance, consensus papers, medical literature, and real-world studies. The goal is to provide a comprehensive, in-depth resource that identifies best practices, gaps, and future directions for keeping AI medical devices safe and effective over time.

Regulatory Frameworks for AI/ML-Enabled Medical Devices

International Guidance and Definitions

Global regulatory bodies recognize the unique nature of AI/ML-based devices. The **International Medical Device Regulators Forum (IMDRF)** has published harmonized terminology and principles for AI-enabled and machine learning–based devices. In particular, an IMDRF guidance (2021) defines key terms: an *ML-enabled Medical Device (MLMD)* is one that "uses machine learning, in part or in whole, to achieve its intended medical purpose." It explicitly notes that AI-based software can continue "learning and iteration as additional data becomes available" ([23] innolitics.com). IMDRF defines:

- **Continuous Learning**: *"Training that leads to change of an MLMD with each exposure to data that takes place on an ongoing basis during the operation phase of the MLMD life cycle."* ([2] innolitics.com).
- **Batch Learning**: training that changes the model at discrete times, based on fixed data sets ([22] innolitics.com).

© 2026 IntuitionLabs.ai - North America's Leading AI Software Development Firm for Pharmaceutical & Biotech. All rights reserved.

- **Locked State**: the device is "in a locked state when changes are not permitted" ([1] innolitics.com). The guidance notes that "locked device" has been used to mean either one the developer does not intend to modify, or any device that does not perform continuous learning ([24] innolitics.com).

These definitions establish a vocabulary for regulators and industry worldwide. IMDRF is also developing *Good Machine Learning Practice (GMLP)* guidelines (including post-market evaluation sections) to align expectations globally. At the national level, regulators have begun issuing AI-specific policies. The FDA and European Commission, for example, have both published or drafted comprehensive guidance for AI devices (see below). Notably, **WHO** and other global health authorities endorse a total-product-life-cycle approach: the WHO's 2021 device surveillance guidance explicitly ties PMS to ongoing compliance with safety and performance requirements (www.who.int).

United States (FDA) Framework

The U.S. Food and Drug Administration has been a leader in addressing AI/ML devices. Its Center for Devices and Radiological Health (CDRH) created a Digital Health Center of Excellence to guide innovation. The FDA has approved hundreds of AI/ML-enabled devices (nearly 950 by August 2024 ([3] www.healthcaredive.com)) across imaging, neurology, cardiology, and other domains. Clearance pathways have included 510(k) submissions (for moderate-risk devices) and De Novo classification (for first-of-kind devices) ([25] pmc.ncbi.nlm.nih.gov).

Pre-market Expectations: In 2019, FDA released a proposed framework for modifications to AI/ML-based Software as a Medical Device (SaMD) ([26] pmc.ncbi.nlm.nih.gov). In early 2025, FDA issued a draft guidance for AI/ML SaMD lifecycle management (subtitled *"Artificial Intelligence-Enabled Device Software Functions: Lifecycle Management and Marketing Submission Recommendations"*). This draft guidance emphasizes that sponsors should include **postmarket performance monitoring and management** plans in their submissions ([4] www.fda.gov). The FDA press release announcing the guidance noted:

> *"The draft guidance includes recommendations for how and when, in marketing submissions, sponsors should describe the postmarket performance monitoring and management of their AI-enabled devices… The proposed recommendations reflect a comprehensive approach to the management of risk throughout the device total product life cycle."* ([4] www.fda.gov).

The guidance (expected finalization in 2025) covers data transparency, bias mitigation, and performance monitoring. It underscores the FDA's view that AI devices require a *life-cycle approach*, not one-time validation. The agency is also convening public webinars on performance monitoring plans.

Notably, FDA's proposed premarket framework introduces a **"Predetermined Change Control Plan" (PCCP)** concept. Under this approach, device developers would define the scope and methods of anticipated algorithm changes (the *SaMD Pre-specifications*), along with an *Algorithm Change Protocol* to control risk ([8] www.ncbi.nlm.nih.gov) ([7] pmc.ncbi.nlm.nih.gov). If variations remain within the pre-specified boundaries, the model could be updated post clearance without a new submission. This plan entails rigorous transparency and real-world performance monitoring to ensure safety ([8] www.ncbi.nlm.nih.gov). In other words, FDA would allow continuous or frequent updates so long as they follow the approved change plan. This marks a significant shift from traditional clearance processes, explicitly accommodating AI adaptivity.

Post-market Requirements: Under U.S. law, all medical device manufacturers must comply with postmarket requirements (21 CFR 822 for section 522 studies, MDR for adverse event reporting, etc.). The FDA's device post-market guidance states that manufacturers must track serious injuries/deaths, malfunctions, device utilization, and must conduct any mandated surveillance or post-approval studies ([14] www.fda.gov). For example, certain AI/ML devices (especially in class III or novel claims) have been subject to FDA-imposed post-approval clinical studies to confirm real-world effectiveness. Additionally, FDA's National Evaluation System for Health Technology (NEST) is exploring real-world data collection to monitor AI/ML device performance.

© 2026 IntuitionLabs.ai - North America's Leading AI Software Development Firm for Pharmaceutical & Biotech. All rights reserved.

Importantly, FDA guidance does not treat all AI modifications as requiring new 510(k). Instead, if modifications are minor or within a cleared PCCP, they may be handled via shorter submissions. A case in point: IDx-DR (AI for retinopathy) was cleared via De Novo in 2018, then benefited from a 2020 510(k) that added a new training mode. FDA accepted the update with only retrospective software performance tests (no new clinical trial), noting that only the feature functionality needed confirmation ([16] pmc.ncbi.nlm.nih.gov). This illustrates the FDA's flexible approach: locked updates can be handled as regulated changes, while continuous mechanisms (once the PCCP concept is fully implemented) could allow many adjustments without full re-reviews.

European Union Framework

The EU approach to AI/ML devices is currently more conservative. The **Medical Device Regulation (MDR 2017/745)** (fully applicable since 2021) governs devices including software. Under MDR, manufacturers must have robust PMS plans, collect data on device performance in clinical use, and file periodic safety update reports (PSURs) for all class IIa, IIb, and III devices (including high-risk software) ([27] pmc.ncbi.nlm.nih.gov). However, neither MDR nor its companion, the In Vitro Diagnostic Regulation (IVDR), contain specific provisions for continuously learning software. AI/ML-based devices are simply treated as software falling under general device rules. Importantly, "devices should not change after market approval" under MDR's conformity assessment philosophy ([28] pmc.ncbi.nlm.nih.gov). This means that any significant modification to an AI device – algorithmic update or retraining – is considered a new version requiring review by a Notified Body.

As described in a recent analysis, this creates tension: **adaptive AI conflicts with MDR's static paradigm.** In the HeartFlow FFR_CT example, regulators noted:

> "Because the algorithm was constantly changing, the clinical performance of the AI system could not be determined in a single point in time, as requested by the MDR… It illustrates the challenge that adaptive AI poses for traditional conformity assessments." ([28] pmc.ncbi.nlm.nih.gov)

In practice, most EU-cleared AI devices remain "locked" in terms of regulatory status. Any planned improvements (for instance, new training data or parameter changes) currently must be submitted to the Notified Body for a revised CE-mark certificate. Some companies are exploring *Continuous Verification* or *Living Documents* for device files, but formal procedures are nascent.

Complicating the landscape is the imminent **EU AI Act**, a horizontal regulation slated to apply from 2026. The AI Act classifies medical AI as "high-risk" (Annex II) and will impose general requirements on high-risk AI systems, including governance, transparency, incident logging, and post-market monitoring ([29] pmc.ncbi.nlm.nih.gov). The AI Act explicitly acknowledges the need for ongoing compliance ("lifecycle management"), though detailed rules on continuous learning are still in discussion. Legal analyses suggest the AI Act *may pave the way for specialized pathways* for continuous learning AI, since MDR/IVDR currently have no such provisions ([5] goodlifesci.sidley.com). As one commentary notes, the AI Act could "even potentially pave the way for a regulatory pathway specifically tailored to continuous learning AI systems" ([5] goodlifesci.sidley.com). Until then, EU practice treats any algorithm change as a change in device design requiring regulatory review.

Other Jurisdictions and Global Initiatives

Other major regulators are also active. In **Canada**, regulators await formal submissions on AI devices; a 2022 CADTH report observed that *no* continuously learning AI medical device had yet been approved there, though locked systems exist ([30] www.ncbi.nlm.nih.gov). Canada's Health Canada likewise applies existing medical device regulations to AI. **Japan's PMDA** and **China's NMPA** are rapidly formulating AI guidance (e.g. China's "Technical Guiding Principle for AI Medical Device" (2020)) and emphasizing "lifespan management". The NMPA has initiatives to "optimize whole life cycle

© 2026 IntuitionLabs.ai - North America's Leading AI Software Development Firm for Pharmaceutical & Biotech. All rights reserved.

regulation" for high-end devices (english.nmpa.gov.cn), likely including AI. Globally, IMDRF and regional harmonization bodies aim to develop shared principles (e.g. IMDRF's Good ML Practices and Europe's MDR guidance).

In sum, regulatory frameworks are evolving. All acknowledge that AI's capacity for change demands special attention. The U.S. is moving toward an explicit lifecycle (post-market aware) model with provisions for managed updates ([8] www.ncbi.nlm.nih.gov), whereas the EU currently relies on general PMS rules and cautious static conformity. Both emphasize the manufacturer's responsibility to plan for and monitor performance over time. Effective PMS for AI will require a blend of these approaches: baseline validation (pre-market), continuous data collection (post-market), risk management of changes, and clear communication with regulators and users.

Post-Market Surveillance: Concepts and Methods

Goals and Processes of PMS

The fundamental goal of PMS is to **protect patient safety** by ensuring devices remain safe and effective in real-world use. This involves:

- **Data Collection:** Gathering information from diverse sources (spontaneous reports, clinical studies, registries, device use logs, quality complaints).
- **Performance Monitoring:** Measuring device outputs (e.g. diagnostic accuracy, false positive rate) in practice.
- **Risk Management:** Analyzing data to detect new or increasing risks (e.g. degradation, unexpected errors, cybersecurity issues) and taking corrective actions (design changes, software updates, user education, device recalls).

For medical devices, PMS is both a regulatory obligation and an ethical imperative. WHO emphasizes that PMS *"ensures devices continue to be safe and well-performing and to ensure actions are undertaken if the risk of continued use … outweighs the benefit"* (www.who.int). Manufacturers must integrate PMS into their Quality Management Systems (e.g. ISO 13485) and Risk Management (ISO 14971) processes. Key regulatory activities include periodic safety update reports (PSURs) or similar documents summarizing real-world data, vigilance (adverse event) reporting, and implementation of corrective & preventive actions (CAPA) when needed (www.who.int).

For **AI/ML-enabled devices**, PMS has added layers:

- **Model Performance Drift:** Unlike static devices, ML algorithms can experience *concept drift* (changes in the external distribution of cases/patients) or *model drift* (parameters shift subtly even without retraining). PMS must include metrics to detect drift (e.g. tracking sensitivity/specificity over time, calibration errors).
- **Data Bias Detection:** In deployment, an AI may encounter patient subgroups underrepresented in training data. Real-world monitoring should include fairness audits (e.g. by demographics) to catch emergent bias or inequity.
- **Software Updates:** PMS feeds back into the development cycle. If drift or bias is detected, the manufacturer may need to re-train the model with new data. For locked devices, this means scheduling a new submission. For continuous devices, this could mean adjusting update algorithms.
- **Usage Patterns:** AI-enabled devices often rely on data inputs or user interactions. PMS may involve analyzing how clinicians use the AI tool (e.g. override rates, user feedback). Continuous learning systems could potentially adapt incorrectly if fed systematically biased user inputs; monitoring must watch for such feedback loops.

© 2026 IntuitionLabs.ai - North America's Leading AI Software Development Firm for Pharmaceutical & Biotech. All rights reserved.

- **Cybersecurity & Data Integrity:** Since AI systems often rely on data connectivity, PMS must include vigilance against data breaches or model corruption (e.g. data poisoning attacks). Notably, regulations like the EU's AI Act and FDA guidance stress secure development and logging to support post-market audit trails.

As regulators recognize, effective AI PMS requires a *Total Product Life Cycle (TPLC)* mindset: planning for surveillance at design stage, using real-world evidence (RWE) sources continuously, and updating both the product and the PM plan over time ([8] www.ncbi.nlm.nih.gov) ([11] pmc.ncbi.nlm.nih.gov).

Post-Market Surveillance Data Sources

Key sources of data for PMS of AI devices include:

- **Spontaneous Reporting Systems:** Reports of adverse events (adverse patient outcomes, device malfunctions) collected via the standard MDR (MedWatch in US, vigilance in EU). However, such reports may miss many AI-specific issues, since misdiagnoses or errors may not be recognized as device-related.
- **Device Registries & Cohort Studies:** Organized collection of cases/platform evaluations, especially for high-risk devices. For example, registries of AI-assisted surgery outcomes or screening results. These allow systematic performance tracking.
- **Real-World Data (RWD):** Analysis of real clinical data (electronic health records, insurance claims, imaging archives) to assess AI performance in situ. The FDA's Sentinel project and other RWD networks could incorporate AI-specific queries. Studies have proposed using RWD to compare AI predictions with actual outcomes (though, as discussed later, causality can be tricky ([11] pmc.ncbi.nlm.nih.gov)).
- **Self-Monitoring Tools:** Some AI devices may include built-in performance logging. For example, an AI imaging tool could record case difficulty metrics, prediction confidence over time, or error rates flagged by users. These logs could be aggregated across hospitals for trend analysis.
- **User Feedback Channels:** Many AI products (especially those integrated into clinical workflow) allow users to flag false positives/negatives or provide corrective input. Systematic capture of this feedback (e.g. an "audit trail" of overrides) is valuable data for PMS, particularly for continuous-learning systems that may incorporate feedback.
- **Periodic Evaluation:** Similar to the lifecycle approach in FDA guidance, manufacturers should perform time-interval performance re-assessments. For instance, every 6-12 months run a validation study on new cases (like a prospective or retrospective reader study or software test) to confirm the AI still meets benchmarks.

Gathering and analyzing these data can be resource-intensive. As **Ansari et al. (NEJM AI)** note, one challenge is that *"positive interventions by clinicians can confound the analysis of AI model performance"*: e.g., if AI kicks off earlier treatment that prevents an adverse outcome, traditional metrics may look worse even though the AI helped ([11] pmc.ncbi.nlm.nih.gov). This highlights that PMS data must be interpreted carefully, often requiring advanced causal analysis. For continuous models, manufacturers may need to withhold model recommendations (or flag their use) in order to unbiasedly measure outcomes ([19] pmc.ncbi.nlm.nih.gov).

Monitoring Model Performance

To maintain safety and effectiveness, continuous tracking of model performance is essential. Key strategies include:

- **Embedding Key Performance Indicators (KPIs):** Manufacturers should define measurable KPIs for the AI device (e.g. sensitivity, false positive rate, calibration error, response time). These metrics should be collected and trended post-market. The FDA suggests a *Performance Monitoring Plan* outlining the KPIs and how they will be tracked ([4] www.fda.gov).
- **Threshold Alerts and "Retrain Triggers":** Predefined thresholds for acceptable performance can be established. If the model's accuracy or error rates fall below a trigger level, this would alert the manufacturer. For continuous systems, this threshold could signal the need for a scheduled retraining or algorithm recalibration.

© 2026 IntuitionLabs.ai - North America's Leading AI Software Development Firm for Pharmaceutical & Biotech. All rights reserved.

- **Statistical Process Control (SPC):** Techniques from manufacturing QA (like control charts) can be applied to per-batch error rates. For example, if an AI triage tool normally flags 5% of exams as positive, a sudden jump might flag a drift.
- **Adversarial and Edge Case Logging:** Some modern AI devices log cases where the model is uncertain (low confidence) or fails (rare data patterns). Regular auditing of these cases helps assess if the model is encountering new types of data.
- **Periodic Audits:** Independent evaluations by new sets of observers (like radiologist re-reads) can periodically benchmark AI performance, akin to post-approval studies. For example, an institution might have a workflow where a fraction of cases are reviewed manually for QA.

Overall, the plan should encompass both automated data analysis and human oversight. In the FDA's draft framework, manufacturers are encouraged to present their PMS methods (including any automated monitoring algorithms) in submissions ([4] www.fda.gov).

Reporting and Risk Management

When surveillance identifies a potential issue, reporting and corrective actions follow. For significant problems (e.g. patient harm or device failure), mandatory reports to regulators are required as usual. Even for issues far short of harm (like unexplained performance declines), FDA recommends formally notifying the agency and users. For example, if an AI diagnostic tool is found to have systematically missed certain lesion types, the manufacturer might issue a safety notice and patch the algorithm.

For continuous-learning devices, risk management plans must account for *the risk associated with adaptation itself*. IMDRF's GMLP draft suggests embedding a risk management file that anticipates how retraining could introduce errors (e.g. overfitting) and how these will be controlled ([31] www.ncbi.nlm.nih.gov). The FDA's Algorithm Change Protocol is one such risk control: changes within the protocol are presumed safe, while deviations trigger a full review ([8] www.ncbi.nlm.nih.gov). The principle is to treat new training as a form of design change that is managed through PMS.

Continuous Learning AI in Medical Devices

Definition and Rationale

A **continuous learning AI** system is one that refines or updates its model on-the-fly as new data are ingested during real-world use. Technically, this might involve online learning algorithms, incremental retraining with incoming data, or periodic automated re-training without human intervention. The key feature is that the model's internal parameters can change on an ongoing basis **after deployment**. In contrast to static (locked) algorithms, continuous learning allows the device to adapt to new trends in patient populations, technology (e.g. imaging hardware upgrades), or unforeseen circumstances (emergent diseases).

The potential benefits of continuous learning are significant. As the literature notes, continuous-learning AI *"is able to improve its predictions and classifications over time"* ([9] www.ncbi.nlm.nih.gov). This could mean better accuracy as more examples accumulate, adaptation to local population specifics, or incremental learning of rare conditions. In the context of medical AI, one often-cited advantage is that a continuously learning algorithm could incorporate real-world outcomes (ground truth) to refine itself. For example, a sepsis prediction model could adjust if it observes that certain lab trends lead to better outcomes with intervention. Likewise, an imaging AI might improve detection of rare tumor subtypes as it sees more cases.

Continuous learning also promises longer device lifespans: instead of a model gradually degrading as practices evolve, an adaptive model could stay *state-of-the-art*. For instance, say an AI ECG tool was developed using older scanners; a

© 2026 IntuitionLabs.ai - North America's Leading AI Software Development Firm for Pharmaceutical & Biotech. All rights reserved.

continuous version could learn to handle data from new ECG devices automatically. In public health, continuous AI could quickly integrate knowledge about new strains of disease (e.g. flu variants) or changing epidemiology.

Risks and Challenges

However, continuous learning also introduces **new safety and validation challenges** ([9] www.ncbi.nlm.nih.gov) ([10] www.ncbi.nlm.nih.gov). As a Canadian analysis observes, the very capability of adaptation *"introduces a number of risks and challenges."* Some specific concerns include:

- **Catastrophic Forgetting:** New training data may override the model's existing knowledge, especially if not carefully managed (a phenomenon called catastrophic forgetting) ([9] www.ncbi.nlm.nih.gov). For example, if a continuous-learning algorithm is retrained on a data stream heavily featuring a new condition, it might lose accuracy on older conditions.
- **Generalizability and Unintended Bias:** A continuous model might adapt too closely to current data, hurting its ability to generalize to other settings. If an AI system is deployed in a new region or hospital with different patient demographics, its adaptive learning could accentuate biases if not corrected. The CADTH report warns that continuous AI may *"not perform as expected when learning from real-world data"* if that data is not representative ([9] www.ncbi.nlm.nih.gov).
- **Unpredictability:** A continuously updating model is inherently a *moving target*. Small changes in data feed can gradually shift the model in unforeseen ways. This unpredictability complicates risk assessment: it becomes harder to guarantee that the model still meets the original safety criteria. Regulators and manufacturers stress the need for *"regular and ongoing evaluation"* of continuous AI systems to detect drift and unintended consequences ([10] www.ncbi.nlm.nih.gov).
- **Validation Timing:** A core question is *when* to re-evaluate effectiveness. Traditional validation (as part of a regulatory submission) provides one snapshot. Continuous learning requires defining triggers for re-validation. Should performance be audited monthly, quarterly, or after a certain volume of new data? These policies are not yet standardized.
- **Data Integrity:** Continuous learning relies on incoming data which may not be curated or fully labeled. There is a risk of "garbage in, garbage out" if poor-quality real-world data inadvertently biases the model. Data drift due to changes in measurement devices or software updates can inadvertently degrade the model.
- **Software Quality and Cybersecurity:** The infrastructure supporting continuous learning (data pipelines, retraining servers) introduces new attack surfaces. For instance, an adversary could attempt *data poisoning* by submitting manipulated cases to alter the model. Regulatory guidance (EU AI Act, FDA) emphasizes robust software development and cybersecurity measures in the lifecycle of AI devices.
- **Transparency and Explainability:** Continuous updates complicate transparency. Clinicians using the device need to know if the algorithm changed. FDA's draft guidance and other experts suggest continuous models should log changes and update logs for review. Moreover, patients and providers must trust that the continuous updates do not make the AI behave unpredictably. Lack of transparency could erode trust rapidly.

Because of these issues, continuous learning AI must have **built-in guardrails**. Proposed risk management strategies include:

- Maintaining a "frozen" version for validation purposes, to compare against the updated one.
- Human oversight mechanisms (e.g. requiring clinician sign-off before algorithm changes are released for clinical use).
- Internal quality controls, such as test suites that each updated model must pass.
- Versioning and audit logs to track exactly how the algorithm has changed (so issues can be traced).

Regulatory Approaches to Continuous Learning

Regulators are actively grappling with how continuous learning fits into device approval. As noted, the current FDA draft concept of a *Predetermined Change Control Plan* is aimed directly at facilitating controlled continuous changes ([8]

© 2026 IntuitionLabs.ai - North America's Leading AI Software Development Firm for Pharmaceutical & Biotech. All rights reserved.

www.ncbi.nlm.nih.gov). Under such a plan, an AI manufacturer would define:

- The scope of future updates (e.g. what data can be used for retraining).
- Success criteria (the model must remain within pre-specified performance bounds).
- A monitoring plan (how to measure performance baseline and detect drift).

The ALGORITHMIC CHANGE PROTOCOL would then lay out exactly how retraining is to occur and how changes are validated. If later the model is retrained within those conditions, it is considered compliant. The FDA expects "transparency and real-world performance monitoring by manufacturers" so that both the FDA and the sponsor can evaluate the model continuously ([8] www.ncbi.nlm.nih.gov). In other words, the liability for safe updates is shared through prior planning.

In Europe, where the AI Act and MDR apply, a dedicated pathway for continuous learning has not yet been established. As the Sidley analysis points out, MDR/IVDR currently only certify *fixed-state* AI: *"AI systems that are retrained after being placed on the market must currently be reviewed by notified bodies whenever their systems are substantially changed."* ([5] goodlifesci.sidley.com). This implies each major re-training is effectively a new device version. Some observers expect that EU regulations will eventually create a special regime for "adaptive AI" (possibly through guidance under the AI Act) that could allow continuous learning under strict conditions. The EU AI Act itself mentions risk management and monitoring for high-risk AI, but details on continuous updates remain under development.

Other countries are building similar models. Japan's PMDA has signaled intention to allow some level of adaptive AI under a tightly controlled plan, akin to FDA's. Global harmonization efforts (IMDRF, GMLP) uniformly stress that *good machine learning practice* includes lifecycle management and performance reassessment plans for adaptive algorithms.

In summary, continuous learning AI sits at the frontiers of regulation. Authorities encourage it insofar as it can improve outcomes, but only with explicit controls. The prevailing view is that continuous learning systems are only acceptable if they adhere to a **proactive plan** that ensures each update is safe. This will likely involve a combination of pre-specification (as in FDA's PCCP) and intensive PMS.

Implementation Considerations

Technical Implementation: Manufacturers designing continuous learning AI should architect them with monitoring in mind. This often means splitting the product into (a) a static core that is rigorously validated, and (b) an update mechanism that is carefully controlled. The update mechanism might feed back cases that were misclassified (with clinician confirmation) into periodic retraining. Crucially, any automatic retraining should include integration tests (for safety) before deploying updated weights. The model may maintain an internal *confidence threshold* above which it is allowed to adapt.

Ethical/Social: Continuous learning raises questions about informed consent and transparency to patients. If an AI system changes its behavior, should users or clinicians be informed? How should this information be presented? Answers will need alignment with data governance and medical ethics, especially when the model "remembers" patient data for learning.

Economic/Operational: Continuous learning systems can potentially save time by reducing the need for frequent regulatory submissions. However, they may impose additional burdens: maintaining infrastructure, documenting changes, and robust quality assurance. Healthcare providers need to trust that a continuously learning system will not degrade – which may require contractual assurances from manufacturers.

In all, continuous learning AI offers potential clinical benefits but also demands a strong PMS framework. Without rigorous surveillance and control, a continuously changing medical AI could introduce new, poorly-understood risks. Conversely, with proper checks, continuity of learning can help an AI device remain state-of-the-art. The remainder of this report analyzes how to achieve that balance in practice.

© 2026 IntuitionLabs.ai - North America's Leading AI Software Development Firm for Pharmaceutical & Biotech. All rights reserved.

Locked (Static) AI Models in Medical Devices

Definition and Rationale

In contrast to continuous-learning systems, a **locked AI model** is one whose parameters do not change during normal operation. Once approved or cleared, the algorithm is fixed. If future improvements are needed, they are implemented by releasing a new software version (often requiring fresh regulatory review). Nearly all currently marketed AI medical devices have been locked at launch, reflecting the traditional paradigm. For example, an image analysis AI is trained on development and validation datasets and then "frozen". Every patient image is processed with the same model. Updates to this model (retrains, tweaks) are batch processes that occur offline.

Locked models are conceptually simpler and easier to understand. Their performance can be fully characterized at approval time through clinical trials or studies. Established regulatory processes for software modifications (like submission of a new 510(k) for changed indications or substantial algorithm changes) already exist. This model mirrors how other software is regulated and often aligns with a medical device's life cycle documentation (such as a static "Design History File").

Benefits of Locked Models

The primary advantage of locking the algorithm is **predictability**. Clinicians and regulators know exactly what the device does and how it was validated. There is no uncertainty from unseen training on new data. This brings several specific benefits:

- **Regulatory Simplicity:** Locked devices follow existing change control processes. Any substantial upgrade triggers a clear path (new 510(k)/PMA update or CE recertification). Regulators and notified bodies can apply familiar risk-based review. As the IMDRF notes, a locked device is one "for which the developer does not have an intention of modifying at the present time" ([32] innolitics.com). This fits squarely into the conventional device framework.
- **Validation and Reproducibility:** The performance metrics (sensitivity, specificity, calibration) measured in premarket studies remain valid in postmarket use, assuming the environment doesn't change. There is a stable reference for clinicians to trust.
- **Transparency and Explainability:** With a fixed model, any explanations or interpretations based on the model's coefficients, if provided, remain valid. Documentation (like risk analyses) needs no amendment unless changes are made.
- **Patient Trust and Physician Confidence:** Users often find it reassuring that the algorithm will not "silently change its mind." There is less feeling of unpredictability compared to an always-learning system.

Limitations and Risks of Locked Models

Locked AI models also have notable drawbacks, particularly over the *long term*. The biggest risk is **obsolescence due to data drift**. Healthcare data distributions can change over time (due to new technology, emerging diseases, demographic shifts). A fixed model trained on data from 2015 might not perform optimally in 2025. For instance, a pneumonia detection AI trained on chest X-rays from older patients might underperform on CT scans of a younger population.

Key issues include:

- **Model Drift:** Without retraining, the deployed model may gradually lose accuracy if the underlying data patterns shift. This is akin to how weather forecasting models may degrade if climate patterns change—they are static and need re-calibration. The CADTH review explicitly warns that *"Locked AI can become dated, in that the training data may no longer be representative of real-world data and can experience model drift where the AI performance degrades over time."* ([12] www.ncbi.nlm.nih.gov).

© 2026 IntuitionLabs.ai - North America's Leading AI Software Development Firm for Pharmaceutical & Biotech. All rights reserved.

- **Performance Gaps:** If data or practice patterns change significantly, locked models might not catch up without manual intervention. For example, consider the IDx-DR diabetic retinopathy AI. A recent real-world study found it failed to analyze the fundus images in 26.1% of patients ([13] pmc.ncbi.nlm.nih.gov). Issues like smaller pupil sizes or different camera quality hindered image analysis. A continuous-learning version could potentially learn to handle those edge cases, but the locked IDx-DR required an external update.
- **Batch Retraining Overhead:** When updates are needed, locked models rely on major process (re-approval). This can be cumbersome. The PLOS analysis of FDA data found that only *eight* AI/ML radiology products had formal post-market improvements approved ([33] pmc.ncbi.nlm.nih.gov) between 201X–202X. Each improvement required an official process (often using retrospective data review) ([34] pmc.ncbi.nlm.nih.gov). The average interval to an update was about 348 days ([35] pmc.ncbi.nlm.nih.gov). In reality, many locked devices may never be updated after initial clearance unless a compelling new use arises.
- **Limited Personalization:** Locked models treat all user populations the same (unless separate versions are released). They cannot learn from local data at a given hospital to improve local performance. For example, if a certain hospital has a unique patient demographic (ethnic mix, disease prevalence) different from the training set, the fixed AI will not adapt to that. Continuous models might mitigate this by adapting to local data streams (though with their own risks).
- **Inefficiency in Learning:** Combining global learning and local deployment is difficult. In a locked approach, each institution may retrain models on local data separately, leading to wasted effort and potentially poorer models that never benefit the broader community. The US regulatory pathway (510k) does not easily allow a manufacturer to harness multi-center data unless re-submitting for each upgrade.

Locked models require effective PMS to track when performance degrades enough to warrant an update. This means manufacturers should **plan for periodic re-validation**. Regulatory agencies have indicated that manufacturers should ideally incorporate a PMS plan reflecting any known issues (e.g. IDx-DR labeling warns about poor image quality). If significant drift is detected, the only recourse is a formal improvement (as with IDx-DR's new training mode clearance ([16] pmc.ncbi.nlm.nih.gov)) or recall.

Post-Market Performance of Locked Models: Real-World Examples

The real-world performance of locked AI devices can provide insights:

- **IDx-DR (Retinopathy AI):** As noted, a large German study (875 patients) found that in 26.1% of cases IDx-DR could not analyze the image ([13] pmc.ncbi.nlm.nih.gov). When images were analyzable, IDx-DR matched ophthalmologist grading in only ~54.2% of cases ([13] pmc.ncbi.nlm.nih.gov). While detection of severe cases was good (94% sensitivity in those), the locked model's inability to analyze a quarter of images is a serious gap. If IDx-DR were continuously learning, it might ask doctors to re-capture images or adjust to suboptimal input, potentially improving yield. Instead, the locked system left these as failures unless the solution was updated in practice (which it was, by adding a training mode in 2020 ([16] pmc.ncbi.nlm.nih.gov)).
- **SkinVision (Melanoma Detection App):** A prospective evaluation of the CE-marked SkinVision app (locked model running on smartphone images) vs dermatologists found alarmingly poor real-world accuracy ([15] pmc.ncbi.nlm.nih.gov). The app's advertised performance did not hold in practice: sensitivity ranged only from 41% to 83%, and specificity 60–83% ([15] pmc.ncbi.nlm.nih.gov). It tended to over-call lesions as cancerous, leading to many false positives. Both patients and doctors had *low confidence in the app*, and no patient trusted it alone ([15] pmc.ncbi.nlm.nih.gov). This underscores that a locked AI, once cleared internationally, may underperform in diverse practice settings. Without a continuous-learning component, SkinVision would require a new algorithm version release (with new training data) to improve.
- **Anterior Segment Imaging AI:** In ophthalmology, some FDA-cleared devices (e.g. an IOL calculator with AI) are locked at clearance. Post-market studies have sometimes revealed calibration drift due to changes in surgical techniques. Manufacturers typically issue updates after collecting new larger datasets. This iterative update process is the only way to "learn" with a locked model, but it is slow and reactive.

These examples illustrate the spectrum: locked models can work well if training data matches real use; otherwise they risk large blind spots. They emphasize the critical need for active monitoring (PMS) to identify issues. In the SkinVision case, presumably user reports or complaints in country-of-use (Europe, for example) could prompt the company to revise the AI (if it chose to). In the IDx-DR case, FDA clearance of an updated software version took years after knowing some

© 2026 IntuitionLabs.ai - North America's Leading AI Software Development Firm for Pharmaceutical & Biotech. All rights reserved.

limitations. Under a continuous paradigm, in contrast, one could imagine ongoing user feedback directly feeding into model refinements much sooner.

Surveillance Strategies for Locked Models

Given the limitations, PMS for locked models focuses on *detection of when performance decays and initiating updates*. Recommended strategies include:

- **Regular Re-Validation Studies:** Schedule periodic studies where the AI outputs are compared to current ground truth. For instance, every 1–2 years, perform a retrospective evaluation on a set of recent cases or images from multiple sites (similarly to how the IDx-DR resubmission evaluation was done ([16] pmc.ncbi.nlm.nih.gov)). This can catch drift.
- **Quality Metrics Tracking:** Incorporate usage analytics into installations. If the AI is deployed as part of an imaging system, collect statistics like the percentage of images flagged, distribution of prediction confidences, and repeat analysis rates (cases where AI requests a retake). Any significant change over time could indicate drift.
- **User Feedback Mechanisms:** Implement easy ways for clinicians to flag false negatives/positives. For insured markets, even litigation data (malpractice claims against device) could be monitored as a signal.
- **Periodic Label Updates:** As new kinds of cases emerge (e.g. new disease variants, new imaging devices), plan a cycle of updating the training labels. For example, if an imaging AI sees scans from a new scanner model, collect a sample and relearn features to accommodate it.
- **Alert Thresholds:** If continuous performance monitoring (via sampled outputs) shows metrics falling outside acceptable ranges, issue safety notices or temporarily disable the AI pending revision.

In essence, the locked-model approach treats performance drift as a *risk* to be mitigated by careful monitoring and timely product updates. This is labor-intensive but aligns with existing responsibilities of manufacturers. Regulatory bodies will expect documented PMS plans addressing locked-model drift and any planned cycle for re-training.

A key point noted in the literature is that *both continuous and locked systems need continuous vigilance*. Even a locked model that "never learns" can still degrade if its context changes. The CADTH report comments: *"Locked AI will also need periodic evaluation and might lead to a situation of needing to distribute updates or make withdrawals across the healthcare system."* ([12] www.ncbi.nlm.nih.gov). Therefore, even "static" AI devices cannot be totally hands-off. The difference is that locked AI must rely on human-initiated updates (batch model retraining), whereas continuous AI can (in theory) adjust automatically if properly governed.

Comparison of Continuous Learning vs. Locked Models

To summarize key contrasts, the table below outlines major differences in approach, benefits, and challenges:

Aspect	Continuous Learning AI	Locked (Static) AI
Definition	Model retrains or updates incrementally with each new data exposure during actual operation ([2] innolitics.com). Sometimes called "adaptive" or "unlocked" AI.	Model parameters remain fixed during operation; any updates occur only when manufacturer deploys a new version through formal channels ([1] innolitics.com).
Validation	Cannot be fully validated "once and for all." Requires plan for ongoing re-validation. FDA proposes *Algorithm Change Protocol* to define how updates will be tested ([8] www.ncbi.nlm.nih.gov).	Validated baseline at approval (benchmarks, clinical trials). Subsequent changes require new validation studies. Easier to verify performance matches documentation.
Regulatory Pathway	Emerging; FDA is moving toward allowing pre-specified updates without new 510(k) if within plan ([8] www.ncbi.nlm.nih.gov). EU currently has no special pathway; updates need re-certification.	Established pathways: changes follow existing device modification rules (e.g. new 510(k) or PMA supplement, or CE recert). No expectation of automatic updates.

© 2026 IntuitionLabs.ai - North America's Leading AI Software Development Firm for Pharmaceutical & Biotech. All rights reserved.

Aspect	Continuous Learning AI	Locked (Static) AI
Performance Concept	Trade-off: may improve and adapt, but risk unpredictability. Must guard against *catastrophic forgetting* and ensure continued generalizability ([9] www.ncbi.nlm.nih.gov).	Stable performance relative to original state, but risk *drift* if environment changes. Could degrade or become outdated if not retrained.
Risk Management	Must specify safety controls in Algorithm Change Protocol. Real-time monitoring required. Transparency of changes (audit trail) is crucial ([8] www.ncbi.nlm.nih.gov).	Traditional risk management applies. Risk of drift handled via periodic risk analyses and manufacturer updates or recalls if needed.
Post-Market Surveillance	Emphasizes automated real-world monitoring of performance metrics, bias, and errors (potentially using AI on real data) ([11] pmc.ncbi.nlm.nih.gov). May incorporate clinician feedback in loop.	Uses standard surveillance (MDR reports, user feedback, periodic audits), but should include revalidation studies and planned updates based on triggered signals.
Adaptation to New Data	High potential for adaptation – can learn from new patient data or imaging sources. For example, could adjust to new disease variants or demographics over time.	No automatic adaptation. Must explicitly collect new data and retrain model offline. Updates lag real-world changes.
Examples	Proposed models: closed-loop insulin dosing that adapts to patient's physiology, or radiology AI that retrains on new institutional data (not yet widely marketed). (No FDA-cleared continuous-learning AI in 2025; under development) ([30] www.ncbi.nlm.nih.gov).	Many current devices: e.g. IDx-DR (diabetic retinopathy AI), fixed mammography or dermatology screening AIs, standard image analysis tools.

The choice between continuous and locked involves a trade-off between **innovation speed vs. control**. Continuous AI can potentially be more responsive to new information, aligning with agile data science workflows. Locked AI offers rigorous oversight at the expense of slower update cycles. From a surveillance standpoint, locked-AI devices essentially shift the risk to periodic review cycles (which may be years apart), whereas continuous AI requires constant watches.

Post-Market Surveillance Considerations: Data and Evidence

Empirical data on AI device performance underscores the importance of robust PMS. Several recent surveys and studies highlight gaps and outcomes:

- **Regulatory Growth and Distribution:** The Healthcare Dive analysis shows FDA had cleared 950 AI/ML devices by Aug 2024 ([3] www.healthcaredive.com). Radiology and imaging account for most; approvals accelerated in 2018–2020 ([18] pmc.ncbi.nlm.nih.gov). However, analyses suggest *data paucity*: one RAPS article reported that most cleared imaging AI devices had limited public evidence of performance in real-world settings . This implies that PMS will have to generate new evidence.

- **Performance Drift in Practice:** Ansari et al. (NEJM AI, 2025) highlight that AI predictive models often face dataset shifts in use. They note that even with continuous monitoring, arrhythmia arises: *"confounding medical interventions…modify outcomes, introducing bias into performance assessment"* ([11] pmc.ncbi.nlm.nih.gov). In other words, an AI that prompted a better treatment might appear to fail at predicting an "outcome" because that outcome was prevented. This insight means PMS for AI cannot simply compare predicted vs. observed outcomes; it must account for the human-in-the-loop effect. The authors call for causal modeling approaches to disentangle true model decay from effective interventions ([19] pmc.ncbi.nlm.nih.gov).

- **Signal Detection:** The field of pharmacovigilance has begun leveraging AI (e.g. large language models) to sift through vast data for adverse event signals ([21] jamanetwork.com). By analogy, specialized AI tools may assist in AIaMD surveillance – for example, analyzing free-text radiology reports to spot patterns of AI misses. However, these tools also have biases and false positives ([21] jamanetwork.com), so any PMS AI would itself need validation.

- **Case-Specific Outcomes:** Case studies of performance highlight both stability and failure modes. For instance, many AI imaging systems perform very well on well-curated test sets (AUC ~0.95 as in some pivotal studies), but real-world conditions may differ markedly. Postmarket audits in some hospitals have found that image quality (which manufacturers rarely account for) is often worse in practice than in studies, reducing AI accuracy. These underscore that PMS should monitor device use in situ, not just rely on premarket specs.

- **Quantitative Evidence:** Where available, quantitative performance data motivate PMS actions. In the SkinVision study, the low specificities (60–83%) imply that for every 100 lesions, up to 40 false alarms would occur ([36] pmc.ncbi.nlm.nih.gov). If such a device were deployed as a screening aid, it would cascade into many unnecessary biopsies or anxious referrals. PMS mechanisms (like requirement to report usage and outcomes) could identify if actual referral rates escalate after deploying such an app. Similarly, IDx-DR's 26% failure rate means one in four screenings gave no result ([13] pmc.ncbi.nlm.nih.gov); a responsible surveillance plan would track analyzable-image rates and prompt analysis if patients fall outside expected ranges.

© 2026 IntuitionLabs.ai - North America's Leading AI Software Development Firm for Pharmaceutical & Biotech. All rights reserved.

Collectively, these data illustrate the variety of evidence that PMS must integrate. Surveys and real-world performance studies complement regulatory clearance data. Notably, **expert consensus** acknowledges that AIaMD PMS is essential: the European Society of Radiology explicitly calls for standardized PMS practices and clinician involvement in monitoring ([20] pmc.ncbi.nlm.nih.gov) ([27] pmc.ncbi.nlm.nih.gov). They point out a startling gap – a 2025 ESR survey found only ~30% of radiologist users were even familiar with PMS requirements ([37] pmc.ncbi.nlm.nih.gov) – underscoring the need for education alongside surveillance systems.

Case Studies

To ground the discussion, we present case vignettes illustrating post-market issues in both locked and (prospectively) continuous systems. These examples span diagnostic domains:

Device/Case Study	Function	Learning Mode	Key PMS Issues / Outcomes
IDx-DR (Digital Diagnostics) ([16] pmc.ncbi.nlm.nih.gov) ([13] pmc.ncbi.nlm.nih.gov)	Automated screening for diabetic retinopathy (retinal image analysis)	Locked (De Novo clearance with retraining update via 510(k))	In real-world use, 26.1% of patient images were unassessable by the system ([13] pmc.ncbi.nlm.nih.gov) (often due to small pupils or poor image quality). After ~975 days, the company added a "training mode" (cleared in December 2020) to improve usability ([16] pmc.ncbi.nlm.nih.gov). This update was reviewed via retrospective performance testing (standalone software assessment) rather than new clinical trial. PMS focus: monitoring image analyzability rates, shadowing sensitivity/specificity over time, and addressing workflow issues.
SkinVision (SkinVision BV) ([15] pmc.ncbi.nlm.nih.gov)	Smartphone app to assess pigmented skin lesions (melanoma screening)	Locked (CE-cleared static algorithm)	A prospective study in 2022 found the app had low accuracy: sensitivity 41–83%, specificity 60–83% ([15] pmc.ncbi.nlm.nih.gov). It overcalled lesions, leading to many false alarms. Both patients and dermatologists showed low confidence, and no patient would trust the app alone ([15] pmc.ncbi.nlm.nih.gov). PMS concern: the CE-marked device's performance fell short of clinical needs; real-world oversight would ideally flag the excessive false positives and prompt algorithm refinement. This required generating new training data (if updated at all).
HeartFlow$_{FFR}$CT (HeartFlow Inc.) ([6] pmc.ncbi.nlm.nih.gov)	CT-based computation of fractional flow reserve (AI/ML analysis of coronary CT angiograms)	Locked (intended to be adaptive but constrained by regs)	The EU regulators noted that HeartFlow's algorithm was "constantly changing" during development, making it hard to freeze for MDR approval ([6] pmc.ncbi.nlm.nih.gov). Under EU rules, such continuous updates would require separate review. Eventually the company demonstrated performance via a series of clinical studies and was CE-marked, but with an explicit statement that updates would need regulatory oversight. PMS challenge: continuous improvement was needed to match new CT technology, but MDR frameworks demanded each change be documented. The case highlights a tension between state-of-the-art algorithm development and static device norms.
Aidence Veye Lung Nodule (Aidence)	AI for lung nodule detection on chest CT	Proposed Continuous / Unknown (incorporating learning from new scans)	In published regulatory analysis, this device is cited as requiring "continuous monitoring and updating" to maintain accuracy ([38] pmc.ncbi.nlm.nih.gov). (Exact mode unspecified, likely locked updates in practice.) PMS focus: ensuring small lung nodules continue to be detected as CT slice thickness and scanning protocols change. In US-510(k) clearance records, software changes (e.g. improved algorithms) were done via supplemental submissions. The device illustrates that meticulous real-time monitoring is needed: radiologists or QA algorithms could spot if nodule detection rates drop, triggering updates.

These examples demonstrate real-world PMS concerns. In each case, either performance issues emerged in practice (SkinVision, IDx-DR) or were anticipated by developers/regulators (HeartFlow, Veye). They reflect common themes:

- **Unanticipated Use Factors:** IDx-DR's high failure rate was due to miotic pupils – a factor not evident in initial testing. PMS highlighted that need for fallback solutions (e.g. training mode that guides image capture).
- **Validation vs. Reality:** SkinVision's marketed accuracy did not hold up, showing that prospective surveillance can reveal gaps even for CE-certified devices.
- **Regulatory Fit:** HeartFlow's development required negotiation on how to treat an evolving algorithm under MDR. Ultimately, they managed via formal clinical evidence submission, but the regulatory friction was clear.
- **Continuous Vigilance:** Aidence's case underscores that for something as subtle as lung nodule detection, static performance evaluation is not enough; monitoring over time is essential.

Table 2: Post-Market Data on Key Cases
We compile the salient data points from the above cases:

© 2026 IntuitionLabs.ai - North America's Leading AI Software Development Firm for Pharmaceutical & Biotech. All rights reserved.

Device	Performance in Use	Update/Surveillance Actions
IDx-DR	26.1% of clinics' images ungradable ([13] pmc.ncbi.nlm.nih.gov); sensitivity ~94% on analysable images; specificity ~90%. Matched ophthalmologists in only ~54% of cases ([13] pmc.ncbi.nlm.nih.gov).	After deployment, manufacturer released an improved software with a "training mode" via 510(k) (Dec 2020) ([16] pmc.ncbi.nlm.nih.gov). FDA evaluated via retrospective SA testing. PMS should monitor analyzability rates and referable-disease detection rates going forward.
SkinVision	Prospective study: sensitivity 41–83%, specificity 60–83% ([39] pmc.ncbi.nlm.nih.gov). Classified far more lesions as high-risk than dermatologists (leading to "clinically harmful" over-detection) ([40] pmc.ncbi.nlm.nih.gov).	The study itself suggests limited trust. For safety, the manufacturer would need to collect more labeled data and recalibrate the algorithm for clinical use. Possible regulatory re-submission (CE refinement) would be needed for any algorithmic change, given the locked status.
HeartFlow FFR_{CT}	Not publicly quantified here. Known to perform well when FDA-approved, but EU approval was delayed due to its evolving nature.	Company conducted multiple validation studies for its adaptive algorithm. Currently, any algorithm improvements are likely treated as design changes requiring Notified Body review. PMS would include clinical registry data on patient outcomes.
Aidence Veye	Specific performance numbers not given. Cited as requiring ongoing updates.	In EU, manufacturer must provide PMS data as part of compliance. In the US, any significant algorithm change would go through 510(k). Monitoring would use CT case follow-ups.

Data Analysis and Evidence-Based Discussion

The above case data illustrate broader findings from analyses of AI devices. In particular, we highlight some patterns and expert recommendations:

- **Rapid Approval vs. Slow Data Accrual:** The conflation of rapid deployment and slow evidence gathering is a recurring concern. As one review notes, **AI devices often hit the market with limited user training data** ([11] pmc.ncbi.nlm.nih.gov). Postmarket data often lag behind. Consequently, PMS plans must prioritize closing evidence gaps as soon as possible. This might involve multi-center registries or even randomized trials post-approval for high-risk AI.
- **Signal of Performance Decay:** Monitoring baseline performance metrics over months/years can reveal downward trends. For example, if an AI screening tool's positivity rate drifts significantly (after adjusting for population changes), that is a red flag. Statistical process control methods from Six Sigma can be applied: control charts with action thresholds can turn continuous performance data into alerts.
- **Clinical Decision Impact:** Some experts argue that PMS for AI should focus on *patient outcomes* rather than raw model metrics. If an AI helps guide a therapy, one should track outcome metrics (e.g. morbidity, mortality) over time to ensure its benefit is sustained ([11] pmc.ncbi.nlm.nih.gov). However, this is complex since outcomes are influenced by many factors beyond the AI. This requires sophisticated causal study designs.
- **Comparison Group:** Given conflicting signals (e.g. Ansari's confounding issue), some recommend comparing outcomes with and without AI assistance in the postmarket period. For example, if an institution sometimes uses the AI (for certain surgeons) and not for others, that natural experiment can inform AI effectiveness. This could be part of a PMS study design.
- **Software Metrics as Signals:** Besides clinical outcomes, software-level metrics can serve as proxies. For instance, a consistently rising false alarm rate might indicate deterioration. A malpractice or complaint database tracking AI-assisted cases can be mined for trends. For example, if multiple claims arise due to a specific failure mode of the AI (e.g. missed lesions in a particular organ), this triggers action.
- **Regulator-Facilitated Studies:** In some cases, regulators may require post-approval studies for AI devices (especially if pivotal trials were small). For example, FDA's PMAs for AI diagnostics sometimes include post-market study conditions. These generate high-quality data for PMS. Transparency reading all postmarket commitments ensures accountability.

Overall, the evidence underscores that **no surveillance approach is foolproof**. Armed only with passive reports, malfunctions might be missed until too late. Combining multiple data sources (active strategizing) is recommended. The ESR recommendations (radiology) advocate that deploying institutions should share data on AI performance (via feedback forms or network) and that vendors should maintain living post-market clinical follow-up mechanisms ([27] pmc.ncbi.nlm.nih.gov).

Implications and Future Directions

© 2026 IntuitionLabs.ai - North America's Leading AI Software Development Firm for Pharmaceutical & Biotech. All rights reserved.

Harmonizing Regulation and Innovation

The dichotomy of continuous vs locked models illustrates a central policy challenge: how to encourage beneficial model evolution without compromising safety. Regulators are converging on the view that *proactive lifecycle management* is key. The EU AI Act and FDA TPLC initiatives push manufacturers to integrate monitoring and update mechanisms. We anticipate:

- **Global Standards (IMDRF GMLP):** The IMDRF AI/ML Working Group's forthcoming Good ML Practice guidelines (currently under consultation) will likely become touchstones for PMS and adaptive algorithms internationally ([41] www.imdrf.org). These will probably mirror FDA's Algorithm Change Protocol concept and urge rigorous PMS planning.
- **International Databases:** Similar to Eudamed (EU's device database) and FDA's public listings, we might see registries of AI algorithm performances, e.g., aggregate metrics per breakout by population. This could facilitate benchmarking (like how wine producers share water quality). Ideally, regulators would require anonymized AI performance logs to aggregate into national-level surveillance.
- **Dynamic Labeling:** Labels (IFUs) for AI devices may become dynamic over time. Rather than static paper inserts, AI device documentation might be updated continuously. For example, Google's AI diagnostic models for imaging now come with recommendations that they only "support" and that final decisions rest with physicians. If a model is updated, the label could note version changes on a website.
- **Stakeholder Roles:** Radiologists and other clinicians are increasingly expected by regulators to participate in PMS. The ESR consensus emphasizes radiologist contribution to monitoring safety ([27] pmc.ncbi.nlm.nih.gov). Nursing staff, technologists, or even patients (via apps) might provide user-driven data. Collaboration between vendors and clinicians on performance logs (e.g. via integrated comment fields in software) is a likely trend.

Technological Aids for PMS

The tools used for PMS will themselves evolve, often leveraging AI for AI oversight. Examples:

- **Anomaly Detection on Data Streams:** Machine learning can be applied to detect anomalies in device usage data. For instance, unsupervised methods could spot when an AI's output distribution shifts. LLMs or big data analytics could scan radiology reports to find narrative mentions of AI errors.
- **Benchmarking Suites:** Simulators and challenge datasets could be run periodically. For example, a consortium might maintain a refreshed dataset of diverse cases; AI devices would have to score periodically on it to demonstrate maintained accuracy.
- **Audit Frameworks for ML:** Analogous to financial audit, ML audit protocols are in development (part of GMLP). These would routinely check training data integrity, version control logs, and performance records of devices in field.
- **Real-World Performance Dashboards:** Cloud-based systems could allow manufacturers to upload de-identified output logs from customer installations for centralized analysis. Regulators or third parties (like patient safety organizations) might run independent analytics on this pooled data.

Economic and Social Considerations

Continuous learning can reduce the time and cost of regulatory submissions for every update (assuming a plan is approved), potentially accelerating innovation. However, it also shifts some burdens onto ongoing processes of validation and monitoring. Cost-sharing models may emerge: for instance, payers might fund post-market registries or data collection for AI devices, given the public health interest. Conversely, if an AI device fails in practice, liability issues become complex. Manufacturers may require clinicians to use devices under specific protocols or disclaim that they operate on trial basis until enough PMS data accumulate.

Another social impact is on **device access**. If continuous-learning devices require internet connectivity for updates and monitoring, resource-limited settings may be disadvantaged or disproportionately affected by lapses. Policymakers will

© 2026 IntuitionLabs.ai - North America's Leading AI Software Development Firm for Pharmaceutical & Biotech. All rights reserved.

need to ensure that PMS requirements do not inadvertently widen global health disparities. Open-source AI models could be subject to community-based surveillance as well.

Finally, **patient engagement** will matter. Patients should be informed when AI devices are used in their care and possibly consent-aware to their AI-driven data being used for continuous learning. Transparency initiatives (e.g. an AI label stating the version and last update date) may become standard. Patient-reported outcomes might also feed into surveillance (e.g. symptom tracking after AI-guided diagnosis).

Research and Evidence Gaps

Key areas needing further study include:

- **Continuous vs Batch Update Outcomes:** Empirical comparisons of systems that were continuously updated versus those updated in batches would inform best practices. Currently, long-term studies of continuous AI in real practice are scarce. Pilot programs (e.g. in radiology networks) could provide data on which approach better maintains accuracy.
- **Cost-Benefit Analyses of Surveillance:** Determining the optimal intensity of PMS is an open question. Overly burdensome monitoring (many false alarms) wastes resources; too slack leads to missed problems. Research can help quantify trade-offs (e.g., how many cases should a monitoring plan sample to detect a given performance drop with high confidence).
- **Standard Metrics:** The field lacks standardized metrics for many AI tasks. If every AI uses a different threshold (e.g. risk score) or defines "positive" differently, pooling data is hard. Harmonizing metrics across devices would ease surveillance. This is an area of active discussion (e.g. open AI challenge datasets).
- **Ethical Monitoring:** Feedback loops in continuous learning raise ethical issues. Studies are needed on consent, privacy, and bias correction strategies in continuous systems. For example, continuous retraining on sensitive populations must be done with care to avoid stealth discrimination.

Conclusion

Post-market surveillance of AI medical devices is a field in rapid flux, mirroring the rapid evolution of the underlying technology. Our comprehensive review finds that **locked** and **continuous** learning models each pose unique challenges and advantages. Locked models offer stability and clear validation, but risk becoming outdated and require reactive updates. Continuous learning models offer adaptability and potential for ongoing improvement, but demand rigorous oversight to prevent unpredictable failures.

Both models necessitate vigilant PMS. Established principles of surveillance (data collection, risk management, corrective action) apply, but must be extended to account for algorithmic behavior. In practice, this means richer data streams, automated monitoring tools, and new regulatory mechanisms. Industry and regulators worldwide recognize this need: the FDA is already imposing lifecycle data requirements ([4] www.fda.gov) and exploring adaptive approval pathways ([7] pmc.ncbi.nlm.nih.gov), the EU is integrating AI oversight into future legislation ([5] goodlifesci.sidley.com), and harmonization efforts are underway.

Key recommendations from this analysis include:

- **Implement comprehensive monitoring plans:** Every AI device should have a documented PMS plan tailored to its update model (continuous vs locked). This plan should specify performance metrics, update triggers, and data sources (e.g. real-world evidence, user logs, registries). For continuous systems, special emphasis is needed on version control and retraining oversight ([8] www.ncbi.nlm.nih.gov).
- **Engage clinicians in surveillance:** As the European radiology community emphasizes, healthcare professionals must be aware of regulations and actively report suspected issues ([20] pmc.ncbi.nlm.nih.gov) ([27] pmc.ncbi.nlm.nih.gov). Training and education on AI device use/PMS is crucial.

© 2026 IntuitionLabs.ai - North America's Leading AI Software Development Firm for Pharmaceutical & Biotech. All rights reserved.

- **Leverage real-world data and AI:** Regulators and manufacturers should use advanced analytics (including AI tools) to process the vast data from device usage for signals ([21] jamanetwork.com). Collaboration on shared real-world datasets could accelerate detection of common issues.
- **Balance innovation with caution:** Policymakers should support adaptive AI but require it to follow pre-approved protocols. Emergency uses of AI (e.g. pandemic response) may need special provisions for rapid updates, but safety oversight cannot be neglected.

Looking to the future, AI in healthcare will only deepen. As models become more powerful (e.g. large multimodal models for medical text and images), the PMS landscape must mature accordingly. We anticipate increased integration of AI surveillance into routine device management. Ultimately, the goal stays constant: ensure that these powerful technologies **do no harm** and improve patient outcomes over the long term. Achieving this will require continuing multi-stakeholder collaboration, investment in evidence generation, and robust regulatory-science frameworks.

Sources: This report draws on numerous expert sources and data. Key references include World Health Organization guidelines (www.who.int), FDA regulatory announcements ([4] www.fda.gov), IMDRF definitions ([1] innolitics.com) ([2] innolitics.com), scientific studies on AI performance ([13] pmc.ncbi.nlm.nih.gov) ([15] pmc.ncbi.nlm.nih.gov), and regulatory analyses ([7] pmc.ncbi.nlm.nih.gov) ([42] pmc.ncbi.nlm.nih.gov). All statements are supported by the cited literature. Each source is indicated inline for verification.

External Sources

[1] https://innolitics.com/articles/imdrf-guidance-machine-learning-enabled-medical-devices-key-terms-and-definitions/#:~:Aspec...

[2] https://innolitics.com/articles/imdrf-guidance-machine-learning-enabled-medical-devices-key-terms-and-definitions/#:~:6...

[3] https://www.healthcaredive.com/news/fda-ai-medical-devices-growth/729351/#:~:The%2...

[4] https://www.fda.gov/news-events/press-announcements/fda-issues-comprehensive-draft-guidance-developers-artificial-intelligence-enabled-medical-devices#:~:The%2...

[5] https://goodlifesci.sidley.com/2023/10/16/europes-regulatory-approach-to-continuous-learning-ai/#:~:The%2...

[6] https://pmc.ncbi.nlm.nih.gov/articles/PMC11634576/#:~:The%2...

[7] https://pmc.ncbi.nlm.nih.gov/articles/PMC11634576/#:~:the%2...

[8] https://www.ncbi.nlm.nih.gov/books/NBK605105#:~:regul...

[9] https://www.ncbi.nlm.nih.gov/books/NBK605105#:~:The%2...

[10] https://www.ncbi.nlm.nih.gov/books/NBK605105#:~:model...

[11] https://pmc.ncbi.nlm.nih.gov/articles/PMC12380398/#:~:postd...

[12] https://www.ncbi.nlm.nih.gov/books/NBK605105#:~:At%20...

[13] https://pmc.ncbi.nlm.nih.gov/articles/PMC12864748/#:~:analy...

[14] https://www.fda.gov/medical-devices/device-advice-comprehensive-regulatory-assistance/postmarket-requirements-devices#:~:Medic...

[15] https://pmc.ncbi.nlm.nih.gov/articles/PMC9367531/#:~:a%20d...

[16] https://pmc.ncbi.nlm.nih.gov/articles/PMC9994700/#:~:match...

[17] https://www.healthcaredive.com/news/fda-ai-medical-devices-growth/729351/#:~:In%20...

© 2026 IntuitionLabs.ai - North America's Leading AI Software Development Firm for Pharmaceutical & Biotech. All rights reserved.

[18] https://pmc.ncbi.nlm.nih.gov/articles/PMC11634576/#:~:Figur...

[19] https://pmc.ncbi.nlm.nih.gov/articles/PMC12380398/#:~:strat...

[20] https://pmc.ncbi.nlm.nih.gov/articles/PMC12701188/#:~:The%2...

[21] https://jamanetwork.com/journals/jamanetworkopen/fullarticle/2822418#:~:Obser...

[22] https://innolitics.com/articles/imdrf-guidance-machine-learning-enabled-medical-devices-key-terms-and-definitions/#:~:Note%...

[23] https://innolitics.com/articles/imdrf-guidance-machine-learning-enabled-medical-devices-key-terms-and-definitions/#:~:Examp...

[24] https://innolitics.com/articles/imdrf-guidance-machine-learning-enabled-medical-devices-key-terms-and-definitions/#:~:Note%...

[25] https://pmc.ncbi.nlm.nih.gov/articles/PMC11634576/#:~:subst...

[26] https://pmc.ncbi.nlm.nih.gov/articles/PMC11634576/#:~:the%2...

[27] https://pmc.ncbi.nlm.nih.gov/articles/PMC12701188/#:~:risk%...

[28] https://pmc.ncbi.nlm.nih.gov/articles/PMC11634576/#:~:The%2...

[29] https://pmc.ncbi.nlm.nih.gov/articles/PMC12701188/#:~:inclu...

[30] https://www.ncbi.nlm.nih.gov/books/NBK605105#:~:,regu...

[31] https://www.ncbi.nlm.nih.gov/books/NBK605105#:~:risks...

[32] https://innolitics.com/articles/imdrf-guidance-machine-learning-enabled-medical-devices-key-terms-and-definitions/#:~:Note%...

[33] https://pmc.ncbi.nlm.nih.gov/articles/PMC9994700/#:~:Abstr...

[34] https://pmc.ncbi.nlm.nih.gov/articles/PMC9994700/#:~:AI%2F...

[35] https://pmc.ncbi.nlm.nih.gov/articles/PMC9994700/#:~:The%2...

[36] https://pmc.ncbi.nlm.nih.gov/articles/PMC9367531/#:~:a%20d...

[37] https://pmc.ncbi.nlm.nih.gov/articles/PMC12701188/#:~:showe...

[38] https://pmc.ncbi.nlm.nih.gov/articles/PMC11634576/#:~:effic...

[39] https://pmc.ncbi.nlm.nih.gov/articles/PMC9367531/#:~:a%20d...

[40] https://pmc.ncbi.nlm.nih.gov/articles/PMC9367531/#:~:class...

[41] https://www.imdrf.org/consultations/good-machine-learning-practice-medical-device-development-guiding-principles#:~:Consu...

[42] https://pmc.ncbi.nlm.nih.gov/articles/PMC11634576/#:~:medic...

© 2026 IntuitionLabs.ai - North America's Leading AI Software Development Firm for Pharmaceutical & Biotech. All rights reserved.

IntuitionLabs - Industry Leadership & Services

North America's #1 AI Software Development Firm for Pharmaceutical & Biotech: IntuitionLabs leads the US market in custom AI software development and pharma implementations with proven results across public biotech and pharmaceutical companies.

Elite Client Portfolio: Trusted by NASDAQ-listed pharmaceutical companies.

Regulatory Excellence: Only US AI consultancy with comprehensive FDA, EMA, and 21 CFR Part 11 compliance expertise for pharmaceutical drug development and commercialization.

Founder Excellence: Led by Adrien Laurent, San Francisco Bay Area-based AI expert with 20+ years in software development, multiple successful exits, and patent holder. Recognized as one of the top AI experts in the USA.

Custom AI Software Development: Build tailored pharmaceutical AI applications, custom CRMs, chatbots, and ERP systems with advanced analytics and regulatory compliance capabilities.

Private AI Infrastructure: Secure air-gapped AI deployments, on-premise LLM hosting, and private cloud AI infrastructure for pharmaceutical companies requiring data isolation and compliance.

Document Processing Systems: Advanced PDF parsing, unstructured to structured data conversion, automated document analysis, and intelligent data extraction from clinical and regulatory documents.

Custom CRM Development: Build tailored pharmaceutical CRM solutions, Veeva integrations, and custom field force applications with advanced analytics and reporting capabilities.

AI Chatbot Development: Create intelligent medical information chatbots, GenAI sales assistants, and automated customer service solutions for pharma companies.

Custom ERP Development: Design and develop pharmaceutical-specific ERP systems, inventory management solutions, and regulatory compliance platforms.

Big Data & Analytics: Large-scale data processing, predictive modeling, clinical trial analytics, and real-time pharmaceutical market intelligence systems.

Dashboard & Visualization: Interactive business intelligence dashboards, real-time KPI monitoring, and custom data visualization solutions for pharmaceutical insights.

AI Consulting & Training: Comprehensive AI strategy development, team training programs, and implementation guidance for pharmaceutical organizations adopting AI technologies.

Contact founder Adrien Laurent and team at https://intuitionlabs.ai/contact for a consultation.

© 2026 IntuitionLabs.ai - North America's Leading AI Software Development Firm for Pharmaceutical & Biotech. All rights reserved.

DISCLAIMER

The information contained in this document is provided for educational and informational purposes only. We make no representations or warranties of any kind, express or implied, about the completeness, accuracy, reliability, suitability, or availability of the information contained herein.

Any reliance you place on such information is strictly at your own risk. In no event will IntuitionLabs.ai or its representatives be liable for any loss or damage including without limitation, indirect or consequential loss or damage, or any loss or damage whatsoever arising from the use of information presented in this document.

This document may contain content generated with the assistance of artificial intelligence technologies. AI-generated content may contain errors, omissions, or inaccuracies. Readers are advised to independently verify any critical information before acting upon it.

All product names, logos, brands, trademarks, and registered trademarks mentioned in this document are the property of their respective owners. All company, product, and service names used in this document are for identification purposes only. Use of these names, logos, trademarks, and brands does not imply endorsement by the respective trademark holders.

IntuitionLabs.ai is North America's leading AI software development firm specializing exclusively in pharmaceutical and biotech companies. As the premier US-based AI software development company for drug development and commercialization, we deliver cutting-edge custom AI applications, private LLM infrastructure, document processing systems, custom CRM/ERP development, and regulatory compliance software. Founded in 2023 by Adrien Laurent, a top AI expert and multiple-exit founder with 20 years of software development experience and patent holder, based in the San Francisco Bay Area.

This document does not constitute professional or legal advice. For specific guidance related to your business needs, please consult with appropriate qualified professionals.

© 2025 IntuitionLabs.ai. All rights reserved.