10.48047/jocaaa.2026.35.03.18

Continuous Postmarket Safety Intelligence: Real-Time Signal Detection from Device Telemetry and Clinical Workflow Data

Shrikant Chikhalkar
Independent Researcher, USA

Abstract

Postmarket safety monitoring of connected medical devices will need to find important signals in large amounts of data from telemetry, operations logs, and clinical workflow events because traditional complaint-based methods for monitoring software-driven devices are not good at spotting new dangers before they harm patients. To address these challenges in regulated environments with non-negotiable requirements for defensibility, reproducibility, and governance alignment, organizations propose a systems-level architecture for continuous postmarket safety intelligence. Such an architecture would include risk-informed signal definitions, privacy-preserving analytic pipelines; various signal detection methods, including rule-based checks, statistical process control and anomaly detection; and human-in-the-loop triage workflows. Escalation pathways map confirmed signals to explicit corrective and preventive action processes, while evidence packages provide auditable records of regulatory dialogue. Security controls that reduce adversarial threats, such as spoofed telemetry and model poisoning, and usability-centered reviewer interfaces and difference-in-differences validation frameworks will help make platform-level performance trackable and improvable, transforming postmarket monitoring from reactive to continuous safety intelligence across the connected-device ecosystem.

Keywords: Postmarket Surveillance, Telemetry Analytics, Anomaly Detection, Signal Governance, Regulatory Defensibility

1. Introduction to Postmarket Safety Challenges in Connected Medical Devices

Connected medical devices are one of the fastest-growing sectors of the healthcare technology ecosystem; they continuously stream operational telemetry in the form of error logs, usage patterns, network connectivity events, performance counters, and clinical workflow signals. The volume, velocity and variety of these data streams represent an opportunity for postmarket surveillance programs. Customary passive sources of postmarket safety information, such as adverse event reporting, complaint handling, and periodic literature reviews, are insufficient to proactively detect new risks posed by software-enabled device ecosystems before harm occurs to patients. This is especially true for the clinics. Even in specialties that have been the focus of intensive guidelines research (such as cardiology), only 19 percent of guidelines are based on randomized controlled trial evidence, meaning most clinical decisions (and thus device-dependent practices) may be based on evidence that lacks generalizability to patients [1].

The basic problem is one of signal-to-noise ratio and clinical relevance: devices in common use in diverse medical, specialty, and hospital-based care settings generate huge volumes of raw data every day, but there are very few different events that clinicians need to know about as potential precursors to a safety-critical failure. Data quality issues in health information systems on which postmarket surveillance is based present another barrier. A systematic review of 227 peer-reviewed studies of digital health data quality summarized six dimensions in the health context (accessibility, accuracy, completeness, consistency, contextual validity, and currency). Consistency stood out as the data quality dimension that

10.48047/jocaaa.2026.35.03.18

most influenced the other data quality dimensions, as well as all five health outcome categories: clinical, clinician, research, business process, and organizational outcomes [2]. Of the above, consistency and completeness were the most common, reported by 72.2% and 60.4% of studies reviewed in [2] respectively.

Automated analysis architectures that take into account these data quality shortcomings are still lacking, which means that safety teams are often overwhelmed with raw telemetry data and have no methodologies to discriminate in a systematic manner instrumentation artifacts from actual changes in behavior, as well as no way to connect anomalies detected with corrective and preventive action workflows.

This article proposes a systems-level architecture for continuous postmarket safety intelligence in regulated environments, including risk-informed signal definitions, privacy-preserving analytic pipelines, human-in-the-loop triage and interpretation of signals, and an auditable chain of evidence leading to action.

2. Data Sources, Signal Semantics, and Harmonization Frameworks

High-quality, complete, semantically consistent data is critical for postmarket safety intelligence. Connected medical devices produce several classes of operationally relevant data. For example, there are structured device logs with state changes, alarms, and warnings; unstructured crash and diagnostic dumps; telemetry on device usage; sequences of screen flows and feature activations; performance metrics; connectivity events; pairing events; and acquisition pipeline health metrics. The importance of this data infrastructure is supported by empirical results from analyzing high-risk device postmarket incidents published in 2024: hardware/mechanical failures are the most frequent cause of incident (25), and field modifications are the most common response (46% of all corrective actions required). Software updates (26%) and recalls (22%) are the second and third most common responses, respectively. The implications are that how well a failure mode can be detected and characterized affects how appropriate the response is likely to be [3].

Another architectural requirement is separating patient-identifying information from the rest of the operational components so that the information needed for operational analysis at the population level can be brought together without exposing protected health information. This is part of a privacy-by-design approach that can be considered integral to applicable regulatory frameworks.

Telemetry data is not semantically consistent across generations of devices, software versions, and deployment scenarios. Without a canonical event taxonomy that includes versioning and backward compatibility, telemetry analytic pipelines are likely to detect false positives due to instrumentation regressions, meaning they may pick up noise in telemetry data from instrumentation instead of actual changes in device behavior. A data contract layer comprises schema validation of incoming events, validation of required fields, rejection of events that are ill-formed or out of sequence, recording instrumentation version metadata, and generating summary quality reports indicating missingness rates, timestamp skews, sampling drifts, and duplication ratios. Signal fidelity is clinically relevant, particularly in a cardiac monitoring context, where CAD systems are filtering ECG signals that may be anomalous for a variety of reasons, including baseline wander, power line interference at known frequencies and harmonics (50/60 Hz), muscle artifacts and electrode contact noise. Classification accuracy has been shown to decrease in the presence of natural noise processes. The literature notes the explicit need for systematic analyses of linear and nonlinear approaches in this area of research [4].

10.48047/jocaaa.2026.35.03.18

Also, heterogeneous device fleets, multiple versions of an application, or low-connectivity environments where devices batch telemetry for long periods need the pipeline to be able to handle a variety of ingestion scenarios without corrupting rolling-window feature computations with out-of-order event sequences. If the harmonization layer is poorly designed, the signal detection efforts that rely on this data may lack reliability and regulatory defensibility.

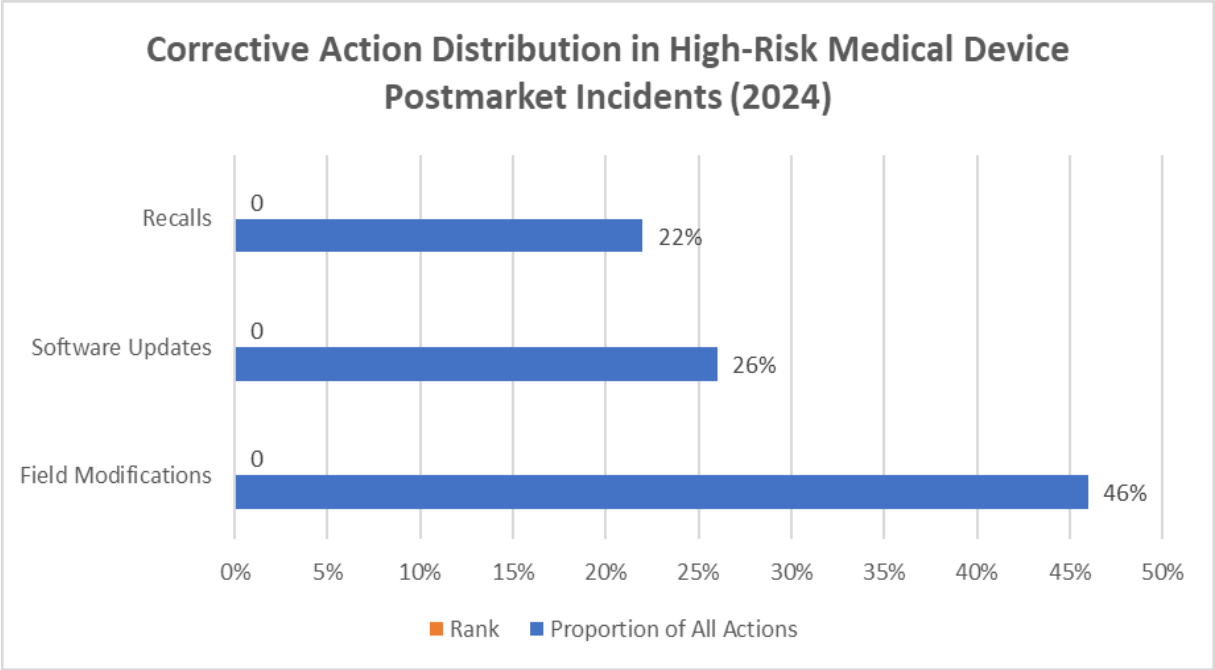

Fig 1: Proportion of Corrective Action Types Across High-Risk Device Postmarket Surveillance Findings [3, 4]

3. Signal Detection Architecture and Feature Extraction Methods

Signal detection architecture separates the steps of data intake, standardization, feature extraction, and feature detection into different processes to make it easier to test each part, manage versions independently, and keep data preparation separate from decision-making.

Telemetry producers are verified at the ingestion stage using WPA2-Enterprise (802.11i) and 802.1x authentication methods that All cryptographic modules are required to be FIPS PUB 140-2 compliant [6]. Ingestion also uses rate-limiting methods to guard against sudden flood attacks and fake high-volume data streams, and it stores events to get ready for later processing. Normalization maps different incoming events to the canonical taxonomy of the harmonization framework. Feature extraction is done via rolling window aggregates, session aggregates, and cohort stratifications. These cohort stratifications are important in healthcare because WLAN enterprise designs need to be checked to ensure they meet a minimum signal strength of negative (-) 65 dBm and have cochannel interference (CCI) below 2%. Thus, the same absolute error rate might be clinically insignificant for one device configuration or care setting but might be strongly meaningful in another.

Detection approaches would ideally combine these four to seek to cover the hazard space. Rule-based checks cover known, well-understood hazards using deterministic detection logic. Statistical process control methods like CUSUM charts and EWMA charts are used to spot changes in how often an event happens or in the performance measurements. Anomaly detection methods try to find new attempts to

10.48047/jocaaa.2026.35.03.18

insert data that are different from the usual patterns seen in the training data. When attacked, only 30-40% of the training labels need to be inverted [5] for the classifier's inference accuracy to plummet. Change-point detection algorithms are often applied to real-time data streams to detect process changes, including software regressions, hardware wear-out, and environmental changes.

One often-missed requirement for an early telemetry analytics effort is triage. Each alert should come with a clear explanation that details what changed, how much it changed, which group it belongs to, how confident we are about the detected issue, and links to example event traces, statistical charts, and query definitions. Black-box adversarial attacks show that remotely hosted models can misclassify inputs at an 84.24% success rate without knowledge of model architecture or parameters. Because alarm results are impervious to justification, human reviewers cannot distinguish between legitimate anomalies and adversarial inputs, leading to extra load on human reviewers and alert fatigue, regardless of accuracy.

The pipeline must support near real-time detection of highly severe threats and batch detection for slower-moving trends. High-severity pathways need a specific time limit set as a separate requirement, and this requirement should be checked through complete performance testing of the system. WLAN modeling tools used for coverage validation should be accurate to within 98-100% [6].

A. Authentication and Cryptographic Standards		
Parameter	**Value / Standard**	**Context**
Authentication Protocol	**WPA2-Enterprise (802.11i)**	Telemetry producer authentication at ingestion layer
Extensible Authentication Protocol	**802.1x**	Mutual authentication for ingestion layer producers
Cryptographic Module Compliance	**FIPS PUB 140-2**	Required standard for all cryptographic modules
B. WLAN Signal and Validation Thresholds		
Parameter	**Value**	**Context**
Minimum WLAN Signal Threshold	**Negative 65 dBm**	Required floor for WLAN enterprise design validation
Maximum Cochannel Interference (CCI)	**< 2%**	Validated CCI ceiling for enterprise WLAN deployments
WLAN Coverage Validation Tool Accuracy	**98–100%**	Target accuracy range for predictive WLAN modeling tools
C. Adversarial Attack Performance Benchmarks		
Parameter	**Value**	**Context**
Training Label Inversion Threshold	**30–40%**	Proportion of training labels that must be inverted to substantially degrade classifier inference accuracy
Black-Box Adversarial Misclassification Rate	**84.24%**	Success rate of black-box adversarial attacks against remotely hosted models without model architecture knowledge

10.48047/jocaaa.2026.35.03.18

Table 1: Key Numerical Parameters in Signal Detection Architecture [5, 6]

4. Safety Governance, Human-in-the-Loop Workflows, and Escalation Pathways

Unlike confirmed safety signals, automated signal detection refers to signals that are subject to follow-up. The organization's signal governance process should explain the responsibilities of signal reviewers, the timelines for escalating and sorting signals, the requirements for documenting signals, and how signal detection connects to quality management systems (QMS) to help with proper follow-up and decision-making for corrective actions.

Routing logic should filter the alerts and send them into the desired reviewer queue based on the signal type and severity. Software quality reviewers are best suited for exception rate, memory usage, or connectivity degradation alerts. Alerts are routed to clinical safety reviewers if they relate to patterns of therapy interruptions, abnormal state transitions, or device behaviors associated with adverse events. This routing architecture corresponds to the IMDRF SaMD clinical evaluation framework. It defines the analytical validation of SaMD as the ability to confirm that the SaMD correctly processes input data to generate accurate, reliable and precise output, and the clinical validation of a SaMD as the ability to determine whether the use of the output achieves the intended purpose in the target population in the context of clinical care [7]. For example, signals implicating output accuracy belong to the former category, while messages implicating patient outcome belong to the latter and would require a higher proportion of clinical reviewer involvement.

Escalation policies must take into account the IMDRF SaMD N12 risk categorization scheme. According to this scheme, SaMD that treat or diagnose critical medical conditions are severity category IV devices, while SaMD that inform the instruction of clinical management of patients with non-serious medical conditions are severity category I devices (see above). Experts who do not have any conflict of interest should independently review high-severity signals, i.e., those with plausible links to patient harm in critical or serious healthcare situations, and these signals should be triaged in rapid review cycles. Lower-severity signals can be triaged on longer time frames.

The platform should enable documented links to investigation artifacts when a signal is deemed a safety issue. These artifacts include things like root-cause analysis, temporary fixes, formal actions to correct and prevent issues, proof of verification, and approvals needed for regulatory defense during inspections after a product is on the market.

Alert fatigue management is another governance issue. Sim et al. state that ML algorithms are probabilistic procedures in which users should expect certain percentages of error; consumers should not insist on certainties that are unwarranted [8]. Since the reviewer can only indicate acceptable error rates and knows that machines have difficulty when the data changes (meaning that tools that work well on one group may not work on another), the feedback from reviewers should adjust the thresholds and include a clear process for tracking improvements by recording how each case was handled—whether it was confirmed, dismissed, or needed more information—along with the reasons for those decisions.

Category	Key Components	Description
Signal Governance Framework	Roles, Timelines, Documentation, QMS Integration	Defines reviewer responsibilities, escalation timelines, documentation standards, and linkage to the QMS for corrective action support
Alert Routing Logic	Severity-Based Filtering,	Routes alerts to software quality or clinical

10.48047/jocaaa.2026.35.03.18

	Reviewer Queue Assignment	safety reviewers based on signal type and severity, aligned with IMDRF analytical and clinical validation categories
Escalation Policies	IMDRF N12 Risk Categories (I–IV)	High-severity signals require rapid, independent expert review; lower-severity signals allow longer triage timeframes
Investigation Artifact Linkage	Root-Cause Analysis, CAPAs, Verification Records	Confirmed safety signals are linked to investigation artifacts including root-cause analyses, corrective actions, and regulatory approval documentation
Alert Fatigue Management	Threshold Adjustment, Disposition Tracking	Reviewer feedback recalibrates signal thresholds to account for ML error rates and data drift; all dispositions (confirmed, dismissed, pending) are recorded with rationale

Table 2: Safety Signal Governance, Triage Workflows, and Escalation Pathways in Postmarket SaMD Surveillance [7, 8]

5. Privacy, Security, and Regulatory Compliance by Design

Privacy and security considerations are not simply a matter of an "opt in" or "opt out" for safety intelligence systems used after market release but are fundamental architectural requirements from the earliest stages of design and implementation for the protection of patients and avoidance of regulatory liability.

Privacy by design for the telemetry analytics of medical devices requires systematic attention to data minimization. Collect telemetry only for the operational features necessary to detect safety signals. Tokenize or pseudonymize device and session identifiers before ingesting data into analytic environments. Aggregate telemetry at the level of population cohorts on dashboards and reports rather than individual-level population profiles. The demographic design of these cohorts is illustrated by the 819 analyzable participants in the Abràmoff et al. pivotal trial, which was 28.6% African American, 16.1% Hispanic, and 63.4% non-Hispanic [9]. Analytic generalizability would have been systematically compromised if safety analytics had been performed using only narrower cohorts. Differential privacy methods used in the reporting sections ensure that there are mathematical protections against identifying the safety signals that have been reported.

When implementing security controls, the entire threat surface needs to be considered. Mutual authentication between the telemetry producers and the ingestion endpoints can help keep attackers from sending fake telemetry data. Encryption in transit and at rest protects data from interception, while tamper-clear logs built on cryptographic chains offer integrity guarantees for audit trails. This division of responsibilities reduces the risk of insider threats by ensuring that no one person can change detection rules or review records without someone else checking their work.

Threat modeling is not limited to data breaches; spoofed telemetry in a control system, for example, could stop a safety alert from triggering or could flood the operator with false alerts. Worst-case scenarios are discussed in the above-mentioned paper from Abràmoff et al. (Abràmoff et al. 2021) (and summarized in the reference). For example, forcing all 40 participants who could not be analyzed to a grading opposed to the predicted AI system outcome caused a reduction in sensitivity from 87.2% to 80.7% and specificity

10.48047/jocaaa.2026.35.03.18

from 90.7% to 89.8% [9]. Such input poisoning could be applied against machine learning-based anomaly detectors by shifting the learned boundary of the normal data likewise. Logging and independent review of privileged access may reduce this.

Regulatory requirements include documentation of postmarket surveillance, complaint handling, and corrective and preventive actions required by quality management systems for medical devices. Cybersecurity guidance covers software and data infrastructure that support safety functions. Platform designs should reflect risk assessments, control mappings, verification materials, and, in keeping with the locked-algorithm, independent-monitoring, and escrowed-approach integrity controls of Abràmoff et al. pivotal 900-subject, 10-site primary care study, be monitored by an independent CRO [9].

6. Verification, Validation, and Evidence Generation for Regulatory Defensibility

Signals in a postmarket safety intelligence platform should be treated like features in a designed system and must go through formal checks to ensure they meet regulatory standards and can be improved over time. Each signal's description must include what it is,

Each signal's specification must include its definition, the hazard or fault mode it detects, the inputs and logic used to calculate it, the thresholds and reasons for those thresholds, and the expected sensitivity and specificity in the operational environment. Verification activities involve running test scenarios against each signal definition, checking both positive and negative cases, using past incident data to confirm that the signal detects the incident, and simulating faults to see how the signal reacts as if it were another device.

An important measurement to evaluate the overall benefit of the platform is whether the alerting system considerably improves the time-to-detection and time-to-containment of actual safety problems. A proper evaluation could be performed with a difference-in-differences (DID) method that compares the outcome over time of a treated group and a parallel control group over an identical period of observation. Like policy evaluation, a DD can control for baseline differences and secular trends that confound before-after comparisons. For instance, this can be used to assess whether device cohorts with different baseline complaint rates improved detection latency [10]. Validation designs in this category compare when alerts are detected to when similar complaints were reported in the past, eliminating any increases from the original platform.

Evidence packages that assist with regulatory talks and internal reviews include explanations of signals, changes made over time, results from testing, data on how things perform in real life, and notes from reviewers showing what steps were taken for each signal being looked at. When regulatory questions arise about whether detection was adequate or response timely, packages must provide traceable responses to these questions.

Human-in-the-loop triage effectiveness can be improved by using reviewer interfaces that make it easier for humans to work, give feedback, and encourage consistent actions from reviewers. Such interfaces help ensure reviewer dispositions are accurate and repeatable across varying alert volumes [11]. As more connected devices are added, creating easy-to-use measures and interfaces for assessing them will be essential for creating reliable, checkable, and upgradable surveillance systems in the future.

10.48047/jocaaa.2026.35.03.18

Category	Component/Activity	Description
Signal Specification Requirements	Signal Definition	Documented description of each signal, including the hazard or fault mode it detects
	Input & Logic	Inputs and calculation logic used to compute the signal
	Thresholds	Defined thresholds with documented rationale
	Performance Targets	Expected sensitivity and specificity in the operational environment
Verification Activities	Synthetic Test Streams	Execute synthetic data against signal definitions using positive and negative test cases
	Historical Incident Replay	Run historical incident data against signal definitions to confirm retrospective detection
	Fault Injection	Simulate device faults to mimic behavior of other devices and test signal response
Validation & Performance Evaluation	Time-to-Detection	Measure whether the alerting platform reduces time-to-detection of real safety problems
	Time-to-Containment	Measure whether the platform reduces time-to-containment of safety events
	Evaluation Method	Difference-in-Differences (DID) comparing treated vs. control device cohorts over identical observation periods
	Confound Control	DID controls for baseline differences and secular trends that confound before-after comparisons
	Benchmark Alignment	Align alert detection timelines with historical complaint-based detection timelines for the same event class
Evidence Packages	Signal Definitions & Revision Histories	Versioned documentation of each signal specification over time
	Verification Test Results	Documented outcomes of all verification test executions
	Real-World Performance Data	Operational data demonstrating signal performance in deployment
	Disposition Records	Reviewer action documentation for each signal under review, supporting audit traceability
Human-in-the-Loop Triage	Reviewer Interface Design	Interfaces that reduce cognitive load, provide feedback, and promote consistent reviewer behavior

10.48047/jocaaa.2026.35.03.18

	Reviewer Accuracy & Repeatability	Ensuring dispositions are accurate and consistent across varying alert volumes
	Scalability	Metrics and usable interfaces designed to scale with expanding connected device ecosystems
Regulatory Defensibility	Traceable Evidence	Packages must provide traceable responses to regulatory questions about detection adequacy and response timeliness
	Auditability	Systems must be defensible, auditable, and upgradable over time

Table 3: Signal Verification, Validation, and Evidence Generation Requirements for Regulatory Defensibility in Postmarket Safety Intelligence Platforms [10, 11]

Conclusion

Continuous postmarket safety intelligence is the next step in regulated industry practices for safely operating connected medical devices in increasingly complex, heterogeneous deployment environments. By considering telemetry signals to be formally specified, hazard-linked engineered features rather than heuristic guidelines, organizations can build defensible detection capabilities that are resilient to regulatory scrutiny and that can be progressively improved over time. Data quality governance, layered detection logic, privacy-preserving pipeline design, and human-in-the-loop escalation combine into a tightly integrated platform that provides end-to-end coverage from alert to validated remediation. Difference-in-differences evaluation and human-centered design give quantifiable detection performance and reliable reviewer workflows for production deployments. As this ecosystem of connected devices evolves with increasing velocity and breadth, the auditability, clinical relevance, and governance alignment of the architecture will be key to developing and sustaining postmarket safety intelligence.

References

[1] Christopher A et al., "A 'green button' for using aggregate patient data at the point of care," ResearchGate, 2014. [Online]. Available: https://www.researchgate.net/profile/Christopher-Longhurst-2/publication/263778510

[2] Rehan Syed et al., "Digital Health Data Quality Issues: Systematic Review," JMIR Publications, 2023. [Online]. Available: https://www.jmir.org/2023/1/e42615/

[3] Oviyaasri Manimuthukumar et al., "Global trends in post-market surveillance of high-risk medical devices: An empirical analysis based on regulatory data," Indian Journal of Medical Research, 2025. [Online]. Available: https://pmc.ncbi.nlm.nih.gov/articles/PMC12624750/

[4] Roshan Joy Martis et al., "Current methods in electrocardiogram characterization," Computers in Biology and Medicine, Volume 48, 2014. [Online]. Available: https://www.sciencedirect.com/science/article/abs/pii/S0010482514000432

[5] Nicolas Papernot et al., "SoK: Towards the Science of Security and Privacy in Machine Learning," arXiv:1611.03814v1, 2016. [Online]. Available: https://arxiv.org/pdf/1611.03814

[6] David H.Hoglund, "Secure and Reliable Medical Device and Mobile Connectivity," AAMI, 2017. [Online]. Available: https://array.aami.org/doi/pdf/10.2345/0899-8205-51.2.130

10.48047/jocaaa.2026.35.03.18

[7] IMDRF, "Software as a Medical Device (SaMD): Clinical Evaluation," 2017. [Online]. Available: hhttps://www.imdrf.org/sites/default/files/docs/imdrf/final/technical/imdrf-tech-170921-samd-n41-clinical-evaluation_1.pdf

[8] Jordan Zheng Ting Sim et al., "Machine learning in medicine: what clinicians should know," National Library of Medicine, 2021. [Online]. Available: https://pmc.ncbi.nlm.nih.gov/articles/PMC10071847/

[9] Michael D. Abràmoff et al., "Pivotal trial of an autonomous AI-based diagnostic system for detection of diabetic retinopathy in primary care offices," npj, 2018. [Online]. Available: https://www.nature.com/articles/s41746-018-0040-6.pdf

[10] Justin B. Dimick and Andrew M. Ryan, "Methods for evaluating changes in health care policy: The difference-in-differences approach," JAMA, 2014. [Online]. Available: https://scholarblogs.emory.edu/mscr/files/2014/12/Difference-in-differences-approach.pdf

[11] Ben Shneiderman and Catherine Plaisant, "Designing the User Interface: Strategies for Effective Human-Computer Interaction." [Online]. Available: http://seu1.org/files/level5/IT201/Book%20-%20Ben%20Shneiderman-Designing%20the%20User%20Interface-4th%20Edition.pdf